VOLUME 27 · NUMBER 5 · FEBRUARY 10 2009

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Colorectal Cancer Risk Prediction Tool for White Men and Women Without Known Susceptibility

Andrew N. Freedman, Martha L. Slattery, Rachel Ballard-Barbash, Gordon Willis, Bette J. Cann, David Pee, Mitchell H. Gail, and Ruth M. Pfeiffer

From the Division of Cancer Control and Population Sciences; Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda; Information Management Systems, Rockville, MD; School of Medicine, University of Utah, Salt Lake City, UT; and the Division of Research, Kaiser Permanente Medical Group, Oakland, CA.

Submitted April 4, 2008; accepted August 19, 2008; published online ahead of print at www.jco.org on December 29, 2008.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Andrew N. Freedman, PhD, National Cancer Institute, National Institutes of Health, EPN 4005 MSC 7344, 6130 Executive Blvd, Bethesda, MD 20892-7344; e-mail: Andrew_Freedman@nih.gov.

The Appendix is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF version (via Adobe® Reader®).

0732-183X/09/2705-686/$20.00

DOI: 10.1200/JCO.2008.17.4797.

ABSTRACT

Purpose
Given the high incidence of colorectal cancer (CRC), and the availability of procedures that can detect disease and remove precancerous lesions, there is a need for a model that estimates the probability of developing CRC across various age intervals and risk factor profiles.

Methods
The development of separate CRC absolute risk models for men and women included estimating relative risks and attributable risk parameters from population-based case-control data separately for proximal, distal, and rectal cancer and combining these estimates with baseline age-specific cancer hazard rates based on Surveillance, Epidemiology, and End Results (SEER) incidence rates and competing mortality risks.

Results
For men, the model included a cancer-negative sigmoidoscopy/colonoscopy in the last 10 years, polyp history in the last 10 years, history of CRC in first-degree relatives, aspirin and nonsteroidal anti-inflammatory drug (NSAID) use, cigarette smoking, body mass index (BMI), current leisure-time vigorous activity, and vegetable consumption. For women, the model included sigmoidoscopy/colonoscopy, polyp history, history of CRC in first-degree relatives, aspirin and NSAID use, BMI, leisure-time vigorous activity, vegetable consumption, hormone-replacement therapy (HRT), and estrogen exposure on the basis of menopausal status. For men and women, relative risks differed slightly by tumor site. A validation study in independent data indicates that the models for men and women are well calibrated.

Conclusion
We developed absolute risk prediction models for CRC from population-based data, and a simple questionnaire suitable for self-administration. This model is potentially useful for counseling, for designing research intervention studies, and for other applications.

J Clin Oncol 27:686-693. © 2008 by American Society of Clinical Oncology

INTRODUCTION

Colorectal cancer (CRC) is the third most commonly diagnosed cancer and the third leading cause of cancer death in the United States. During 2008, an estimated 148,810 new cases of CRC will be diagnosed, and 49,960 persons will die as a result of the disease.[1] Approximately one in 18 persons in the United States will develop CRC during his or her life.

Currently, several CRC screening strategies are available, including the fecal occult blood test (FOBT), double-contrast barium enema, flexible sigmoidoscopy, colonoscopy, virtual colonography, and combinations of these tests. Many of these strategies have been shown to be effective for reducing CRC incidence and mortality.[2,3]

Given the high incidence of CRC, its significant cost to society, and the availability of screening tests, a model that estimates an individual's probability of developing CRC using risk factor information that can be obtained easily in a clinical setting may aid physicians and their patients in deciding on screening regimens, and can also be useful in designing chemoprevention and screening intervention trials.[4]

Several risk and protective factors for CRC have been consistently identified in epidemiologic studies, including physical activity, cigarette smoking, and body mass index (BMI).[5,6] However, no quantitative risk model that takes competing causes of death into account is currently available to estimate the absolute risk or probability of developing CRC. Existing models are qualitative and based on expert opinion,[7] or applicable only to special populations.[8-10] We present a model that, given a set of risk and protective factors and age, predicts the absolute

risk of developing CRC over a given time period, accounting for competing causes of death. We used data from two population-based case-control studies to assess risk or protective factors. After describing the study populations, we describe the development of the CRC absolute risk model and give examples of risk estimates for various combinations of factors. We also present a short, self-administered risk assessment questionnaire that can be used to obtain information about risk factors for the model.

METHODS

Study Populations Used to Estimate Relative Risk

We used data from two population-based case-control studies, one for colon cancer[11-13] and one for rectal cancer,[14-16] to estimate relative risks (RRs) of CRC. Controls for both case-control studies were matched to cases on sex and 5-year age groups. These studies were conducted by investigators at the Universities of Utah (Salt Lake City, UT), Minnesota (Minneapolis, MN), and the Kaiser Permanente Medical Care Program (KPMCP) of Northern California (Oakland, CA). The Appendix (online only) provides additional details.

We restricted our analyses to non-Hispanic white men and women age 50 years and older, who comprised the majority of participants in both studies. We differentiated proximal (cecum through transverse colon), distal (splenic flexure, descending, and sigmoid colon), and rectal (rectosigmoid junction and rectum) tumor sites because incidence rates for these sites differ dramatically by age, and because risk factors and prior screening may have different effects on each site (Fig 1).

