# Uncontrolled Confounding in Studies of Screening Effectiveness: An Example of Colonoscopy

**Ronald C. Eldridge, MPH**[1], **Chyke A. Doubeni, MD, MPH**[2], **Robert H. Fletcher, MD, MSc**[3], **Ann G. Zauber, PhD**[4], **Douglas A. Corley, MD, PhD, MPH**[5], **V. Paul Doria-Rose, DVM, PhD**[6], and **Michael Goodman, MD, MPH**[1]

[1]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA

[2]Department of Family Medicine and Community Health, the Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

[3]Department of Population Health, Harvard Medical School, Boston, MA

[4]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY

[5]Kaiser Permanente Division of Research, Oakland, CA

[6]Applied Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD

## Abstract

**Background**—The purpose of this communication is to estimate the expected magnitude of error produced by uncontrolled confounding from health behaviors in observational medical record-based studies evaluating effectiveness of screening colonoscopy.

**Methods**—We used data from the prospective NIH-AARP Diet and Health Study to assess the impact of health behavior related factors (lifestyle, education, and use of non-steroidal anti-inflammatory drugs [NSAID]) on the association between colonoscopy and CRC mortality. We first examined the difference between adjusted and unadjusted results within the cohort data, and then estimated a broader range of likely confounding errors based on the Breslow-Day approach that uses prevalence of confounders among persons with and without exposure, and the rate ratio reflecting the association between these confounders and the outcome of interest. As dietary factors and habits are often inter-correlated, we combined these variables (physical activity, body mass index, waist-to-hip ratio, alcohol consumption, and intakes of red meat, processed meat, fiber, milk, and calcium) into a "healthy lifestyle score" (HLS).

**Results**—The estimated error (a ratio of biased-to-true result) attributable to confounding by HLS was 0.959–0.997 indicating less than 5% departure from the true effect of colonoscopy on CRC mortality. The corresponding errors ranged from 0.970 to 0.996 for NSAID, and from 0.974 to 1.006 for education (all 3% difference). The results for other CRC screening tests were similar.

**Conclusion**—Health behavior-related confounders, either alone or in combination, seem unlikely to strongly affect the association between colonoscopy and CRC mortality in observational studies of CRC screening.

Corresponding author Michael Goodman, MD, MPH, Department of Epidemiology, Emory University School of Public Health, 1518 Clifton Rd., NE, Atlanta, GA 30322, Tel: 404.727.2734, Fax: 404.727.8737.

# INTRODUCTION

Colonoscopy is the most commonly used procedure for colorectal cancer (CRC) screening in the US.[1] Unlike other tests currently recommended for CRC screening,[1, 2] such as fecal occult blood testing (FOBT)[3–5] and flexible sigmoidoscopy,[6–8] its use is not currently supported by evidence from randomized trials. Although randomized controlled trials of screening colonoscopy are under way, their results will not be available until at least 2022–2025.[9–11] Therefore, in the foreseeable future, colonoscopy screening policy will be based on evidence from observational studies, where unmeasured confounding is always a concern.[12–14]

Lifestyle factors, such as physical activity, diet, and alcohol consumption are associated with CRC incidence and, to a lesser extent, CRC mortality.[15–17] Studies also show that use of aspirin and other non-steroidal anti-inflammatory drugs (NSAIDs) are associated with a decrease in CRC incidence and mortality.[18] Another factor demonstrably associated with CRC mortality is education.[19–21] Lifestyle factors, NSAID use and education are also presumed to be related to healthier behaviors that include screening practices, and for this reason all three may act as confounders in the association between colonoscopy and mortality.[22–30] However, these variables are often not available in medical records.

The current study investigates to what extent unmeasured confounding by lifestyle-related risk factors, education and the use of NSAIDs may affect the association between colonoscopy and other CRC screening methods and mortality in observational studies. These studies of this question should help future research, and may also be useful in interpreting the previously published observational studies.[13, 25, 31]

# METHODS

## Overview

The three categories of possible confounders examined in the present study are lifestyle (including diet and habits), education, and use of NSAID. As dietary factors and habits are often inter-correlated, we combined there variables into a "healthy lifestyle score" (HLS).

The data for analyses came from the prospective NIH-AARP Diet and Health Study. We first estimated the differences in results for crude, partially adjusted, and fully adjusted models evaluating the association between colonoscopy and mortality within the NIH-AARP cohort. This approach allowed us to assess the impact of confounding that is specific to the NIH-AARP population.

