Supplementary Table 1. Information about genotyping and imputation within DACHS

Genotyping platform	Number of cases	Number of controls	Recruitment period	Imputation
Illumina HumanCytoSNP	1,593	1,645	2003-2008	Cosmopolitan panel of reference haplotypes from Phase 1 of the 1,000 Genome Project
Illumina HumanOmniExpress	654	473	2007-2010	
Illumina HumanOmniExpress	1,122	598	2010-2015	Haplotype Reference Consortium (Version r1.1.2016)
Illumina Infinium OncoArray	851	622	2003-2016	

Triallelic SNPs and those not assigned an rs-number were excluded, as were genotyped SNPs when they had a low call rate (<98%), lack of Hardy-Weinberg equilibrium in controls (p<1x10-4) or low minor allele frequency (<0.1%). See Peters et al[1] and Schumacher et al[2] for more information about genotyping and imputation. Please note that data from Illumina HumanCytoSNP and the Illumina HumanOmniExpress with recruitment period from 2007-2010 have been used for previous analyses within the DACHS study[3].

References

1. 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.

2. 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.

3. 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;10:143-152.

Supplementary Table 2. Risk of colorectal cancer according to polygenic risk score and adherence to a healthy lifestyle

	Colorectal cancer	
Subgroup	**n_{cases}(%)/$n_{controls}$(%)**	**OR (95% CI)[1]**
Low genetic risk		
Unfavourable lifestyle	264(29)/215(19)	1.00 (Ref.)
Intermediate lifestyle	350(39)/386(35)	0.75 (0.59-0.96)
Favourable lifestyle	287(32)/513(46)	0.45 (0.35-0.58)
P trend		<0.0001
Intermediate genetic risk		
Unfavourable lifestyle	434(32)/253(22)	1.00 (Ref.)
Intermediate lifestyle	470(34)/384(34)	0.71 (0.57-0.88)
Favourable lifestyle	464(34)/501(44)	0.60 (0.48-0.75)
P trend		<0.0001
High genetic risk		
Unfavourable lifestyle	623(32)/264(24)	1.00 (Ref.)
Intermediate lifestyle	684(35)/359(33)	0.86 (0.70-1.05)
Favourable lifestyle	644(33)/463(43)	0.61 (0.49-0.75)
P trend		<0.0001

[1]Model adjusted for matching factors age and sex; and the factors school education, family history of colorectal cancer, colonoscopy, participation in a health check-up, ever regular use of non-steroidal anti-inflammatory drugs (NSAIDs).

Abbreviations: CI: confidence intervals; OR: odds ratio; Ref.: Reference

Supplementary Table 3. Association between the healthy lifestyle score and colorectal cancer in the DACHS study by polygenic risk score

		Colorectal cancer		$P_{interaction}$
	Lifestyle score	**n_{cases}(%)/$n_{controls}$(%)**	**OR (95% CI)[1]**	
Polygenic risk score				
<median	0 or 1	135 (9)/78(5)	1.00 (Ref.)	
	2	342(23)/257(15)	0.73 (0.52-1.04)	
	3	566(37)/568 (34)	0.55 (0.40-0.76)	
	4	384(25)/524(31)	0.42 (0.30-0.58)	
	5	101(7)/242(15)	0.25 (0.17-0.37)	
	P trend		<0.0001	
				0.08
≥median	0 or 1	216(8)/89(5)	1.00 (Ref.)	
	2	628(23)/308(18)	0.86 (0.64-1.15)	
	3	938(35)/561(34)	0.73 (0.55-0.96)	
	4	699(26)/477(29)	0.63 (0.47-0.84)	
	5	211(8)/234(14)	0.38 (0.27-0.53)	
	P trend		<0.0001	

[1]Model adjusted for matching factors age and sex; and the factors school education, family history of colorectal cancer, colonoscopy, participation in a health check-up, ever regular use of non-steroidal anti-inflammatory drugs (NSAIDs).