Risk Factors to Estimate RR Models

We assessed a variety of factors that have been consistently associated with colon or rectal cancer,[6,17,18] including age; history of CRC in first-degree relatives; history of sigmoidoscopy and/or colonoscopy; history of polyps; use of multivitamins; red meat, vegetable, and fruit consumption; alcohol intake; BMI (kg/m^2); cigarette smoking; use of aspirin and other nonsteroidal anti-inflammatory drugs (NSAIDs); current leisure-time vigorous activity; and estrogen status as assessed by menopausal status and hormone-replacement therapy (HRT) use. Although additional nutrient variables including calcium, dietary fiber, and iron have been related to CRC risk in some studies,[6,18] a detailed dietary assessment and supporting nutrient database would be needed to capture these intakes accurately, making such an assessment unfeasible in most clinical settings. Therefore, we did not include dietary variables that require a detailed assessment. The Appendix includes further description of the variables we evaluated in developing our RR models.

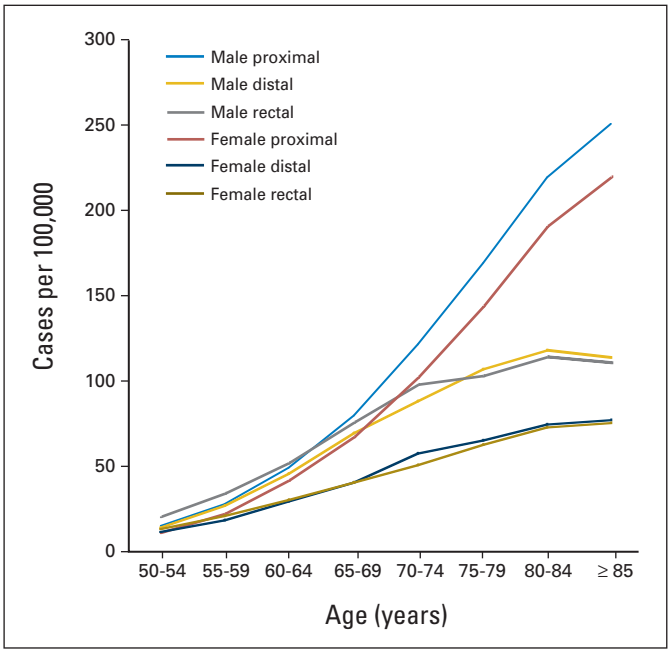

Fig 1. Colorectal cancer incidence by tumor site for white non-Hispanic men and women (13 Surveillance, Epidemiology, and End Results [SEER] sites 1992-2002).

Projecting Probabilities (absolute risk) of Developing CRC

Our approach[19] included (1) estimating RR parameters from population-based case-control data separately for proximal, distal, and rectal cancer; (2) estimating baseline age-specific cancer hazard rates (based on the National Cancer Institute's Surveillance Epidemiology and End Results (SEER) Program incidence rates) and attributable risks (ARs) from the case-control data; and (3) combining competing risks, RRs, and baseline hazards to estimate the probability of developing the first of proximal, distal, or rectal cancer over a prespecified time interval (eg, 5, 10, or 20 years) given a person's age and risk factors. The advantage of modeling the sites separately is that the covariates have different associations by various sites, and thus discriminatory accuracy can be improved by separately modeling each site.[20] The Appendix contains additional details.

Estimating the RR Models

We analyzed proximal and distal colon cancer cases separately and used all eligible controls from the colon cancer study[11-13] to estimate separate RR models. Age was included in the models in two categories (≤ 65 and > 65) when significant, to account for the matching. We determined RR estimates for all factors described earlier herein and assessed interactions among these factors and with age. Because a substantial number of participants in both studies had missing information on sigmoidoscopy/colonoscopy, we included an "unknown" category for that variable in all models. The following variables were coded as 0, 1, 2 for one *df* (trend) models: men proximal—smoking, BMI and family history; men distal—BMI and family history; men rectal—current vigorous exercise; women proximal—current vigorous exercise and family history; and women distal—family history. Odds ratios (ORs) estimating RRs and corresponding 95% CIs were computed from unconditional logistic regression models. Variable selection for inclusion in the final model was based on Wald tests for individual parameters as well as information on previously established risk factors. Statistical analyses were performed using SAS software (version 8.2; SAS Institute Inc, Cary, NC).

Estimating the Baseline Age-Specific CRC Hazard Rates

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 SEER incidence rates by 1 – [the estimate of the AR for each CRC cancer site][19] (Appendix).

The age- and sex-specific SEER incidence rates for proximal, distal, and rectal cancer (Appendix) were obtained for white non-Hispanics in 13 SEER registries between 1992 and 2002 that cover 14% of the US population.[21] Cancers of the appendix, second primary, and recurrence of colorectal cancers were not included in the computation of these rates. Competing mortality hazards from causes other than CRC were obtained from US mortality data between 1990 and 2002 (Appendix).

CIs on the Absolute Risk Estimates

We extended the influence function method of Graubard and Fears,[22] for the three outcomes (proximal, distal, or rectal cancer) to estimate the variance of the absolute risk estimate (Appendix). Approximate normality of the estimates was used to obtain 95% CIs for the estimated absolute risks.