To allow generalization of results beyond the NIH-AARP cohort we estimated a range of likely confounding errors based on the Breslow-Day method.[32] With the Breslow-Day approach, magnitude of confounding error is calculated based on the difference in prevalence of the confounder among persons with and without exposure of interest (denoted $P_1$ and $P_0$, respectively), and based on the rate ratio ($RR_{confounder}$) reflecting the association between the confounder and the outcome of interest (in this case CRC mortality). The input parameters for these calculations were obtained from the NIH-AARP data. To provide a range of possible confounding error we used the upper and lower 95% limits rather than point estimates in all calculations.

## Study Population

Established in 1995–1996, the prospective NIH-AARP Diet and Health Study mailed out 3.5 million questionnaires to persons between the ages of 50 and 71 years. The questionnaire asked about demographic characteristics, diet, and other health related behaviors.[33] A

follow-up questionnaire that included questions on cancer screening practices was mailed in 1996–97 to 542,095 subjects who responded to the initial mail-out and 334,906 valid questionnaires were returned.

We excluded cohort members if they: had questionnaires filled out by a proxy (n=10,383); reported poor health (n=4,983) or a history of cancer (n=4,391); provided uninformative or contradictory responses to the CRC screening questions (n=10,286); or had missing data on lifestyle score factors (n=147,941) or aspirin/NSAID use (n=1,482). Thus, the final analytic data set contained 155,440 subjects.

### Data on Colorectal Cancer Screening

The NIH-AARP study collected information on whether, in the past three years, a participant underwent testing for blood in the stool and if he or she had received any of the following procedures: flexible sigmoidoscopy, colonoscopy, proctoscopy, or an endoscopy of unidentified type.

These items were used to define a dichotomous variable "colonoscopy", and two additional derived variables categorized participants with respect to their receipt of "FOBT" and "any CRC screening" procedure. Each screening procedure was compared to the "no screening" reference category

### Colorectal Cancer Deaths & Follow-up Time

Causes of death were ascertained by linking cohort data to the National Death Index through December 31, 2008. Participants who died from CRC were identified using the International Classification of Diseases 9th edition (ICD-9) codes 153.0–154.8 or ICD-10 codes C18.0–C20.0 for underlying cause of death. The time under observation (mean=11 years) was calculated by subtracting each subject's date of return for the follow-up questionnaire from the date of death, date of move outside of cohort residency, or December 31, 2008, whichever was the earliest. A total of 602 cohort members died from CRC with a cumulative 10-year mortality estimate of 0.4%, which is in agreement with national estimates.[34]

### Healthy Lifestyle Score

As lifestyle characteristics are often inter-correlated, we combined candidate variables into a single *a priori* "healthy lifestyle score" (HLS) and used the resulting score as the confounder of interest. The findings of the 2007 expert report issued jointly by the American Institute for Cancer Research (AICR) and the World Cancer Research Fund (WCRF)[15] served as the basis for selecting the candidate risk factors. The AICR/WCRF report identified ten lifestyle-related factors for which the association with CRC incidence was deemed to be "convincing" or "probable". Nine of these factors – body mass index (BMI), waist-to-hip ratio, intakes of alcohol, red meat or processed meat products, physical activity, and consumption of milk, calcium and fiber – were incorporated into the HLS. Garlic intake was also listed in the AICR/WCRF report, but was not included in the score because it was not measured in the NIH-AARP questionnaire.

The nine variables were combined into a one-dimensional score, henceforth called the "unweighted HLS". Each of the nine components was divided into gender-specific quartiles where the "healthiest" quartile of each variable (lowest for the detrimental factors and highest for the beneficial factors) was given 4 points, and the "unhealthiest" quartile was given 1 point. The HLS quartile cutoffs for each variable were based on the total study population (n=304,863), not the analytical cohort (n=155,440), but the differences between the two sets of cutoffs were minimal. The nine components were then summed to produce

the overall unweighted score with a possible range from 9 through 36, where higher HLS values were hypothesized to confer a lower risk of CRC.

For the alternative version of the score, henceforth called the "weighted HLS", we combined diet- and body composition-related variables into separate sub-scores and the two sub-scores were then included along with physical activity into the overall HLS. With the exception of physical activity, all other variables in the weighted HLS were expressed for each study subject as the number of standard errors between the mean and the observed value. The standardization was performed separately for men and women. The sex-specific standardized variables for BMI and waist-to-hip ratio were summed to form the body composition sub-score, multiplied by −1 to maintain consistent direction of the hypothesized association, and the resulting sub-score was divided into quartiles. A similar approach was used to develop a diet sub-score that was based on intakes of red and processed meat, milk, calcium, fiber and alcohol. The standardized values for harmful factors (red and processed meat, alcohol intake) were also multiplied by −1.