Abbreviations: CI: confidence intervals; OR: odds ratio; Ref.: Reference

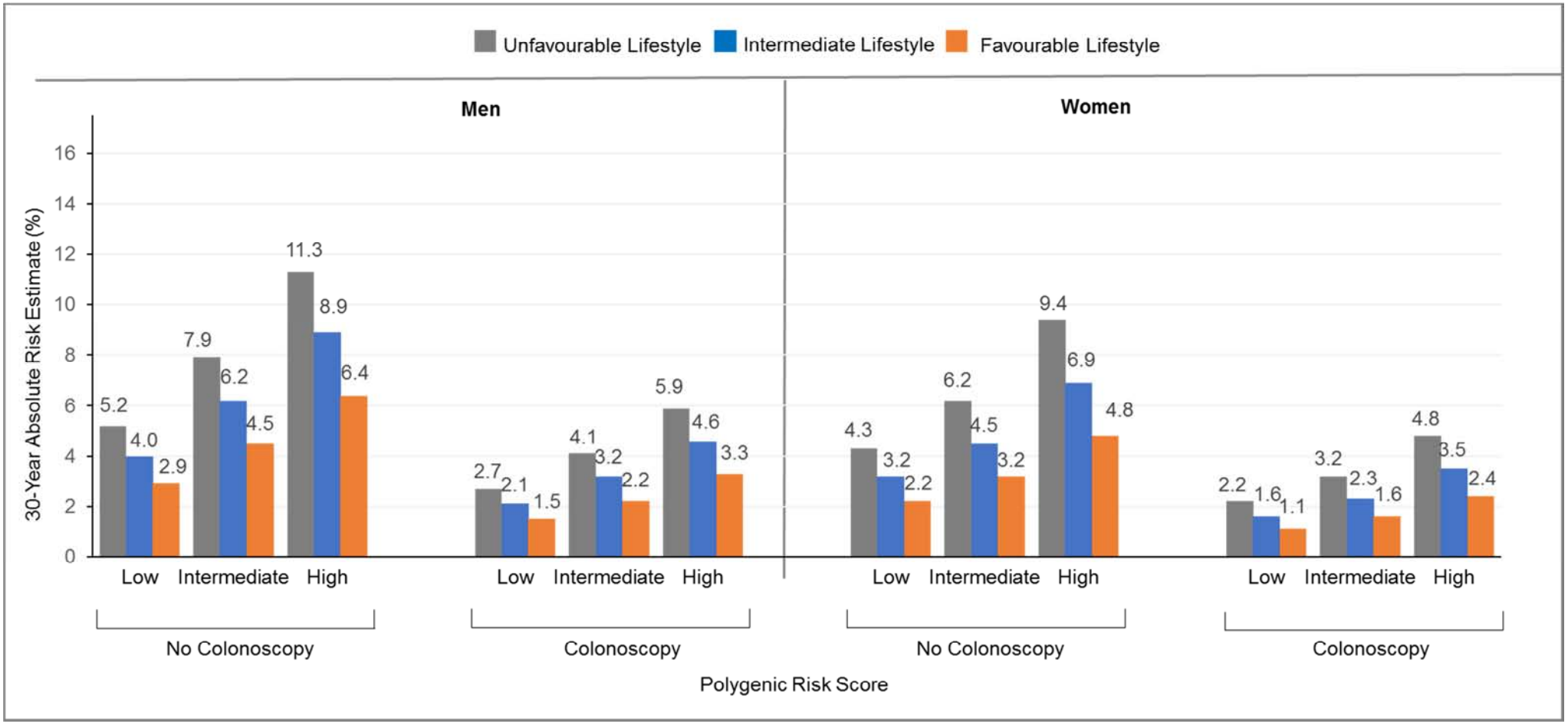
Supplementary Table 4. Association between the healthy lifestyle score and colorectal cancer in the DACHS study by polygenic risk score (tertiles)

