

Variation of Positive Predictive Values of Fecal Immunochemical Tests by Polygenic Risk Score in a Large Screening Cohort

Tobias Niedermaier, PhD, MPH¹, Yesilda Balavarca, PhD², Anton Gies, RPh, PhD², Korbinian Weigl, MPH, PhD¹, Feng Guo, MSc, PhD¹, Elizabeth Alwers, MD, PhD¹, Michael Hoffmeister, PhD, MPH¹ and Hermann Brenner, MD, MPH, PhD^{1,2,3}

INTRODUCTION: Prevalence of colorectal neoplasms varies by polygenic risk scores (PRS). We aimed to assess to what extent a PRS might be relevant for defining personalized cutoff values for fecal immunochemical tests (FITs) in colorectal cancer screening.

METHODS: Among 5,306 participants of screening colonoscopy who provided a stool sample for a quantitative FIT (Ridascreen Hemoglobin or FOB Gold) before colonoscopy, a PRS was determined, based on the number of risk alleles in 140 single nucleotide polymorphisms. Subjects were classified into low, medium, and high genetic risk of colorectal neoplasms according to PRS tertiles. We calculated positive predictive values (PPVs) and numbers needed to scope (NNS) to detect 1 advanced neoplasm (AN) by the risk group, and cutoff variation needed to achieve comparable PPVs across risk groups in the samples tested with Ridascreen (N = 1,271) and FOB Gold (N = 4,035) independently, using cutoffs yielding 85%, 90%, or 95% specificity.

RESULTS: Performance of both FITs was very similar within each PRS group. For a given cutoff, PPVs were consistently higher by 11%–15% units in the high-risk PRS group compared with the low-risk group (all *P* values < 0.05). Correspondingly, NNS to detect 1 advanced neoplasm varied from 2 (high PRS, high cutoff) to 5 (low PRS, low cutoff). Conversely, very different FIT cutoffs would be needed to ensure comparable PPVs across PRS groups.

DISCUSSION: PPVs and NNS of FITs varied widely across people with high and low genetic risk score. Further research should evaluate the relevance of these differences for personalized colorectal cancer screening.

SUPPLEMENTARY MATERIAL accompanies this paper at <http://links.lww.com/CTG/A754>.

Clinical and Translational Gastroenterology 2022;13:e00458. <https://doi.org/10.14309/ctg.0000000000000458>

INTRODUCTION

Fecal immunochemical tests (FITs) are widely recommended by expert panels (1–4) and increasingly used for colorectal cancer (CRC) screening (5). Although cutoffs vary between tests and screening programs, typically, the same cutoffs are used for the entire screening population within screening programs. However, if prevalence of neoplasms varies between subgroups of the screening population, positive predictive values (PPVs) are expected to likewise vary between those subgroups. Correspondingly, also their reciprocal values, numbers needed to undergo colonoscopy after a positive FIT (number needed to scope [NNS]) to detect 1 neoplasm would vary. In particular, for a uniform cutoff, PPVs are expected to be higher for high-risk

groups than for low-risk groups (unless sensitivity and specificity also vary substantially between high- and low-risk groups). From both an individual patient and public health perspective, it seems preferable to use cutoffs that yield the same PPV in all subgroups to avoid follow-up colonoscopies with very low yield of advanced neoplasms in the low-risk groups and not to miss follow-up colonoscopies with relatively high yield in the high-risk groups.

In recent years, genomewide association studies have identified a rapidly increasing number of single nucleotide polymorphisms (SNPs) that are associated with increased risk of CRC (6–15). Although risk information conveyed by individual SNPs is small, combining multiple SNPs in polygenic risk scores (PRS) has been shown to enable risk stratification with major variation

¹Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; ²Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany; ³German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany. **Correspondence:** Tobias Niedermaier, PhD, MPH. E-mail: t.niedermaier@dkfz.de

Received December 6, 2021; accepted December 28, 2021; published online January 19, 2022

© 2022 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of The American College of Gastroenterology

of CRC risk between subjects with high and low PRS (16). It has been shown that such PRS also enable relevant risk stratification by prevalence of advanced neoplasms among participants of screening colonoscopy (17), suggesting that PRS could be useful for risk-adapted CRC screening. The aim of this study was to evaluate to what extent PPVs and NNS for detecting advanced neoplasms are expected to vary according to PRS-defined risk groups if the same FIT cutoff is used across all risk groups, and how much FIT cutoffs would have to differ to achieve comparable PPVs and NNS across PRS-defined risk groups.

METHODS

Study design and population

All analyses are based on data from the BliTz study, which has been described previously elsewhere (17–20). In brief, BliTz is an ongoing prospective CRC screening study conducted in Southern Germany. Participants of screening colonoscopy aged 50–79 years, representing the target population for CRC screening, are recruited at a visit at 1 of 20 participating gastroenterology practices before screening colonoscopy. Stool and blood samples are collected before bowel preparation. Participants selected for the current analysis were recruited between November 2005 and January 2019.

Participants were asked to fill out a self-administered questionnaire on potential CRC risk factors, basic sociodemographic information, and medical history. Two different FITs and testing procedures were conducted during the recruitment period: From November 2005 to November 2008, participants filled small containers with raw native fresh stool samples, stored them in provided plastic bags, kept them in the freezer, and brought them to the practice visit at the time of colonoscopy. The stool-filled containers were immediately stored in the freezer (–15 to –40 °C) at the practice visit, shipped on dry ice to a central laboratory within 1 to few days, and analyzed with Ridascreen Hemoglobin (R-Biopharm, Darmstadt, Germany). From November 2008 until February 2012, stool collection was performed in the same way as described above, but another quantitative FIT was applied (FOB Gold; Sentinel Diagnostics, Milano, Italy). From February 2012 until January 2019, participants collected stool samples in collection tubes containing hemoglobin-stabilizing buffer (10-mg stool in 1.7-mL buffer; Sentinel Diagnostics, Milano, Italy; Ref. 11561H). The tubes were to be sealed in envelopes and mailed to the study center at the German Cancer Research Center, where they were kept at 2–8 °C in the refrigerator before transporting in a cold chain to the central laboratory (Labor Limbach, Heidelberg, Germany) for analysis with FOB Gold. It has been shown that diagnostic performance of FIT in the BliTz study was virtually identical between the 2 different sample handling methods used during the recruitment period (21) and that repeated thawing and freezing of stool samples had little impact on accuracy measures of FITs (22). In this study, all analyses were performed and are presented separately for the cohort screened with Ridascreen Hemoglobin (until November 2008) and the cohort screened with FOB Gold (from November 2008 on). The BliTz study has been approved by the Ethics Committees of the Medical Faculty Heidelberg (178/2005) and of the responsible state physicians' chambers (Baden-Württemberg, M118-05-f; Rheinland-Pfalz, 837.047.06(5145); Saarland, 217/13; Hessen, MC 254/2007) and is registered in the German Clinical Trials Register (DRKS-ID: DRKS00008737).