Risk Assessment Questionnaire

We constructed a short, self-administered risk assessment questionnaire to capture the information used in the CRC risk prediction models. We tested

the questionnaire using cognitive interviewing techniques.[23] Cognitive interviews involved four "rounds" of nine participants each. We used verbal probing and think-aloud techniques to evaluate sources of misunderstanding and inaccuracy in reporting. Based on the results of our cognitive testing, the questionnaire was improved after each round to address potential sources of response errors[24] and until it was usually understandable and yielded responses that appeared to classify individuals into risk categories accurately.

RESULTS

RR Models

In our analyses, we included 1,599 colon cancer cases (665 men, 708 women) with 1,974 controls (1,058 men, 916 women) and 664 rectal cancer cases (397 men, 267 women) with 859 controls (478 men, 381 women). Table 1 displays frequencies for age and recruitment site for men and women respectively, and for proximal, distal, and rectal tumor sites.

Several factors were not related to CRC risk in our data, including FOBT; multivitamin use; alcohol use; and red meat and fruit consumption. We examined variables for smoking status and smoking pack-years for each risk model, but included only the variables "smoking duration and usual number of cigarettes smoked per day for current and former smokers" when significant, because this variable seemed to have the strongest effect of all the smoking variables, as noted previously in these studies.[25]

For men, the predictors of proximal colon cancer in the final model were prior negative sigmoidoscopy and/or colonoscopy, polyp history, number of relatives with CRC, aspirin and NSAID use, usual number of cigarettes smoked per day and years of smoking in current and former smokers, BMI, and servings of vegetables per day (Table 2). For example, men with two or more first-degree relatives with CRC were more likely to be diagnosed with proximal colon cancer than those without a positive family history (OR = 3.28; 95% CI, 1.84 to 5.84), whereas regular users of aspirin or NSAIDS were less likely to be diagnosed with proximal colon cancer (OR = 0.65; 95% CI, 0.51 to 0.82). For distal colon cancer, the final model included the same factors except for smoking and vegetable consumption (Table 2). The RR model for rectal cancer included sigmoidoscopy and/or colonoscopy, polyp history, number of relatives with CRC, NSAID use, and current vigorous leisure-time activity (Table 2). Only four controls and two cases in the rectal case-control study reported having two or more family members with CRC; therefore, family history was incorporated into the rectal RR model as "one or more relatives with CRC." None of the models showed significant two-way interactions among these variables. All three RR models for men included separate intercepts for study center.

For women, the proximal colon RR model included prior negative sigmoidoscopy and/or colonoscopy, polyp history, number of relatives with CRC, estrogen status within the last 2 years, current vigorous activity, aspirin and NSAID use, and servings of vegetables per day (Table 3). The distal colon RR model included sigmoidoscopy and/or colonoscopy and polyp history, number of relatives with CRC, aspirin and NSAID use, and estrogen status within the last 2 years, an age indicator (≥ 65 years), BMI, and an interaction between BMI and estrogen status (Table 3). Older obese women (≥ 30 kg/m^2) had an increased risk of CRC (OR = 2.68; 95% CI, 1.39 to 5.20). The main effect of BMI, however, was not statistically significant (OR = 1.08; 95% CI, 0.75 to 1.54). No other statistically significant interactions were found in any of the other models. Sigmoidoscopy and/or colonoscopy, polyp history, number of relatives with CRC, estrogen status within the last 2 years, current

Table 1. Demographic Characteristics of the Colon and Rectal Case–Control Study Populations (restricted to white men and women, age ≥ 50 years)

Characteristic	Colon Cancer Case-Control Study (1991-1994)			Rectal Cancer Case-Control Study (1997-2002)	
	Proximal Cases	Distal Cases	Controls	Rectal Cases	Controls
Men, No.					
Total	429	462	1058	397	478
Location					
Kaiser (HMO)	196	236	415	240	285
Utah (population)	86	75	239	157	193
Minnesota (population)	147	151	404	—	—
Age, years					
50-59	81	103	189	106	127
60-69	148	188	410	163	186
≥ 70	200	171	459	128	165
Women, No.					
Total	374	334	916	267	381
Location					
Kaiser (HMO)	163	154	341	155	220
Utah (population)	65	63	200	112	161
Minnesota (population)	143	117	375	—	—
Age, years					
50-59	66	69	158	69	98
60-69	118	152	341	98	130
≥ 70	187	113	417	100	153

Abbreviation: HMO, health maintenance organization.

Table 2. Relative Risk Estimates for Proximal, Distal, and Rectal Cancers for White Men Age ≥ 50 Years