		Colorectal cancer		$P_{interaction}$
	Lifestyle score	**n_{cases}(%)/$n_{controls}$(%)**	**OR (95% CI)[1]**	
Polygenic risk score				
Low genetic risk	0 or 1	76(8)/46(4)	1.00 (Ref.)	
	2	188(21)/169(15)	0.62 (0.39-0.97)	
	3	350 (39)/386(35)	0.52 (0.34-0.78)	
	4	230(26)/348(31)	0.37 (0.24-0.57)	
	5	57(6)/165(15)	0.18 (0.11-0.31)	
	Per 1 point increase in score		0.70 (0.63-0.77)	
	P trend		<0.0001	
Intermediate genetic risk	0 or 1	127(9)/56(5)	1.00 (Ref.)	
	2	307(22)/197(17)	0.71 (0.48-1.04)	
	3	470(34)/384(34)	0.54 (0.38-0.79)	0.02
	4	363(27)/351(31)	0.50 (0.34-0.73)	
	5	101(7)/150(13)	0.35 (0.23-0.55)	
	Per 1 point increase in score		0.80 (0.74-0.87)	
	P trend		<0.0001	
High genetic risk	0 or 1	148(8)/65(6)	1.00 (Ref.)	
	2	475(24)/199(18)	1.06 (0.74-1.51)	
	3	684(35)/359(33)	0.89 (0.63-1.25)	
	4	490(25)/302(28)	0.74 (0.52-1.05)	
	5	154(8)/161(15)	0.42 (0.28-0.62)	
	Per 1 point increase in score		0.80 (0.74-0.87)	
	P trend		<0.0001	

[1]Model adjusted for matching factors age and sex; and the factors school education, family history of colorectal cancer, colonoscopy, participation in a health check-up, ever regular use of non-steroidal anti-inflammatory drugs (NSAIDs).
Abbreviations: CI: confidence intervals; OR: odds ratio; Ref.: Reference

Supplementary Table 5. Sensitivity Analysis: Recalculation of 30-year absolute risk estimates of colorectal cancer for 50 year old men and women using a relative risk (RR) estimate for colonoscopy history closer to findings of a large cohort study (RR=0.50).

Subgroup	Cases n(%)/Controls n(%)		30 Year Risk, %	
	Men	Women	Men	Women
No colonoscopy				
Low genetic risk				
Unfavourable lifestyle	164(40)/88(26)	40(15)/26(12)	5.2	4.3
Intermediate lifestyle	157(38)/123(36)	101 (38)/53(24)	4.0	3.2
Favourable lifestyle	87(21)/134(39)	123(47)/143(64)	2.9	2.2
Intermediate genetic risk				
Unfavourable lifestyle	260(41)/97(30)	76(21)/20(9)	7.9	6.2
Intermediate lifestyle	216(34)/120(37)	131(36)/78(35)	6.2	4.5
Favourable lifestyle	153(24)/108(33)	156(43)/126(56)	4.5	3.2
High genetic risk				
Unfavourable lifestyle	341(40)/115(36)	118(21)/27(13)	11.3	9.4
Intermediate lifestyle	302(35)/105(33)	199(35)/57(28)	8.9	6.9
Favourable lifestyle	208(24)/99(31)	248(44)/119(59)	6.4	4.8
Colonoscopy				
Low genetic risk				
Unfavourable lifestyle	53(36)/80(23)	7(9)/21(11)	2.7	2.2
Intermediate lifestyle	59(40)/145(41)	33(40)/65(34)	2.1	1.6
Favourable lifestyle	35(24)/129(36)	42(51)/107(55)	1.5	1.1
Intermediate genetic risk				
Unfavourable lifestyle	75(34)/110(29)	23(15)/26(12)	4.1	3.2
Intermediate lifestyle	80(36)/126(34)	60(28)/60(28)	3.2	2.3
Favourable lifestyle	67(30)/138(37)	88(57)/129(60)	2.2	1.6
High genetic risk				
Unfavourable lifestyle	120(38)/100(29)	44(21)/22(10)	5.9	4.8
Intermediate lifestyle	117(36)/127(37)	66(32)/70(32)	4.6	3.5
Favourable lifestyle	90(28)/118 (34)	98(47)/127(58)	3.3	2.4



Supplementary Figure 1. Sensitivity Analysis: Recalculation of 30-year absolute risk estimates of colorectal cancer for 50 year old men and women using a relative risk (RR) estimate for colonoscopy history closer to findings of a large cohort study (RR=0.50).