The following exclusion criteria were applied to ensure representativeness of samples for an average-risk screen-naive population and to minimize the number of missed neoplasms: age <50 or ≥80 years; history of CRC or inflammatory bowel disease; colonoscopy in the preceding 5 years; inadequate bowel preparation; incomplete colonoscopy (caecum not reached); and nondefined polyp. To maximize efficiency given limited resources, genotyping was conducted in all advanced neoplasia (AN, CRC, or advanced adenoma [AA]) cases and in a random sample of control subjects, matched by sex, rather than in all participants. The current analysis thus focuses on a subset of participants from whom stool samples and genotyping data are available. We excluded participants whose samples failed quality check during genotyping, leaving a final set of 1,271 participants tested with Ridascreen (172 with AN and 1,099 without AN) and 4,035 participants tested with FOB Gold (523 with AN and 3,512 without AN) (Figure 1).

Genotyping and selection of SNPs

Blood cell DNA was extracted, and genotyping was performed using the OncoArray-500 k V1.0 BeadChip (Illumina, San Diego, CA) for 957 (18.0%) subjects and Global Screening Array (Illumina) for 4,349 (82.0%) subjects. Genotyping was performed in 2014 (18.0% of the samples), in 2017 (31.5%), and in 2020 (50.5%) and hence on average 5.9 (median [interquartile range] 6 [3–8]) years after recruitment. Genotyping quality was assessed as described previously (23). For all rounds of genotyped data, imputation of the missing genotypes was performed using the Haplotype Reference Consortium (version r1.1.2016) as the reference panel. PLINK v1.90 was used to extract SNPs for the regions of interest.

Identification and selection of common risk variants

For the PRS, 140 CRC risk variants that were previously discovered and validated for the European descent population (15) were considered to generate the PRS. One SNP (rs6928864) was not available and thus replaced with another SNP (rs6904092, linkage disequilibrium, $D' = 1$ and $r^2 = 1$), and thus, 140 SNPs were finally included (see Supplementary Table 1, <http://links.lww.com/CTG/A754>). Scores were constructed in PLINK using “mode” function in which missing values of SNPs are first imputed and then summed up, which give an actual range of 112–160 risk alleles. An unweighted PRS based on the number of risk alleles was calculated and used for main analyses. Participants were categorized according to tertiles of the PRS among those free of AN, with categories low-risk (112–132 risk alleles), medium-risk (133–138 risk alleles), and high-risk (139–160 risk alleles). Sensitivity analyses were conducted using a weighted PRS, in which weights were applied to each risk allele, equaling the log odds ratio (OR) of the respective SNP as found in the discovery study of each SNP.

Colonoscopy findings

Colonoscopy and histology records were obtained from the gastroenterology practices on completion of colonoscopy. Colonoscopy and histology reports were reviewed, and findings at colonoscopy were independently extracted in a standardized manner by 2 trained investigators who were blinded with respect to stool test results. Participants were classified according to the most advanced finding at colonoscopy (CRC, AA, non-AA, other,