Variable	Proximal		Distal		Rectal*	
	OR	95% CI	OR	95% CI	OR	95% CI
Sigmoidoscopy and/or colonoscopy and polyp history						
Sigmoidoscopy and/or colonoscopy in last 10 years, and no history of polyps	1.00		1.00		1.00	
No sigmoidoscopy and/or colonoscopy in last 10 years	1.42	1.09 to 1.88	2.83	2.10 to 3.81	3.86	2.71 to 5.48
Sigmoidoscopy and/or colonoscopy in last 10 years and history of polyps	1.77	1.17 to 2.66	1.34	0.82 to 2.21	1.92	1.07 to 3.45
Sigmoidoscopy and/or colonoscopy and polyps unknown	1.58	1.02 to 2.41	2.61	1.72 to 3.97	0.51	0.14 to 1.81
No. of relatives with CRC						
0	1.00		1.00		1.00	
1	1.81	1.35 to 2.42	1.68	1.24 to 2.27	1.49	0.91 to 2.46
≥ 2	3.28	1.84 to 5.84	2.81	1.53 to 5.16		
Current leisure-time activity, h/wk						
0					1.00	
> 0 and ≤ 2					0.83	0.72 to 0.95
> 2 and ≤ 4					0.69	0.52 to 0.90
> 4					0.57	0.38 to 0.85
Aspirin/NSAID use						
Nonuser	1.00		1.00		1.00	
Regular user	0.65	0.51 to 0.82	0.71	0.57 to 0.90	0.66	0.46 to 0.95
Smoking, cigarettes/d						
Never smoker	1.00					
> 0 and < 11	1.30	1.05 to 1.61				
≥ 11 and ≤ 20	1.70	1.11 to 2.60				
> 20	2.22	1.17 to 4.20				
Years of smoking						
0	1.00					
> 0 and < 15	0.60	0.34 to 1.06				
≥ 15 and < 35	0.88	0.50 to 1.55				
≥ 35	0.67	0.38 to 1.21				
Vegetable intake, servings/d						
< 5	1.00					
≥ 5	0.58	0.41 to 0.80				
Body mass index, kg/m^2						
≤ 24.9	1.00		1.00			
25.0 to ≤ 30	1.26	1.07 to 1.49	1.38	1.17 to 1.62		
> 30	1.59	1.14 to 2.21	1.90	1.38 to 2.61		

NOTE. Adjusted for study center. Number of cases and controls for covariates used in relative risk estimation can be found in Appendix Table A5, online only.
Abbreviations: OR, odds ratio; CRC, colorectal cancer; NSAID, nonsteroidal anti-inflammatory drug.
†The variable aspirin/NSAID use includes only NSAIDs and not aspirin for the rectal relative risk model.

vigorous leisure-time activity, aspirin and NSAID use, and BMI were all included in the rectal cancer RR model for women (Table 3). Again, because of small numbers in the rectal RR model, family history of colorectal cancer was reduced to "one or more relatives with CRC." All three RR models for women included separate intercepts for study center.

Estimates of the Baseline Age-Specific CRC Hazard Rates

To compute baseline hazard rates, we estimated separate ARs for the three models for men and women. Because the ARs did not differ by study site, we obtained combined AR estimates separately for proximal, distal and rectal cancer. The AR estimates for men were 0.86 (95% CI, 0.79 to 0.91) for proximal, 0.72 (95% CI, 0.63 to 0.80) for distal, and 0.90 (95% CI, 0.69 to 0.97) for rectal cancer. For women, the AR estimates were 0.81 (95% CI, 0.69 to 0.90) for proximal, 0.82 (95% CI, 0.73 to 0.89) for distal in women younger than 65 years, 0.85 (95% CI, 0.76 to 0.91) for distal in women age 65 years or older, and 0.93 (95% CI, 0.57 to 0.99) for rectal cancer.

Examples of Individual Absolute Risk Estimates for CRC

Table 4 presents estimates of the 10- and 20-year projected absolute risks of developing CRC for white men with various ages and risk factor profiles. The first risk profile, the lowest risk example, describes a 50-year-old man who had a colonoscopy in the last 10 years without evidence of polyps. He has no family history of CRC, vigorously exercises 5 hours/week, takes aspirin daily, never smoked, eats more than five servings of vegetables/day, and has a BMI of 24 kg/m^2. His 10-year predicted absolute risk of developing CRC is only 0.16% (95% CI, 0.11 to 0.22), and his 20-year risk is 0.53% (95% CI, 0.38 to 0.73). In risk profile 9, a high-risk example, we consider a 60-year-old man who had a colonoscopy in the last 10 years and was found to have a polyp. He has two relatives with CRC, does not exercise regularly, does not take aspirin or NSAIDS regularly, smokes more than 20 cigarettes a day, eats fewer than five servings of vegetables per day, and has a BMI of 31 kg/m^2. His 10-year predicted absolute risk of developing CRC is 7.14% (95% CI, 3.9 to 12.8), and his 20-year risk is 16.7% (95% CI, 9.1 to 28.5).

Table 3. Relative Risk Estimates for Proximal, Distal, and Rectal Cancers for White Women Age ≥ 50 Years