Supplementary Methods

Assessment of lifestyle factors in the DACHS study

Extensive information on smoking history was collected during interviews. Participants provided information on their current as well as prior smoking behaviour and if applicable the year in which they stopped smoking. Participants were classified as non-smokers, if they had never smoked regularly or were former smokers and smoked <30 pack years; and as smokers if they were smoking at the time of diagnosis or recruitment to the study or were former smokers and smoked ≥30 pack years (classification of former smokers based on findings from Tsoi et al [1]) (**Supplementary Methods Table 1)**. Further details have also been reported previously[2].

Information on alcohol consumption was assessed, where participants were asked how many drinks (beer [0.33L], wine [0.25L] or liquor [0.02L]) they had consumed on average per week at ages 20, 30, 40, 50, 60, 70, and 80, and in the last 12 months. Ethanol content of the beverage types (assuming 4, 8.6, and 33g of pure ethanol in 100ml of beer, wine or liquor, respectively) was derived from food composition tables and the average lifetime alcohol consumption was calculated based on self-recalled alcohol consumption at ages 20, 30, 40, 50, 60, 70, and 80 years. The mean daily lifetime amount of ethanol was calculated by dividing the total weekly ethanol amount by seven days. Participants were classified as having moderate alcohol consumption if they were adherent to the World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) recommendations: ≤24g/day men, ≤12g/day women[3] (**Supplementary Methods Table 1)**. Further details have also been reported previously[4].

Participants were asked about the hours per week they spent with different physical activities over the past decades (i.e. hard exhausting work, light work spent walking or standing, walking, cycling, or doing sports). Based on task-specific metabolic equivalent of task (MET)

values (3.3 MET-h/week for each hour walking, 6 MET-h/week for each hour cycling and 8 MET-h/week for each hour of sports), average recent non-occupational physical activity (walking, cycling or doing sports only) was calculated for each participant. Occupational activity (hard exhausting work, light work spent walking or standing) was not included in our physical activity variable given that most study participants were no longer engaged in occupational activity. Reported information from the most recent decade preceding the participants current age was used to derive the activity specific recent average MET-h/week (e.g. for patients aged 60-69, information from age 60 was used). Further details on the assessment of physical activity in the DACHS study have been reported previously[5]. Participants were classified as being physically active if they met the WHO Global Recommendations on Physical Activity for Health (at least 150 minutes of moderate-intensity or 75 minutes of vigorous-intensity aerobic physical activity throughout the week or an equivalent combination of moderate and vigorous intensity physical activity [at least ~500 MET minutes][6]) (**Supplementary Methods Table 1)**.

Dietary information was obtained by a 23-item food frequency questionnaire (FFQ), and consumption was assessed in six categories of predefined responses ranging from “never” to “multiple times per day”. Participants were asked to report their average frequency of consumption over the previous 12 months (controls) or before colorectal cancer (CRC) diagnosis (cases). A diet quality score was created based on the availability of data from the FFQ and the updated evidence from the 2017 WCRF/AICR diet recommendations for prevention of CRC[7]. Six main food groups (red and processed meat [as a negative component], fish, wholegrains, dairy foods, fruits, and vegetables [positive components]) were included in the diet quality score. Points were assigned depending on the frequency of consumption of the food groups and then summed up. The diet quality score ranged from 0 (lowest) to 50 (highest). Further details on the derivation of the diet quality score were published recently[8]**.** If information on any of the dietary items used to build the diet quality score were missing, the participants were excluded. Participants with a diet quality score in

the highest 40% were considered to have a healthy diet (**Supplementary Methods Table 1)**. Further details on the assessment of diet in the DACHS study have been published previously[9, 10].

Participants reported their current weight and height and their past weight at each decade from age 20 to 80 years. Body mass index (BMI; kg/m^2) was calculated from recent weight and height (5-14 years earlier). Participants with a BMI <18.5kg/m^2 were excluded. Participants with a healthy weight (BMI >18.5 – <25kg/m^2) were assigned one point (**Supplementary Methods Table 1)**. Further details on the assessment of BMI in the DACHS study have been published previously[11].