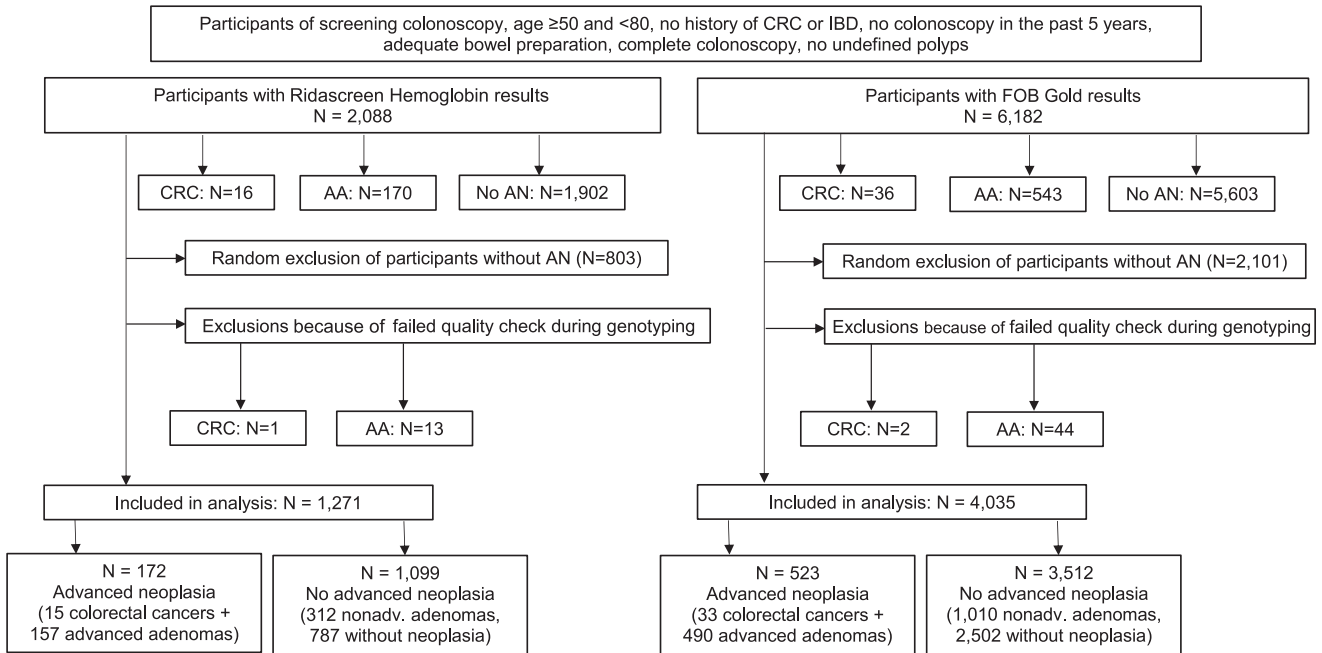


Figure 1. Flow chart of the study participants included in this analysis. AA, advanced adenoma; AN, advanced neoplasia (CRC or AA); CRC, colorectal cancer; IBD, inflammatory bowel disease.

or none of above). Adenomas were defined as advanced if they were ≥ 1 cm, had tubulovillous or villous histology, or high-grade dysplasia.

Laboratory analyses

Fecal hemoglobin concentrations were measured using Ridascreen Hemoglobin, which is based on an enzyme immunoassay (participants recruited until December 2008), and FOB Gold, which is based on a latex agglutination assay (participants recruited from December 2008 onward). In both groups of participants, fully automated FIT analyses were conducted, blinded with respect to colonoscopy results, using Tecan Freedom EVOlyzer[®] (Ridascreen Hemoglobin) and Abbott Architect c8000 (FOB Gold), respectively. The median time (interquartile range) between fecal sampling and laboratory analysis was 9 (6–13) days.

Statistical analyses

Results of screening colonoscopy and FIT were used to derive PPVs and NNS of FIT for detection of 1 AN. To make results comparable across the 2 FITs, cutoffs were preset to yield defined levels of specificities. We chose specificities of 85%, 90%, and 95% (corresponding to low, intermediate, and high cutoff) to reflect the high levels of specificity required for population-based screening.

PPVs for AN were calculated for overall FIT results and for FIT in each PRS group. We assumed the same sensitivity between PRS groups for a given FIT and specificity. Because of the oversampling of AN cases (in all of whom genotyping was conducted) compared with controls (where genotyping was conducted only in a random subset of all subjects), we could not calculate PPVs and 95% confidence intervals (CIs) directly as proportions from contingency (2-by-2) tables of true and false positives and negatives. Instead, we derived sensitivities and specificities from such

tables and estimated the prevalence of AN, which is additionally needed for the PPV calculations as follows: Prevalence of AN in subjects with medium PRS was assumed to equal that of the entire study population, i.e. $\text{prevalence}_{\text{medium}} = \text{overall prevalence}$. Prevalence in subjects with low PRS was approximated by $\text{prevalence}_{\text{low}} = \text{prevalence}_{\text{medium}} \times \text{OR}_{\text{low/medium}}$, where $\text{OR}_{\text{low/medium}}$ is the odds ratio of AN in the low PRS group compared with the medium PRS group. Similarly, prevalence in subjects with high PRS was approximated by $\text{prevalence}_{\text{high}} = \text{prevalence}_{\text{medium}} \times \text{OR}_{\text{high/medium}}$. The ORs were derived from multivariable logistic regression models with AN as outcome and PRS as categorical explanatory variable (medium PRS as the reference group), adjusted for age and sex. We then calculated PPVs and asymptotic standard logit CIs (24) from the estimated prevalences of having AN according to the PRS group and the 2-by-2 tables using the function “BDtest” from the R package “bdpv” (25). That function works with the logit transformations of adjusted sensitivity and specificity (observed cell counts plus continuity correction in each cell) from which the SD of the (logit) PPV is calculated using the delta method. CIs are then obtained by $\exp(\text{estimate}) / (1 + \exp(\text{estimate}))$, where $\text{estimate} = \text{logit PPV} \pm 1.96 \times \text{SD}(\text{logit}(\text{PPV}))$.

As the main outcome, we calculated PPVs of FITs for different cutoffs and PRS categories. Secondary outcomes were differences in cutoffs for different PRS groups achieving comparable PPVs for each FIT and differences in PPVs according to PRS groups when using uniform cutoffs. These differences were tested for statistical significance using 95% bias-corrected and accelerated bootstrap (26) CIs (R package boot (27)).

PPVs of FIT were examined according to PRS strata and FIT cutoffs for both FITs. Statistical tests were 2-sided with the alpha level set at 0.05. All statistical analyses were conducted using R (28).

Table 1. Characteristics of the study population with PRS and FIT data available

Characteristic	Ridascreen Hemoglobin		FOB Gold	
	Total, N = 1,271	%	Total, N = 4,035	%
Sex				
Male	680	53.5	2,076	51.4
Female	591	46.5	1,959	48.6
Age (yr)				
50–54	40	3.1	176	4.4
55–59	407	32.0	1,706	42.3
60–64	289	22.7	852	21.1
65–69	317	24.9	654	16.2
70–74	164	12.9	457	11.3
75–79	54	4.2	190	4.7
Most advanced finding at screening colonoscopy ^a				
CRC	15	1.2	33	0.8
AA	157	12.4	490	12.1
Advanced neoplasia (CRC or AA)	172	13.5	523	12.9
Nonadvanced adenoma	312	24.5	1,010	25.0
No neoplasm	787	61.9	2,502	62.0

AA, advanced adenoma; CRC, colorectal cancer; FIT, fecal immunochemical test; PRS, polygenic risk score.
^aParticipants with colorectal cancer or advanced adenoma are overrepresented because genotyping was performed in all colorectal cancer and advanced adenoma cases but only in a subset of participants free of advanced neoplasms.