## Statistical Analyses

To assess the impact of confounding on the association between colonoscopy and CRC mortality specifically in the NIH-AARP cohort, we constructed Cox PH models that evaluated this association in several ways: 1) without any adjustment, 2) adjusting for a partial list of covariates that included age, sex and hormone replacement therapy use (HRT) coded as a single three-category variable (male, female with HRT and female without HRT), education, race, diabetes, family history of CRC), 3) by adding HLS, education, and NSAID use to the model, and 4) by including all covariates in the model. Another set of analyses was conducted for FOBT and for "any screening".

## Estimating Impact of Confounding using Breslow-Day method

Breslow and Day[32] mathematically describe the effect of a dichotomous confounding variable on the association between dichotomous exposure and outcome variables as:

$$RR_{biased} = RR_{true} \times \frac{(1-P_1) + P_1 \times RR_{confounder}}{(1-P_0) + P_0 \times RR_{confounder}} \quad (1)$$

In the above equation $P_1$ and $P_0$ are the proportions of subjects with the confounder among the exposed and the unexposed, respectively; $RR_{biased}$ denotes the risk or rate ratio that would have been observed in a study that fails to control for the confounder of interest, $RR_{true}$ is the corresponding estimate that fully adjusts for that confounder; and $RR_{confounder}$ reflects the association between the confounder and the outcome of concern. The magnitude and direction of error introduced by unaccounted confounding can be expressed as:

$$Error = \frac{RR_{biased}}{RR_{true}} = \frac{(1-P_1) + P_1 \times RR_{confounder}}{(1-P_0) + P_0 \times RR_{confounder}}, \quad (2)$$

When $RR_{true} = 1$, the entire observed association between exposure and outcome ($RR_{biased}$) is explained by confounding (Error = $RR_{biased}$). If the confounder of interest is not dichotomous, but is represented by an ordinal variable (e.g., divided into quartiles) the above equation is extended as follows:

$$\frac{\text{RR}_{\text{biased}}}{\text{RR}_{\text{true}}} = \frac{(1-P_1^2-P_1^3-P_1^4)+P_1^2 \times \text{RR}_{\text{confounder}}^2+P_1^3 \times \text{RR}_{\text{confounder}}^3+P_1^4 \times \text{RR}_{\text{confounder}}^4}{(1-P_0^2-P_0^3-P_0^4)+P_0^2 \times \text{RR}_{\text{confounder}}^2+P_0^3 \times \text{RR}_{\text{confounder}}^3+P_0^4 \times \text{RR}_{\text{confounder}}^4}, \quad (3)$$

where the superscripts denote quartile-specific values of $\text{RR}_{\text{confounder}}$, $P_1$ and $P_0$.

If the impact of an unmeasured confounder is unknown, one can calculate a plausible range of confounding error expressed as the ratio $\text{RR}_{\text{biased}}/\text{RR}_{\text{true}}$ or simply $\text{RR}_{\text{biased}}$ (if $\text{RR}_{\text{true}}$ is assumed to be 1.0) by estimating a range of likely values for $\text{RR}_{\text{confounder}}$, $P_1$ and $P_0$. In the current study, the exposure of interest is colonoscopy, the outcome of concern is CRC mortality and the confounders under consideration are a composite measure of lifestyle termed the "healthy lifestyle score' (HLS) and regular (at least weekly) use of NSAIDs. All parameters and their ranges are estimated using the data from the NIH-AARP Diet and Health Study cohort.

To obtain rate ratio estimates for the relation between HLS and CRC mortality (denoted $\text{RR}_{\text{confounder}}$ in equations 1, 2 and 3), we divided the un-weighted HLS into ordinal quartiles and the weighted HLS into alternative non-quartile *a priori* cutoffs. For each HLS score, we examined the association with CRC mortality using Cox proportional hazards (PH) models that adjusted for age, sex, HRT, education, race, diabetes, family history of CRC, aspirin/NSAID use and any CRC screening.[35] Each $\text{RR}_{\text{confounder}}$ estimate was accompanied by the corresponding 95% confidence interval (CI). All modeling was performed using SAS 9.2 statistical software, and all models were examined for PH assumptions, interactions, collinearity, and goodness-of-fit.