Estimation of absolute risk for developing CRC

We estimated the 30-year absolute risk of developing CRC for men and women separately given the healthy lifestyle score, the polygenic risk score, and colonoscopy status following the methods presented by Freedman et al[12].

The estimated absolute risk of CRC included:

1) **Estimating relative risks (RR) and attributable risk (AR) parameters from the case-control data (DACHS study data)**

Estimating the RR models

Since we were interested in calculating absolute risk estimates of CRC according to adherence to a healthy lifestyle, the polygenic risk score and colonoscopy, only these factors in addition to age (in two categories: ≤65, >65) were included in the relative risk model. We analysed all CRC cases together and used eligible controls from the DACHS study to estimate separate RR models for men and women. Odd ratios and 95% confidence intervals (CIs) were computed from unconditional logistic regression models. Statistical analyses were performed using R software version 2.15.3 (R Foundation for Statistical Computing, Vienna, Austria).

Attributable risk estimates

The age- and sex-specific attributable risks were calculated via

$$AR_j = 1 - \frac{1}{n_j}\sum_{i=1}^{n_j}\frac{1}{rr_{ij}},$$

where j is one of the four subgroups (men, ≤65 years; men, >65 years; women, ≤65 years; women, >65 years), n_j is the number of cases in the j^{th} subgroup and rr_{ij} is the relative risk of the i^{th} case in the j^{th} subgroup. The standard error of the attributable risk estimate was

calculated using the influence function approach proposed by Graubard and Fears[13] (see also Freedman et al[12]). Using the estimate of the standard error, confidence intervals were constructed assuming that distributions of the logit-transformed attributable risk are normally distributed.

2) Estimating baseline age-specific cancer hazard rates (based on the German Centre for Cancer Registry Data, incidence rates, Robert Koch Institute (the German Federal Institute within the portfolio of the Federal Ministry of Health).

Estimating the baseline age-specific CRC hazard rates

As described by Freedman et al[12], the baseline hazard rate was defined as the hazard rate for individuals each of whose risk factors are at the lowest risk level. The age-specific baseline hazard rates were computed by multiplying the age-specific incidence rates (from the Robert Koch Institute) by 1 – [the estimate of the AR] **(Supplementary Methods Table 2)**.

The age and sex-specific incidence rates for colon and rectal cancer were obtained for Germany between 2003 and 2014 **(Supplementary Methods Table 3)**. For competing risks, the mortality rates for non CRC-specific causes were calculated by subtracting the age- and sex-specific mortality rates for CRC obtained also from the German Centre for Cancer Registry Data, Robert Koch Institute from the overall mortality rates **(Supplementary Methods Table 4)**.

3) Combining competing risks, RRs and baseline hazards to estimate the probability of developing CRC over a pre-specified time period (i.e. 30 years) given a person's age and risk factors.

Absolute risk estimates

To calculate the absolute risk of CRC for a person given the age a and risk factors $\mathbf{x}$, we proceeded as follows. First, we calculated the relative risk of this person $rr(\mathbf{x})$ using the corresponding logistic model. Denoting the (sex-specific) baseline hazard rate for age k by h_{1k} and the (sex-specific) hazard rate of other-cause mortality by h_{2k}, the absolute risk of CRC in the following τ years can then be calculated using formula (2) in Pfeiffer and Petracci[14]:

$$r(a, \tau, \mathbf{x}) = \sum_{k=a}^{a+\tau-1} \frac{h_{1k} rr(\mathbf{x})}{h_{1k} rr(\mathbf{x}) + h_{2k}} [1 - \exp\{-(h_{1k} rr(\mathbf{x}) + h_{2k})\}]$$

$$\exp\left\{-\sum_{l=a}^{j-1} (h_{1l} rr(\mathbf{x}) + h_{2l})\right\}.$$

The standard error was calculated using the influence function approach proposed by Pfeiffer and Petracci[14]. Using the estimate of the standard errors, confidence intervals for the absolute risk estimates were constructed assuming that distributions of the logit-transformed absolute risk are normally distributed. To check the robustness of the obtained standard errors, we additionally calculated bootstrap standard errors for the absolute risk estimates. A comparison of the bootstrap standard errors and the standard errors obtained by the influence function approach are provided in **Supplementary Methods Table 5**.