RESULTS

The study included 1,271 participants tested with Ridascreen Hemoglobin and 4,035 participants tested with FOB Gold (Table 1). Because of genotyping in a random sample of participants without AN only, participants with AN were overrepresented in both cohorts ($n = 172$ and 523 , including 15 and 33 CRC cases and 157 and 490 participants with AA as the most advanced finding at colonoscopy, respectively). In both cohorts, a slight majority of participants were men (53.5% and 51.4%), and the mean age was 63.1 and 61.8 years, respectively.

Risk allele distribution and estimated neoplasia prevalence

Age- and sex-adjusted odds ratios for having any AN indicated a higher risk in the high PRS group ($OR_{\text{high/medium}} = 1.29$, 95% CI 1.07–1.56) and a lower risk in the low PRS group ($OR_{\text{low/medium}} = 0.75$, 95% CI 0.61–0.92). These ORs were used to approximate the PRS-specific prevalence of AN in this study. Estimated prevalences for AN were 9.33% in subjects with medium PRS (just like in the overall study population, where 57 CRC and 751 AA cases were found among 8,662 participants), 12.1% ($9.33\% \times 1.29$) in high-risk subjects, and 7.0% in low-risk subjects.

Overall and PRS-specific performance of FIT

In participants screened with Ridascreen Hemoglobin, we found a steady increase in PPVs with higher genetic risk. For example, at

an intermediate cutoff (90% specificity), PPV was 24% with low PRS compared with 30% and 37% with medium and high risk, respectively (Table 2). Similar increases with PRS were found at a low cutoff (85% specificity), despite overall lower PPVs (19%, 25%, and 31%), and at a high cutoff (95% specificity) with overall higher PPVs (34%, 42%, and 49%) with low, medium, and high PRS, respectively. Differences in PPVs between PRS groups were 5%–7% units for comparisons of medium vs low PRS, 6%–7% units for high vs medium PRS, and 11%–15% units for high vs low PRS. All differences in PPVs were statistically significant.

Almost identical results were found in the larger sample of participants tested with FOB Gold using the same PRS: PPVs steadily increased with higher cutoff (27%, 32%, and 42% among those with medium PRS) and with higher PRS (26%, 32%, and 39% for low, medium, and high risk, respectively, at an intermediate cutoff). Again, all differences in PPV between PRS groups were statistically significant. Using the weighted PRS yielded very similar results when comparing high vs low PRS (see Supplementary Table 2, <http://links.lww.com/CTG/A754>) for both Ridascreen and FOB Gold.

Table 3 shows the corresponding NNS to detect 1 AN. Although up to 4 individuals at low genetic risk need to be scoped with Ridascreen at 90% specificity (4.1, 95% CI 3.5–5.0), this number gradually decreased to 2.7 for individuals with high PRS. At 85% specificity, NNS was only slightly higher in the high PRS group (3.3) than at 95% specificity in the low PRS group (2.9). A comparable pattern was found in those tested with FOB Gold, with NNS ranging from 4.6 (low PRS, low specificity) to 2.0 (high PRS, high specificity), again with similar NNS in those with high PRS at a low FIT cutoff (3.0) and those with low PRS and high FIT cutoff (2.9).

In both FITs, the same PPV was reached at much lower cutoffs in subjects with high and medium PRS compared with the low-risk group (Figure 2). For example, to achieve a PPV of 25%, a Ridascreen cutoff of 4.5 $\mu\text{g/g}$ was required in the low PRS group, compared with 1.7 and 0.8 $\mu\text{g/g}$ in the medium and high PRS groups, respectively (Table 4). For FOB Gold, the corresponding numbers were 10.9, 7.9, and 6.5 $\mu\text{g/g}$. Also in both samples, differences in cutoffs between PRS groups increased with increasing PPV. Differences between PRS groups were statistically significant in nearly all comparisons.

DISCUSSION

In this article, PPVs for screen-detected AN were investigated independently for 2 quantitative FITs, Ridascreen Hb and FOB Gold, according to FIT cutoff and a PRS. In both cohorts, PPVs for having AN were consistently higher (approximately 1.5-fold) in individuals in the upper compared with the lower tertile of genetic risk. Accordingly, considerably higher cutoffs were required in lower risk subjects to achieve comparable PPVs as in higher risk subjects. Differences in cutoffs between low- and high-risk groups were between 1.6 and 9.5 $\mu\text{g/g}$ (Ridascreen) and between 2.9 and 14.2 $\mu\text{g/g}$ (FOB Gold) for PPVs between 20% and 35%, whereas use of the same cutoffs for all screenees resulted in substantially different PPVs across groups. Using the weighted PRS resulted in very similar PPVs and cutoff differences as the unweighted PRS.

Our findings may be relevant for personalized screening. For example, subjects might be referred to colonoscopy not solely because of a positive FIT but according to their estimated risk of having AN. Baseline risks could be stratified into high, medium,

Table 2. Comparison of positive predictive values for detecting advanced neoplasia of both FITs achieved at different PRS categories and specificities

FIT	Cutoff ^a	Positive predictive value for advanced neoplasia (%) (95% CI)			Difference in positive predictive value for advanced neoplasia (%) (95% CI) ^b		
		Low PRS	Medium PRS	High PRS	Medium vs low	High vs medium	High vs low
Ridascreen Hemoglobin	Low	19 (16–23)	25 (21–29)	31 (26–35)	5 (2–9)	6 (2–9)	11 (7–16)
	Intermediate	24 (20–29)	30 (25–36)	37 (31–43)	6 (2–10)	6 (2–10)	13 (8–18)
	High	34 (27–42)	42 (34–50)	49 (41–57)	7 (2–12)	7 (2–12)	15 (10–20)
FOB Gold	Low	22 (20–24)	27 (25–30)	33 (31–36)	6 (2–10)	6 (2–11)	11 (8–17)
	Intermediate	26 (23–28)	32 (29–35)	39 (36–42)	6 (2–11)	7 (2–12)	13 (9–18)
	High	35 (31–39)	42 (38–47)	49 (45–54)	6 (2–13)	7 (2–13)	15 (10–20)

CI, confidence interval; FIT, fecal immunochemical test; PRS, polygenic risk score.