Variable	Proximal		Distal		Rectal	
	OR	95% CI	OR	95% CI	OR	95% CI
Sigmoidoscopy and/or colonoscopy and polyp history						
Sigmoidoscopy and/or colonoscopy in last 10 years, and no history of polyps	1.00		1.00		1.00	
No sigmoidoscopy and/or colonoscopy in last 10 years	1.82	1.32 to 2.51	3.44	2.31 to 5.11	2.99	1.91 to 4.69
Sigmoidoscopy and/or colonoscopy in last 10 years and history of polyps	2.62	1.52 to 4.50	4.35	2.35 to 8.03	3.19	1.41 to 7.25
Sigmoidoscopy and/or colonoscopy and polyps unknown	0.61	0.17 to 1.04	3.17	1.09 to 4.02	0.37	0.04 to 3.14
No. of relatives with CRC						
0	1.00		1.00		1.00	
1	1.51	1.11 to 2.03	1.45	1.04 to 2.00	1.53	0.92 to 2.55
≥ 2	2.27	1.25 to 4.14	2.09	1.09 to 4.02		
Current vigorous leisure exercise, h/wk						
0	1.00				1.00	
> 0 and ≤ 2	0.86	0.75 to 1.00			0.69	0.48 to 1.00
> 2 and ≤ 4	0.75	0.56 to 1.00			0.79	0.45 to 1.37
> 4	0.65	0.52 to 0.99			0.63	0.36 to 1.10
Aspirin/NSAID use						
Nonuser	1.00		1.00		1.00	
Regular user	0.63	0.49 to 0.81	0.70	0.53 to 0.91	0.70	0.50 to 0.97
Vegetable intake, servings/d						
< 5	1.00					
≥ 5	0.72	0.51 to 1.02				
BMI, kg/m^2						
≤ 29.9			1.00		1.00	
≥ 30			1.08	0.75 to 1.54	1.40	0.95 to 2.06
Age, years						
≤ 65			1.00			
> 65			0.55	0.41 to 0.74		
Estrogen status within the last 2 years						
Negative	1.00		1.00		1.00	
Positive	0.68	0.52 to 0.90	0.48	0.33 to 0.68	0.67	0.48 to 0.94
BMI-estrogen interaction			2.68	1.39 to 5.20		

NOTE. Adjusted for study center. Number of cases and controls for covariates used in relative risk estimation can be found in Appendix Table A6, online only.
Abbreviations: OR, odds ratio; CRC, colorectal cancer; NSAID, nonsteroidal anti-inflammatory drug; BMI, body mass index.

Similar 10- and 20-year projected absolute risks of developing CRC for white women with various ages and risk factor profiles are presented in Table 5.

Risk Assessment Questionnaire

The short, paper-based, self-administered risk assessment questionnaire we constructed to capture the information required by the models (Tables 2 and 3) requires 5 to 8 minutes to complete. A Web version of the questionnaire is available at http//www.cancer.gov/colorectalcancerrisk/.

DISCUSSION

We present a model that predicts the probability or absolute risk of developing CRC for men and women age 50 years and older. We combined separate RRs and ARs and baseline hazards for proximal, distal, and rectal cancers to project the risk of the earliest of these tumors. We also developed a short, simple, self-administered risk assessment questionnaire that can be used to obtain information for risk estimation.

In related work, we used independent data from the National Institutes of Health (NIH)-AARP Diet and Health Cohort Study[26] of men and women age 55 years and older to assess the performance of our models.[27] We found that the models had discriminatory accuracy comparable with absolute risk models for other cancers and were well calibrated.

Although the models were developed from cases and controls age 50 years and older, one could project risk for younger individuals by assuming that our relative and ARs apply to younger populations and by using younger age-specific SEER rates. However, such assumptions would need to be checked in independent data because risk factors and biologic mechanisms may differ among those developing CRC at younger ages.

Although absolute risk models exist for breast cancer and lung cancer,[19,28] this is the first such model for CRC. The four other CRC risk prediction models currently available either apply to special populations, such as patients who were referred by general practitioners to gastroenterologists for symptoms,[8] provide a qualitative index of CRC risk,[7,10] or predict different outcomes, such as the risk of having an advanced polyp or cancer in the proximal portion of the colon.[9]

Our model estimates the probability of developing CRC over a prespecified time interval from data collected from two large US population-based case-control studies of colon and rectal cancer, incidence data from 13 SEER registries, which are generally representative of the US population[29] and from national mortality rates. Thus,

Table 4. Examples of 10- and 20-Year Absolute Risk Estimates for CRC for White Men of Different Ages and Risk Factor Profiles

Patient Profile No.	Age (years)	Sigmoidoscopy and/or Colonoscopy in the Last 10 Years	Polyp in the Last 10 Years	No. of First-Degree Relatives With CRC	Current Leisure-Time Activity (h/wk)	Regular User of Aspirin/NSAIDs (at least 3 times/wk)	Smoking (usual No. of cigarettes/day)	Vegetable Intake (servings/d)	Body Mass Index (kg/m^2)	10-Year Absolute Risk		20-Year Absolute Risk	
										%	95% CI	%	95% CI
1	50	Yes	No	0	5	Yes	0	6	24	0.16	0.11% to 0.22%	0.53	0.38% to 0.73%
2	50	No	—	0	1	No	8	3	27	1.06	0.90% to 1.25%	3.48	2.92% to 4.13%
3	50	Yes	Yes	2	0	No	> 20	4	31	2.71	1.49% to 4.87%	9.14	5.02% to 16.05%
4	55	Yes	No	0	6	Yes	0	6	24	0.26	0.19% to 0.36%	0.70	0.53% to 1.04%
5	55	No	—	1	3	Yes	0	3	33	2.02	1.45% to 2.81%	5.57	3.97% to 7.76%
6	55	Yes	Yes	2	0	No	> 20	4	32	4.59	2.51% to 8.25%	12.98	7.11% to 22.50%
7	60	Yes	No	0	5	Yes	0	6	24	0.40	0.29% to 0.56%	0.93	0.66% to 1.33%
8	60	No	—	1	4	Yes	0	3	32	3.07	2.19% to 4.28%	6.97	4.93% to 9.77%
9	60	Yes	Yes	2	0	No	> 20	2	31	7.14	3.89% to 12.77%	16.65	9.11% to 28.46%
10	65	Yes	No	0	5	Yes	0	6	24	0.54	0.38% to 0.77%	1.07	0.74% to 1.55%
11	65	No	—	0	1	No	10	2	26	3.51	2.89% to 4.26%	6.78	5.50% to 8.32%
12	65	Yes	Yes	2	0	No	> 20	4	31	9.95	5.37% to 17.8%	19.4	10.6% to 32.7%