Supplementary Methods Table 1. Description of the lifestyle factors used to derive the healthy lifestyle score

Lifestyle factor	Points	Description
Smoking	0	Smoking: current smoker or former smoker (≥ 30 pack years)
	1	Non-smoking: never smoker or former smoker (< 30 pack years)
Alcohol intake	0	Did not meet recommendations on alcoholic drinks [1]
	1	Met recommendation on alcoholic drinks[1]
Diet quality	0	Unhealthy diet quality: diet score < 34
	1	Healthy diet quality: diet score ≥ 34[2]
Physical activity	0	Did not meet physical activity guidelines[3]
	1	Met physical activity guidelines[3]
Body Mass Index	0	Overweight or obese (BMI ≥ 25kg/m^2)
	1	Healthy weight (18.5 < BMI < 25kg/m^2)

[1] World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) (2007) Recommendation on alcoholic drinks: ≤24g/day men, ≤12g/day women

[2] Diet score in the highest 40%

[3] The WHO Global Recommendations on Physical Activity for Health (2010) recommend adults to engage in at least 150 minutes of moderate-intensity or 75 minutes of vigorous-intensity aerobic physical activity throughout the week or an equivalent combination of moderate and vigorous intensity physical activity (At least ~500 MET minutes)

Supplementary Methods Table 2. Attributable risk estimates stratified by sex and age

Sex, age group	N	n(cases)	Attributable Risk (AR) (95% CI)
Men, >65	2933	1596	0.75 (0.71-0.79)
Men, ≤65	1713	988	0.78 (0.73-0.81)
Women, >65	1851	1062	0.71 (0.65-0.76)
Women, ≤65	1061	574	0.73 (0.68-0.78)

Abbreviations: CI, confidence interval

Supplementary Methods Table 3. Colorectal cancer incidence (cases per 100,000) in Germany: Data from the German Center for Cancer Registry Data, Robert Koch Institute 2003-2014

Age, y	Men	Women
30-34	3.4	3.5
35-39	6.7	6.3
40-44	12.6	11.4
45-49	26.2	21.2
50-54	53.7	39.4
55-59	103.9	63.9
60-64	170.6	95.7
65-69	248.0	136.3
70-74	328.5	186.1
75-79	414.3	253.8
80-84	496.4	337.5
85+	526.6	403.4

Supplementary Methods Table 4. Mortality (deaths per 100,000) for men and women in Germany: Data from the German Center for Cancer Registry Data, Robert Koch Institute 2003-2014

	Male			Female		
Age, y	**All-cause**	**CRC**	**Other-cause**	**All-cause**	**CRC**	**Other-cause**
30-34	74.7	0.8	73.9	32.1	0.6	31.5
35-39	105.0	1.5	103.5	57.2	1.3	55.9
40-44	172.1	3.0	169.1	94.7	2.6	92.1
45-49	305.9	6.8	299.1	170.8	5.2	165.6
50-54	521.4	14.8	506.6	274.1	10.1	263.9
55-59	804.7	29.6	775.1	412.1	17.4	394.7
60-64	1249.0	53.3	1195.7	641.9	27.8	614.1
65-69	1874.8	84.6	1790.2	942.3	43.3	899.0
70-74	2812.2	124.1	2688.1	1496.2	66.1	1430.1
75-79	5128.7	181.4	4947.3	3034.6	105.7	2928.9
80-84	8059.0	259.6	7799.4	5850.8	171.0	5679.9
85+	15551.0	360.9	15190.1	14964.1	286.8	14677.3

Abbreviations: CRC, colorectal cancer

Supplementary Methods Table 5.