^aCutoffs were defined to yield 85% (low), 90% (intermediate), and 95% (high) specificity and were as follows (expressed in µg hemoglobin per gram of stool): Ridascreen Hemoglobin 1.7 µg/g (low), 4.0 µg/g (intermediate), and 12.0 µg/g (high); FOB Gold 8.5 µg/g (low), 11.6 µg/g (intermediate), and 22.7 µg/g (high).

^bStatistically significant differences (α = 0.05) between positive predictive values are indicated in bold.

and low risk according to PRS and further refined by FIT. With low genetic risk, a higher FIT cutoff would be required to achieve the same PPV as in high-risk individuals. In clinical practice, one might consider colonoscopy referral only given a certain probability of having AN, i.e., by raising the FIT cutoff in low-risk individuals, and lowering the cutoff in high-risk individuals. Estimation of such a referral threshold may be based on FIT and PRS result, like in our study, but could be extended to further CRC risk factors or preventive factors (individually or in combination), such as sex, age, having a family history of CRC, or having ever had a colonoscopy (29). The priority of including a risk or preventive factor in an overall score should depend not solely on strengths of associations but also address difficulty/costs of measurements. With such a stepwise approach, some criteria or tests might be expendable if the initial test(s) already indicate(s) a high risk. For instance, an old male screenee with a very high Hb concentration would be considered at high risk even if his genetic risk was low, and costly genotyping would be unnecessary. Applicability of such an approach in different countries, with their heterogeneous screening programs (organized vs opportunistic, with varying extent of additional screening for those with family history of CRC or for polyp surveillance etc.) remains to be determined.

Although several previous studies examined combinations of FIT and other diagnostic tests, mainly stool tests (30) but also blood tests (31) for accuracy of colorectal neoplasia detection, to the best of our knowledge, no study to date investigated PPVs of FIT alone and according to genetic risk. Although other studies assessed PPVs of FIT when combined with another marker (32–35), these studies typically focused on changes in sensitivity, specificity, and PPV by simple combination of both markers, such as definition of test results as positive if 1 of 2 tests was positive, rather than possible adaptation of FIT cutoffs.

To the best of our knowledge, this is the first study investigating genetic risk group-specific FIT cutoffs to achieve prespecified PPVs. Few further studies examined FIT combined with risk scores, although not with a PRS (36–39), but studied different outcomes such as area under the receiver operating characteristics curves (36,37) and sensitivities (38,39). In addition to FIT, those studies used conventional CRC risk factors (36), calcium intake, family history and age (37), high blood pressure, age and abdominal obesity (38), and the Asia-Pacific Colorectal Screening scoring system (39). Discriminatory ability of PRS alone (without FIT) has been examined in further studies (40–43), which focused on CRC as the outcome. Peng et al. (44) found substantial differences in cutoffs required to achieve

Table 3. Comparison of NNSs for detecting advanced neoplasia of both FITs achieved at different PRS categories and specificities

FIT	Cutoff ^a	NNS for advanced neoplasia (%) (95% CI)			Difference in number needed to scope for advanced neoplasia (%) (95% CI) ^b		
		Low PRS	Medium PRS	High PRS	Medium vs low	High vs medium	High vs low
Ridascreen Hemoglobin	Low	5.1 (4.4–6.1)	4.0 (3.4–4.7)	3.3 (2.8–3.8)	1.1 (0.3–2.1)	0.8 (0.2–1.2)	1.9 (1.0–2.8)
	Intermediate	4.1 (3.5–5.0)	3.3 (2.8–3.9)	2.7 (2.3–3.2)	0.8 (0.2–1.6)	0.6 (0.2–0.9)	1.4 (0.8–2.2)
	High	2.9 (2.4–3.7)	2.4 (2.0–2.9)	2.0 (1.7–2.5)	0.5 (0.1–1.0)	0.3 (0.1–0.6)	0.9 (0.5–1.3)
FOB Gold	Low	4.6 (4.3–5.1)	3.7 (3.4–4.0)	3.0 (2.8–3.2)	1.0 (0.2–1.8)	0.7 (0.2–1.0)	1.7 (0.9–2.4)
	Intermediate	3.9 (3.5–4.3)	3.1 (2.8–3.4)	2.6 (2.4–2.8)	0.8 (0.2–1.5)	0.5 (0.2–0.8)	1.3 (0.7–1.9)
	High	2.9 (2.6–3.3)	2.4 (2.1–2.7)	2.0 (1.9–2.2)	0.5 (0.1–1.0)	0.3 (0.1–1.0)	0.9 (0.5–1.3)

CI, confidence interval; FIT, fecal immunochemical test; NNS, numbers needed to scope; PRS, polygenic risk score.

^aCutoffs were defined to yield 85% (low), 90% (intermediate) and 95% (high) specificity and were as follows (expressed in µg hemoglobin per gram of stool): Ridascreen Hemoglobin 1.7 µg/g (low), 4.0 µg/g (intermediate), and 12.0 µg/g (high); FOB Gold 8.5 µg/g (low), 11.6 µg/g (intermediate), and 22.7 µg/g (high).

^bStatistically significant differences (α = 0.05) between NNS are indicated in bold.