NOTE. Categories for each variables used in the CRC absolute risk model: sigmoidoscopy and/or colonoscopy and polyp history in the last 10 years—no, sigmoidoscopy and/or colonoscopy; yes, sigmoidoscopy and/or colonoscopy, no polyps; yes, sigmoidoscopy and/or colonoscopy, yes polyps; unknown sigmoidoscopy and/or colonoscopy and polyps. Number of relatives with CRC—0, no relatives; 1 relative; ≥ 2 relatives. Current leisure-time activity—0 h/wk; 0 to 2 h/wk; 2 to 4 h/wk; > 4 h/wk. Regular user of aspirin/NSAIDs—regular user, at least 3 times/wk; nonuser, less than 3 times/wk. Smoking—usual number of cigarettes/d when smoking: 0, never smoker; 1 to 10 cigarettes/d; 11 to 20 cigarettes/d, > 20 cigarettes/d. Vegetable intake—< 5 servings/d; 5+ servings/d. Body mass index—underweight and normal weight < 30 kg/m^2; overweight ≥ 30 kg/m^2.
Abbreviations: CRC, colorectal cancer; NSAID, nonsteroidal anti-inflammatory drug.

our risk prediction models would be expected to apply to the general non-Hispanic white US population.

We used factors in our models that, in addition to having strong predictive ability, can also be ascertained easily in a clinical setting. Thus, we did not include some factors that may have predictive value, such as and calcium intake or long-term vigorous activity[30-33] but which would require a much more complex questionnaire.[34]

Our risk prediction model has some limitations. Because the majority of participants in the case-control studies were white, we could not estimate RRs for other racial or ethnic groups. A first step to developing models for other racial/ethnic groups could be to combine RR and AR estimates for whites with SEER rates for blacks, Asians, or Hispanics. However, the assumption of constant AR and RR estimates across racial groups needs to be validated in minority populations. Our model is not applicable to individuals with ulcerative colitis, Crohn's disease and familial adenomatous polyposis, because these conditions carry a high risk of CRC, and individuals with these conditions were excluded from the studies. Additionally, our model is not applicable to individuals with hereditary nonpolyposis CRC.

Because we used US mortality data from 1990 to 2002 for our competing mortality hazards, we did not adjust these estimates for potential confounders such as BMI given that our sensitivity analyses

Table 5. Examples of 10- and 20-Year Absolute Risk Estimates for CRC for White Women of Different Ages and Risk Factor Profiles

Patient Profile No.	Age (years)	Sigmoidoscopy and/or Colonoscopy in the Last 10 Years	Polyp in the Last 10 Years	No. of First-Degree Relatives With CRC	Current Leisure-Time Activity (h/wk)	Regular User of Aspirin/NSAIDs	Vegetable Intake (servings/d)	Body Mass Index (kg/m^2)	Estrogen Status Within the Last 2 Years	10-Year Absolute Risk		20-Year Absolute Risk	
										%	95% CI	%	95% CI
1	50	Yes	No	0	5	Yes	6	28	Positive	0.09	0.07% to 0.13%	0.31	0.22% to 0.43%
2	50	No	—	1	3	Yes	7	29	Negative	0.71	0.50% to 1.01%	2.25	1.61% to 3.13%
3	50	Yes	Yes	2	0	No	3	32	Negative	2.51	1.50% to 4.19%	8.11	4.89% to 13.17%
4	55	Yes	No	0	5	Yes	6	28	Positive	0.15	0.11% to 0.21%	0.43	0.31% to 0.62%
5	55	No	—	1	1	No	2	31	Positive	1.85	1.30% to 2.64%	4.99	3.55% to 6.98%
6	55	Yes	Yes	2	0	No	3	32	Negative	4.13	2.46% to 6.84%	11.38	6.88% to 18.23%
7	60	Yes	No	0	5	Yes	6	28	Positive	0.22	0.16% to 0.32%	0.57	0.40% to 0.82%
8	60	No	—	1	1	No	3	32	Positive	2.61	1.85% to 3.68%	6.37	4.57% to 8.82%
9	60	Yes	Yes	2	0	No	3	32	Negative	6.01	3.60% to 9.87%	14.72	8.94% to 23.29%
10	65	Yes	No	0	5	Yes	6	28	Positive	0.30	0.21% to 0.44%	0.69	0.48% to 1.00%
11	65	No	—	1	4	Yes	6	28	Negative	2.13	1.54% to 2.94%	4.64	3.38% to 6.35%
12	65	Yes	Yes	2	0	No	3	32	Negative	8.16	4.89% to 13.31%	17.49	10.66% to 27.37%