Subgroup	Standard errors obtained by the influence function approach		Standard errors obtained by bootstrapping	
	Male	Female	Male	Female
No colonoscopy				
Low genetic risk				
Unfavourable lifestyle	0.0043	0.0059	0.0044	0.0060
Intermediate lifestyle	0.0033	0.0032	0.0033	0.0032
Favourable lifestyle	0.0025	0.0020	0.0025	0.0020
Intermediate genetic risk				
Unfavourable lifestyle	0.0059	0.0078	0.0060	0.0080
Intermediate lifestyle	0.0047	0.0043	0.0048	0.0044
Favourable lifestyle	0.0036	0.0026	0.0036	0.0027
High genetic risk				
Unfavourable lifestyle	0.0082	0.0114	0.0084	0.0118
Intermediate lifestyle	0.0064	0.0062	0.0066	0.0063
Favourable lifestyle	0.0050	0.0038	0.0050	0.0038
Colonoscopy				
Low genetic risk				
Unfavourable lifestyle	0.0017	0.0024	0.0017	0.0025
Intermediate lifestyle	0.0013	0.0014	0.0013	0.0014
Favourable lifestyle	0.0009	0.0008	0.0010	0.0008
Intermediate genetic risk				
Unfavourable lifestyle	0.0024	0.0032	0.0024	0.0033
Intermediate lifestyle	0.0019	0.0018	0.0019	0.0018
Favourable lifestyle	0.0014	0.0011	0.0014	0.0011
High genetic risk				
Unfavourable lifestyle	0.0033	0.0048	0.0034	0.0049
Intermediate lifestyle	0.0025	0.0027	0.0025	0.0026
Favourable lifestyle	0.0019	0.0016	0.0019	0.0016

References

1. Tsoi KK, Pau CY, Wu WK, et al. Cigarette smoking and the risk of colorectal cancer: a meta-analysis of prospective cohort studies. Clin Gastroenterol Hepatol 2009;7:682-688.e1-5.

2. Walter V, Jansen L, Hoffmeister M, et al. Smoking and survival of colorectal cancer patients: population-based study from Germany. Int J Cancer 2015;137:1433-45.

3. World Cancer Research Fund/American Institute for Cancer Research. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. Washington DC, 2007.

4. Walter V, Jansen L, Ulrich A, et al. Alcohol consumption and survival of colorectal cancer patients: a population-based study from Germany. Am J Clin Nutr 2016;103:1497-506.

5. Walter V, Jansen L, Knebel P, et al. Physical activity and survival of colorectal cancer patients: Population-based study from Germany. Int J Cancer 2017;140:1985-1997.

6. World Health Organization. Global Recommendations on Physical Activity for Health. Geneva, 2010.

7. World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Report: Diet, Nutrition, Physcial Activity and Colorectal Cancer. 2017.

8. Carr PR, Weigl K, Jansen L, et al. Healthy Lifestyle Factors Associated With Lower Risk of Colorectal Cancer Irrespective of Genetic Risk. Gastroenterology 2018;155:1805-1815 e5.

9. Carr PR, Jansen L, Walter V, et al. Associations of red and processed meat with survival after colorectal cancer and differences according to timing of dietary assessment. Am J Clin Nutr 2016;103:192-200.

10. Carr PR, Jansen L, Bienert S, et al. Associations of red and processed meat intake with major molecular pathological features of colorectal cancer. Eur J Epidemiol 2017;32:409-418.

11. Walter V, Jansen L, Hoffmeister M, et al. Prognostic relevance of prediagnostic weight loss and overweight at diagnosis in patients with colorectal cancer. Am J Clin Nutr 2016;104:1110-1120.

12. Freedman AN, Slattery ML, Ballard-Barbash R, et al. Colorectal cancer risk prediction tool for white men and women without known susceptibility. J Clin Oncol 2009;27:686-93.

13. Graubard BI, Fears TR. Standard errors for attributable risk for simple and complex sample designs. Biometrics 2005;61:847-55.

14. Pfeiffer RM, Petracci E. Variance computations for functional of absolute risk estimates. Stat Probab Lett 2011;81:807-812.