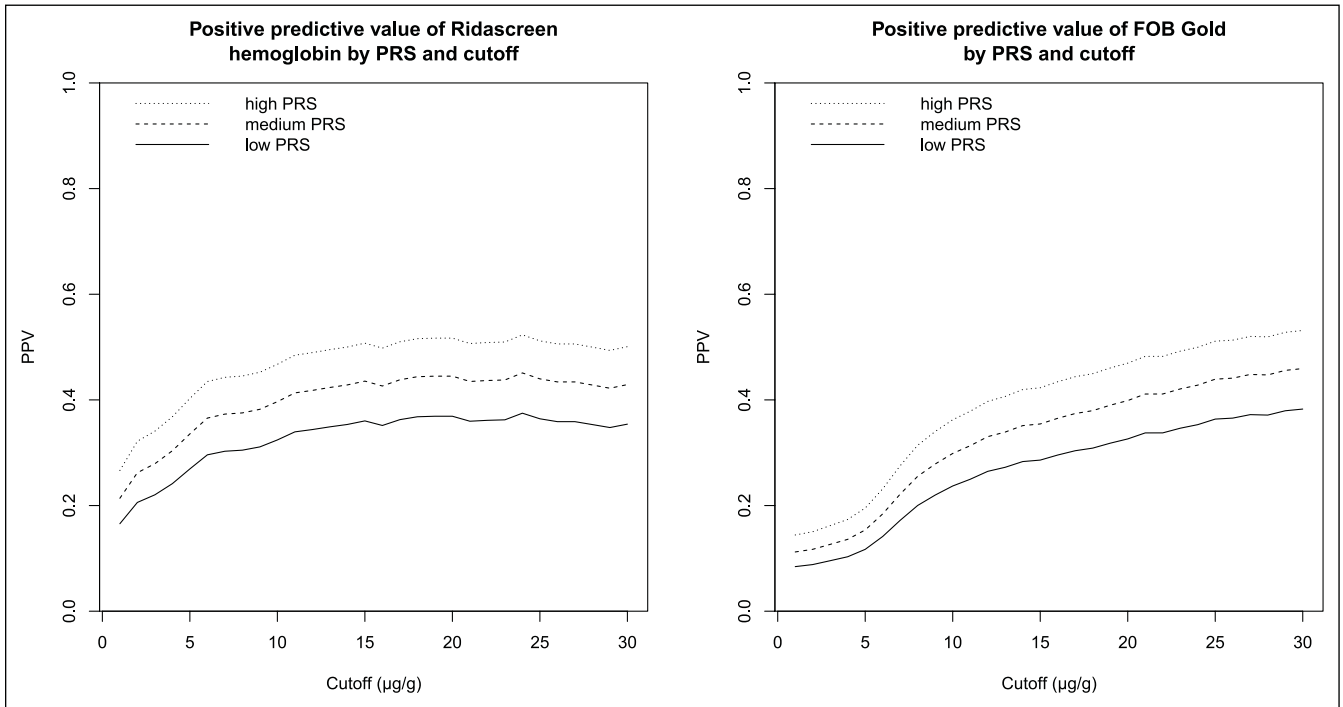


Figure 2. Positive predictive values for advanced neoplasia detection according to examined FIT, cutoff, and PRS category. FIT, fecal immunochemical test; Hb, hemoglobin; PPV, positive predictive value; PRS, polygenic risk score.

comparable PPVs for AN in low-, medium-, and high-risk groups defined by environmental risk scores, similar to our study. Combining information of environmental and genetic risk might result in even stronger variation of PPVs at fixed cutoffs or of cutoffs at fixed PPVs across risk groups.

Our study has several strengths. PPVs for AN detection were estimated using FIT alone and additional polygenic risk stratification, achieving more accurate risk prediction than FIT alone. We corrected for oversampling of AN cases in PPV estimations. Participants were recruited from an

asymptomatic population undergoing CRC screening. Thus, spectrum bias potentially introduced when including symptomatic CRC cases, which might be easier to detect, was avoided. Very large numbers of genotyped individuals with and without AN were included, yielding narrow CIs for PPVs. Results from screening colonoscopy were available for all subjects, ruling out verification bias (45). Two samples using different FITs were used, and we found very high agreement in findings of the 2 cohorts. Blinding of laboratory personnel toward colonoscopy results was ensured. Potential overfitting

Table 4. Combinations of FIT cutoffs achieving comparable positive predictive values and numbers needed to scope

Positive predictive value for AN detection	FIT cutoff (µg/g)			Difference in cutoff (µg/g) (95% CI)		
	Low PRS	Medium PRS	High PRS	Medium vs low	High vs medium	High vs low
Ridascreen Hemoglobin						
20%	1.8	0.8	0.2	1.0 (0.1 to 3.1)	0.6 (0.3 to 1.8)	1.6 (0.8 to 4.5)
25%	4.5	1.7	0.8	2.8 (0.4 to 8.3)	0.9 (0.1 to 2.0)	3.7 (1.1 to 7.2)
30%	6.6	3.9	1.6	2.7 (-17.2 to 5.9)	2.3 (0.9 to 6.8)	5.0 (1.8 to 22.1)
35%	13.0	5.3	3.5	7.7 (-4.0 to 31.1)	1.8 (0.4 to 12.2)	9.5 (2.5 to 32.9)
FOB Gold						
20%	8.0	6.5	5.1	1.5 (0.5 to 4.1)	1.4 (0.2 to 3.2)	2.9 (1.3 to 4.4)
25%	10.9	7.9	6.5	3.0 (0.4 to 7.8)	1.4 (0.5 to 2.6)	4.4 (1.9 to 9.0)
30%	16.5	10.1	7.6	6.4 (1.2 to 15.0)	2.5 (0.5 to 3.6)	8.9 (3.2 to 16.6)
35%	23.4	13.8	9.2	9.6 (1.4 to 22.0)	4.6 (1.4 to 8.3)	14.2 (5.3 to 28.3)

95% CI for differences in cutoffs was calculated using bias-corrected and accelerated (BCa) bootstrap and at least 2,000 bootstrap replicates. Statistically significant differences ($\alpha = 0.05$) between cutoff are indicated in bold.

AN, advanced neoplasia (colorectal cancer or advanced adenoma); CI, confidence interval; FIT, fecal immunochemical test; PRS, polygenic risk score.

by combining FIT and PRS was circumvented by using a simple tertile-based PRS categorization.