NOTE. Categories for each variables used in the CRC absolute risk model: sigmoidoscopy and/or colonoscopy and polyp history in the last 10 years—no, sigmoidoscopy and/or colonoscopy; yes, sigmoidoscopy and/or colonoscopy, no polyps; yes, sigmoidoscopy and/or colonoscopy, yes polyps; unknown sigmoidoscopy and/or colonoscopy and polyps. Number of relatives with CRC—0, no relatives; 1 relative; ≥ 2 relatives. Current leisure-time activity—0 h/wk; 0 to 2 h/wk; 2 to 4 h/wk; > 4 h/wk. Regular user of aspirin/NSAIDs—regular user, at least 3 times/wk; nonuser, less than 3 times/wk. Vegetable intake—< 5 servings/d; 5+ servings/d. Body mass index—underweight and normal weight < 30 kg/m^2; overweight ≥ 30 kg/m^2. Estrogen status within the last 2 years—negative; positive.
Abbreviations: CRC, colorectal cancer; NSAID, nonsteroidal anti-inflammatory drug.

indicated that changes in the risk estimates were small (data not shown). Although our two case-control studies were conducted at slightly different time periods, we believe any changes in the distribution of risk factors would have a minimal effect on our risk estimates, considering that RRs and ARs were estimated separately for the two studies.

Another limitation of our model is that we estimated our RRs and ARs from case-control rather than from cohort studies. Although case-control studies have been used in the development of risk prediction models for melanoma,[35] breast,[19,36,37] bladder,[38] and lung cancer,[39] a general criticism is that such estimates could be subject to recall bias. However, recall bias likely plays a minor role in our models as most of the RRs we found, including BMI, physical activity, HRT, and aspirin and NSAIDs use, were consistent with RRs summarized in a recent comprehensive review of the epidemiologic literature.[6] Although not covered in this review, our risk estimates for screening and polyp history,[2,40,41] smoking,[42] and family history[43,44] are also consistent with many previously published findings, including results from cohort studies. Additionally, the models were well calibrated in an independent validation study using the AARP cohort.[27]

In summary, we developed an absolute CRC risk projection model for white men and women age 50 years or older that may aid physicians and their patients in deciding on screening regimens, and can also be useful in designing chemoprevention and screening intervention trials.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design: Andrew N. Freedman, Martha L. Slattery, Rachel Ballard-Barbash, Mitchell H. Gail, Ruth M. Pfeiffer
Financial support: Andrew N. Freedman, Rachel Ballard-Barbash
Administrative support: Andrew N. Freedman, Martha L. Slattery, Rachel Ballard-Barbash, David Pee
Provision of study materials or patients: Martha L. Slattery, Bette J. Cann
Collection and assembly of data: Martha L. Slattery, Gordon Willis, Bette J. Cann
Data analysis and interpretation: Andrew N. Freedman, Martha L. Slattery, Rachel Ballard-Barbash, Gordon Willis, Bette J. Cann, David Pee, Mitchell H. Gail, Ruth M. Pfeiffer
Manuscript writing: Andrew N. Freedman, Martha L. Slattery, Rachel Ballard-Barbash, Gordon Willis, Bette J. Cann, Mitchell H. Gail, Ruth M. Pfeiffer
Final approval of manuscript: Andrew N. Freedman, Martha L. Slattery, Rachel Ballard-Barbash, Gordon Willis, Bette J. Cann, David Pee, Mitchell H. Gail

REFERENCES

1. Cancer Facts and Figures 2008. Atlanta, GA, American Cancer Society, 2008

2. Winawer SJ, Zauber AG, Fletcher RH, et al: Guidelines for colonoscopy surveillance after polypectomy: A consensus update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society. Gastroenterology 130:1872-1885, 2006

3. Levin B, Lieberman DA, McFarland B, et al: Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: A joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. CA Cancer J Clin 58:130-160, 2008

4. Freedman AN, Seminara D, Gail MH, et al: Cancer risk prediction models: A workshop on development, evaluation, and application J Natl Cancer Inst 97:715-723, 2005

5. Giovannucci E: Modifiable risk factors for colon cancer. Gastroenterol Clin North Am 31:925-943, 2002

6. World Cancer Research Fund (WCRF) Panel (Marmot M, Chair): Food, Nutrition, Physical Activity and the Prevention of Cancer: A global perspective. Washington, DC, WCRF/American Institute of Cancer Research, 2007

7. Colditz GA, Atwood KA, Emmons K, et al: Harvard Report on Cancer Prevention Volume 4: Harvard Cancer Risk Index—Risk Index Working Group, Harvard Center for Cancer Prevention. Cancer Causes Control 11:477-488, 2000

8. Selvachandran SN, Hodder RJ, Ballal MS, et al: Prediction of colorectal cancer by a patient consultation questionnaire and scoring system: A prospective study. Lancet 360:278-283, 2002

9. Imperiale TF, Wagner DR, Lin CY, et al: Using risk for advanced proximal colonic neoplasia to tailor endoscopic screening for colorectal cancer. Ann Intern Med 139:959-965, 2003

10. Driver JA, Gaziano JM, Gelber RP, et al: Development of a risk score for colorectal cancer in men. Am J Med 120:257-263, 2007

11. Slattery ML, Edwards SL, Ma KN, et al: Colon cancer screening, lifestyle, and risk of colon cancer. Cancer Causes Control 11:555-563, 2000

12. Slattery ML, Ballard-Barbash R, Edwards S, et al: Body mass index and colon cancer: An evaluation of the modifying effects of estrogen. Cancer Causes Control 14:75-84, 2003