Our study also has limitations. Prevalences of AN in PRS strata were unknown by design of the study and were therefore estimated from overall prevalence and ORs for high and low vs average genetic risk using logistic regression models for PPV estimates. We refrained from estimating PRS-specific sensitivities of FIT because of comparably small case numbers in each PRS group. Genotyping was solely performed in screening participants from South-Western Germany, the majority of whom is of European descent. Thus, generalizability of our results is limited. Also, although results of 2 FITs were very similar, replication in further studies using different (quantitative) FITs is warranted. By design, qualitative FITs would not be suitable for risk-adapted cutoffs. Replication studies should also comprise subjects from other countries and continents and use different PRS. Finally, sessile-serrated lesions were not recorded consistently as such in the early years of study recruitment, which is why we could not consider them in our analyses. It is known that FITs have poor sensitivity for sessile-serrated lesions (46).

Further large-scale studies are required for comparisons of adherence, optimal screening intervals, and resulting (cost)effectiveness of individual and combined CRC screening using PRS, FIT, and potential further markers. According to a previous large-scale study (16), having a family history and PRS were surprisingly weakly correlated, indicating further potential for improvement in risk stratification by combining those factors.

Use of PRS for adjustment of FIT cutoffs in routine practice would require 1-time genetic testing. Although currently not routinely implemented, determination of PRS for risk stratification and related indications, such as determining risk-adapted starting ages of screening, has recently received increased attention, and genetic testing is increasingly commercially offered and used in several countries, such as the United States. With genetic information from such testing available, adaptation of FIT cutoffs according to genetic risk could be performed at virtually no extra cost. If such information is not available, it might be obtained by a single test at first use of FIT. In case of further improvements in genetic risk prediction, individuals with high genetic risk might be recommended to undergo colonoscopy directly irrespective of FIT. If and under which conditions genetic testing might be implemented in screening practice requires further research and societal discussion which has to take additional aspects, such as ethical, legal, logistic, and economic aspects into account.

Notwithstanding the need of such further research and discussion, our results may provide important information for future screening programs aimed to personalize screening. In particular, they provide important empirical evidence for the variation of FIT cutoffs that would be needed to ensure defined levels of PPV and NNS according to PRS-defined risk groups and for the strong variation of PPVs and NNS of FITs for AN detection at fixed FIT cutoffs.

CONFLICTS OF INTEREST

Guarantor of the article: Tobias Niedermaier, PhD, MPH.

Specific author contributions: H.B.: designed the study. K.W. and F.G.: implemented the polygenic risk score. T.N. and Y.B.: conducted the statistical analyses. T.N.: drafted the manuscript. Y.B., A.G., K.W., F.G., E.A., M.H., and H.B.: contributed to important intellectual content and critically revised the manuscript. All authors approved the final draft submitted.

Financial support: This study was partly funded by grants from the German Research Council (DFG, grant No. BR1704/16–1), the Federal Ministry of Education and Research (BMBF, grant no. 01GL1712), and the German Cancer Aid (No. 70113330).

Potential competing interests: None to report.

ACKNOWLEDGEMENTS

We thank Silvia Calderazzo and Christian Stock for statistical advice. We gratefully acknowledge the excellent cooperation of gastroenterology practices and clinics in patient recruitment and of Labor Limbach in sample collection. We gratefully acknowledge Dr. Katja Butterbach, Dr. Katarina Cuk, and Ulrike Schlesselmann for their excellent work in laboratory preparation of blood samples. We also gratefully acknowledge Isabel Lerch, Susanne Köhler, Dr. Utz Benschaid, Jason Hochhaus, and Maria Kuschel for their contribution in data collection, monitoring, and documentation.

REFERENCES

- Sung JJ, Ng SC, Chan FK, et al. An updated Asia Pacific Consensus Recommendations on colorectal cancer screening. *Gut* 2015;64:121–32.
- Bacchus CM, Dunfield L, Gorber SC, et al. Recommendations on screening for colorectal cancer in primary care. *CMAJ* 2016;188:340–8.
- Rex DK, Boland CR, Dominitz JA, et al. Colorectal cancer screening: Recommendations for physicians and patients from the U.S. Multi-Society Task Force on colorectal cancer. *Gastroenterology* 2017;153:307–23.
- Wolf AMD, Fontham ETH, Church TR, et al. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. *CA Cancer J Clin* 2018;68:250–81.
- Ebell MH, Thai TN, Royalty KJ. Cancer screening recommendations: An international comparison of high income countries. *Public Health Rev* 2018;39:7.
- Tenesa A, Farrington SM, Prendergast JG, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat Genet* 2008;40:631–7.
- Tomlinson IP, Webb E, Carvajal-Carmona L, et al. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet* 2008;40:623–30.
- Peters U, Hutter CM, Hsu L, et al. Meta-analysis of new genome-wide association studies of colorectal cancer risk. *Hum Genet* 2012;131:217–34.
- Peters U, Jiao S, Schumacher FR, et al. Identification of genetic susceptibility loci for colorectal tumors in a genome-wide meta-analysis. *Gastroenterology* 2013;144:799–807.e24.
- Hsu L, Jeon J, Brenner H, et al. A model to determine colorectal cancer risk using common genetic susceptibility loci. *Gastroenterology* 2015;148:1330–9.e14.
- Schumacher FR, Schmit SL, Jiao S, et al. Genome-wide association study of colorectal cancer identifies six new susceptibility loci. *Nat Commun* 2015;6:7138.
- Schmit SL, Edlund CK, Schumacher FR, et al. Novel common genetic susceptibility loci for colorectal cancer. *J Natl Cancer Inst* 2019;111:146–57.
- 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.
- Law PJ, Timofeeva M, Fernandez-Rozadilla C, et al. Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nat Commun* 2019;10:2154.
- Thomas M, Sakoda LC, Hoffmeister M, et al. Genome-wide modeling of polygenic risk score in colorectal cancer risk. *Am J Hum Genet* 2020;107:432–44.
- Weigl K, Chang-Claude J, Knebel P, et al. Strongly enhanced colorectal cancer risk stratification by combining family history and genetic risk score. *Clin Epidemiol* 2018;10:143–52.
- Weigl K, Thomsen H, Balavarca Y, et al. Genetic risk score is associated with prevalence of advanced neoplasms in a colorectal cancer screening population. *Gastroenterology* 2018;155:88–98.e10.

18. Gies A, Cuk K, Schrotz-King P, et al. Direct comparison of diagnostic performance of 9 quantitative fecal immunochemical tests for colorectal cancer screening. *Gastroenterology* 2018;154:93–104.
19. Amitay EL, Cuk K, Niedermaier T, et al. Factors associated with false-positive fecal immunochemical tests in a large German colorectal cancer screening study. *Int J Cancer* 2018;144:2419–27.
20. Niedermaier T, Amitay EL, Gies A, et al. Impact of inadequate bowel cleansing on colonoscopic findings in routine screening practice. *Clin Transl Gastroenterol* 2020;11:e00169.
21. Chen H, Werner S, Brenner H. Fresh vs frozen samples and ambient temperature have little effect on detection of colorectal cancer or adenomas by a fecal immunochemical test in a colorectal cancer screening cohort in Germany. *Clin Gastroenterol Hepatol* 2017;15:1547–56.e5.
22. Gies A, Niedermaier T, Weigl K, et al. Effect of long-term frozen storage and thawing of stool samples on faecal haemoglobin concentration and diagnostic performance of faecal immunochemical tests. *Clin Chem Lab Med* 2019;58:390–8.
23. Anderson CA, Pettersson FH, Clarke GM, et al. Data quality control in genetic case-control association studies. *Nat Protoc* 2010;5:1564–73.
24. Mercaldo ND, Lau KF, Zhou XH. Confidence intervals for predictive values with an emphasis to case-control studies. *Stat Med* 2007;26:2170–83.
25. Schaarschmidt F. [Package "bdpv"]: Inference and Design for Predictive Values in Binary Diagnostic Tests. 2019 (<https://cran.r-project.org/web/packages/bdpv/>).
26. Efron B. Better bootstrap confidence intervals. *J Am Stat Assoc* 1987;82:171–85.
27. Canty ARB. boot: Bootstrap functions. 2017;2018. In: Davison AC, Hinkleyfrom DV (eds). *Bootstrap Methods and Their Application*. 1997, CUP, originally written by Angelo Canty for S: Cambridge University Press.
28. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria, 2021.
29. Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet* 2014;383:1490–502.
30. Niedermaier T, Weigl K, Hoffmeister M, et al. Fecal immunochemical tests combined with other stool tests for colorectal cancer and advanced adenoma detection: A systematic review. *Clin Transl Gastroenterol* 2016;7:e175.
31. Niedermaier T, Weigl K, Hoffmeister M, et al. Fecal immunochemical tests in combination with blood tests for colorectal cancer and advanced adenoma detection-systematic review. *United Eur Gastroenterol J* 2018;6:13–21.
32. Parente F, Marino B, Ilardo A, et al. A combination of faecal tests for the detection of colon cancer: A new strategy for an appropriate selection of referrals to colonoscopy? A prospective multicentre Italian study. *Eur J Gastroenterol Hepatol* 2012;24:1145–52.
33. Harada T, Yamamoto E, Yamano HO, et al. Analysis of DNA methylation in bowel lavage fluid for detection of colorectal cancer. *Cancer Prev Res (Phila)* 2014;7:1002–10.
34. Otero-Estevez O, De Chiara L, Rodriguez-Berrocal FJ, et al. Serum sCD26 for colorectal cancer screening in family-risk individuals: Comparison with faecal immunochemical test. *Br J Cancer* 2014;6:375–81.
35. Otero-Estevez O, De Chiara L, Barcia-Castro L, et al. Evaluation of serum nucleoside diphosphate kinase A for the detection of colorectal cancer. *Scientific Rep* 2016;6:26703.
36. Yen AM, Chen SL, Chiu SY, et al. A new insight into fecal hemoglobin concentration-dependent predictor for colorectal neoplasia. *Int J Cancer* 2014;135:1203–12.
37. Stegeman I, de Wijkerslooth TR, Stoop EM, et al. Combining risk factors with faecal immunochemical test outcome for selecting CRC screenees for colonoscopy. *Gut* 2014;63:466–71.
38. Wu T, Kuo K, Wu Y, et al. Diagnostic accuracy of a single qualitative immunochemical fecal occult blood test coupled with physical measurements. *Chin Med J (Engl)* 2014;127:4164–70.
39. 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* 2015;150:617–25.
40. Jo J, Nam CM, Sull JW, et al. Prediction of colorectal cancer risk using a genetic risk score: The Korean cancer prevention study-II (KCPS-II). *Genomics Inform* 2012;10:175–83.
41. Jung KJ, Won D, Jeon C, et al. A colorectal cancer prediction model using traditional and genetic risk scores in Koreans. *BMC Genet* 2015;16:49.
42. Hosono S, Ito H, Oze I, et al. A risk prediction model for colorectal cancer using genome-wide association study-identified polymorphisms and established risk factors among Japanese: Results from two independent case-control studies. *Eur J Cancer Prev* 2016;25:500–7.
43. Iwasaki M, Tanaka-Mizuno S, Kuchiba A, et al. Inclusion of a genetic risk score into a validated risk prediction model for colorectal cancer in Japanese men improves performance. *Cancer Prev Res (Phila)* 2017;10:535–41.
44. Peng L, Balavarca Y, Niedermaier T, et al. Risk-adapted cutoffs in colorectal cancer screening by fecal immunochemical tests. *Am J Gastroenterol* 2020;115:1110–6.
45. Kohn MA, Carpenter CR, Newman TB. Understanding the direction of bias in studies of diagnostic test accuracy. *Acad Emerg Med* 2013;20:1194–206.
46. Chang LC, Shun CT, Hsu WF, et al. Fecal immunochemical test detects sessile serrated adenomas and polyps with a low level of sensitivity. *Clin Gastroenterol Hepatol* 2017;15:872–9.e1.

Open Access This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.