13. Coates AO, Potter JD, Caan BJ, et al: Eating frequency and the risk of colon cancer. Nutr Cancer 43:121-126, 2002

14. Slattery ML, Curtin KP, Edwards SL, et al: Plant foods, fiber, and rectal cancer. Am J Clin Nutr 79:274-281, 2004

15. Murtaugh MA, Ma KN, Benson J, et al: Antioxidants, carotenoids, and risk of rectal cancer. Am J Epidemiol 159:32-41, 2004

16. Slattery ML, Caan BJ, Benson J, et al: Energy balance and rectal cancer: An evaluation of energy intake, energy expenditure, and body mass index. Nutr Cancer 46:166-171, 2003

17. Potter JD: Colorectal cancer: Molecules and populations. J Natl Cancer Inst 91:916-932, 1999

18. Tomeo CA, Colditz GA, Willett WC, et al: Harvard Report on Cancer Prevention Volume 3: Prevention of colon cancer in the United States. Cancer Causes Control 10:167-180, 1999

19. Gail MH, Brinton LA, Byar DP, et al: Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 81:1879-1886, 1989

20. Colditz GA, Rosner BA, Chen WY, et al: Risk factors for breast cancer according to estrogen and progesterone receptor status. J Natl Cancer Inst 96:218-228, 2004

21. Surveillance Epidemiology and End Results (SEER): http://seer.cancer.gov/

22. Graubard BI, Fears TR: Standard errors for attributable risk for simple and complex sample designs. Biometrics 61:847-855, 2005

23. Willis GB: Cognitive Interviewing: A Tool for Improving Questionnaire Design. Thousand Oaks, CA, SAGE Publications, 2004

24. Smith T, Dean D: Colorectal Cancer Health Assessment Testing of a Self-Administered Survey: Summary Report—Technical Report Prepared for the National Cancer Institute. Rockville, MD, Westat, 2005

25. Slattery ML, Potter JD, Friedman GD, et al: Tobacco use and colon cancer. Int J Cancer 70:259-264, 1997

26. NIH-ARP Diet and Health Study. http://dietandhealth.cancer.gov/

27. Park Y, Freedman AN, Gail MH, et al: Validation of a colorectal cancer risk prediction model among whites 50 years old and over. J Clin Oncol 27:694-698, 2009

28. Bach PB, Kattan MW, Thornquist MD, et al: Variations in lung cancer risk among smokers. J Natl Cancer Inst 95:470-478, 2003

29. Hankey BF, Ries L, Edwards BK: The Surveillance, Epidemiology, and End Results Program: A national resource. CEBP 8:1117-1121, 1999

30. Slattery M: Physical activity and colorectal cancer. Sports Med 34:239-252, 2004

31. Slattery ML, Edwards SL, Ma KN, et al: Physical activity and colon cancer: A public health perspective. Ann Epidemiol 7:137-145, 1997

32. Lee IM, Paffenbarger RS Jr, Hsieh C: Physical activity and risk of developing colorectal cancer among college alumni. J Natl Cancer Inst 83:1324-1329, 1991

33. Severson RK, Nomura AM, Grove JS, et al: A prospective analysis of physical activity and cancer. Am J Epidemiol 130:522-529, 1989

34. Ferrari P, Friedenreich C, Matthews CE: The role of measurement error in estimating levels of physical activity. Am J Epidemiol 166:832-840, 2007

35. Fears TR, Guerry D IV, Pfeiffer RM, et al: Identifying individuals at high risk of melanoma: A practical predictor of absolute risk. J Clin Oncol 24:3590-3596, 2006

36. Gail MH, Costantino JP, Pee D, et al: Projecting individualized absolute invasive breast cancer risk in African American women. J Natl Cancer Inst 99:1782-1792, 2007

37. Claus EB, Risch N, Thompson WD: The calculation of breast cancer risk for women with a first degree family history of ovarian cancer. Breast Cancer Res Treat 28:115-120, 1993

38. Wu X, Lin J, Grossman HB, et al: Projecting individualized probabilities of developing bladder cancer in white individuals. J Clin Oncol 25:4974-4981, 2007

39. Spitz MR, Hong WK, Amos CI, et al: A risk model for prediction of lung cancer. J Natl Cancer Inst 99:715-726, 2007

40. Martínez ME, Sampliner R, Marshall JR: Adenoma characteristics as risk factors for recurrence of advanced adenomas. Gastroenterology 120:1077-1083, 2001

41. Bonithon-Kopp C, Piard F, Fenger C, et al: Colorectal adenoma characteristics as predictors of recurrence. Dis Colon Rectum 47:323-333, 2004

42. Giovannucci E, Martinez ME: Tobacco, colorectal cancer, and adenomas: A review of the evidence. J Natl Cancer Inst 88:1717-1730, 1996

43. St John DJ, McDermott FT, Hopper JL, et al: Cancer risk in relatives of patients with common colorectal cancer. Ann Intern Med 118:785-790, 1993

44. Fuchs CS, Giovannucci EL, Colditz GA, et al: A prospective study of family history and the risk of colorectal cancer. N Engl J Med 331:1669-1674, 1994

Acknowledgment

We thank Anne Rodgers for her helpful comments and editorial assistance.