We also obtained estimates of $P_1$ and $P_0$ (i.e., proportions of persons in each ordinal category of HLS values among the colonoscopy screened and non-screened study subjects, respectively) and their corresponding 95% CIs using OpenEpi statistical calculator.[36] We then used the upper and lower limits of the 95% CI for $\text{RR}_{\text{confounder}}$, and $P_1$ and $P_0$ estimates to calculate a range of likely confounding errors introduced by HLS (equation 3) and by the dichotomous aspirin/NSAID variable (equation 2). The same analyses were also performed for FOBT and for "any screening".

## RESULTS

### Assessment of confounding within the NIH-AARP cohort

As shown in Table 1, approximately 62% of study subjects were male, most cohort members (95%) were whites, and 70% had at least some college education. About 70% of the analytic cohort received at least one of the CRC screening tests. Notably, only 15% underwent a colonoscopy within three years prior to questionnaire administration, which is in agreement with national data.[37]

The observed effects of confounding by HLS and NSAID use on the association between colonoscopy and CRC mortality in NIH-AARP cohort were rather small. The unadjusted RR for a Cox proportional hazards model comparing colonoscopy screening in the previous three years to no screening was 0.465 (95% CI: 0.346–0.625). Controlling for all covariates with the exception of HLS and NSAID use, produced a partially adjusted RR of 0.402 (95% CI: 0.297–0.545) indicating that colonoscopy was associated with a 60% reduction in CRC mortality over 10 years. After including HLS into the model, the RR was 0.405 (95% CI: 0.299–0.549) and did not differ by HLS type (error 0.402/0.405=0.993). The results with and without adjustment for NSAID use produced an identical error estimate. When adjusting for both HLS and NSAID use, the error estimate was 0.990 (0.402/0.406). Analyses for FOBT and "any screening" produced very similar results (data not shown).

The individual components of the HLS and their association with colonoscopy are reported in Tables 2 and 3. For the variables that make up the HLS, red meat consumption, fiber consumption, and physical activity appeared to demonstrate the greatest differences by colonoscopy status. More frequent NSAID use was also more common in colonoscopy-screened study participants than in their non-screened counterparts.

The unweighted HLS was inversely associated with CRC mortality. When comparing the "healthier" categories 20–22, 23–25 and 26+ to the reference category <20 the adjusted $RR_{confounder}$ values (95% CIs) were 0.88 (0.71–1.09); 0.66 (0.52–0.82) and 0.57 (0.46–0.72) respectively. The results for the weighted HLS were similar (Table 4). Regular (at least weekly) NSAID use was also significantly associated with decreased CRC mortality (RR=0.80; 95% CI, 0.68–0.94).

### Assessment of Confounding using Breslow-Day Method

As shown in Table 5, the plausible range of Breslow-Day error estimates ($RR_{biased}/RR_{true}$) was generally small. In the analyses for HLS the the range of error was 0.960–0.995; and 0.959–0.997 for the unweighted and weighted score, respectively. These estimates remained virtually the same when colonoscopy was replaced with "any screening" or FOBT. In the analyses for NSAID use, the plausible range of error for the association between colonoscopy and CRC mortality was between 0.970 and 0.996. The corresponding errors for education ranged from 0.974 to 1.006. Assuming the confounding biases for HLS, aspirin/ NSAID use and education are independent and combine in a multiplicative fashion, the $RR_{biased}/RR_{true}$ estimates resulting from failure to control for all three factors would range from 0.906 (0.959×0.970×0.974) to 0.999 (0.997×0.996×1.006), which correspond to a range of errors from 0.1% to 9.4%.

## DISCUSSION

All calculated confounding error estimates in this study were small, and almost all in the hypothesized direction. On balance, our data and the results reported by others[16, 17, 38–40] indicate that it would take unrealistic differences between $P_1$ and $P_0$, and very pronounced $RR_{confounder}$ values to produce a confounding error of substantial magnitude. For example, for a dichotomous confounding factor to produce an error of 20%, assuming $RR_{confounder}$ of 0.7, $P_0$ of 10%, 20% and 30%, the factor would require $P_1–P_0$ differences of 65% (75%– 10%), 63% (83–20%) and 61% (91%–30%), respectively. The same calculations for a much stronger $RR_{confounder}$ of 0.5 would require the corresponding $P_1–P_0$ differences of 38%, and 36% and 34%, lower, but still clearly outside the realistically expected range.

Our analyses also showed that confounders such as age, sex, HRT, race, and family history of CRC should always be considered because in the AARP data they changed the RR estimate by 13.5% (from 0.465 to 0.402). It is important to point out, however, that in contrast to HLS and NSAID use, these other factors are less likely to remain uncontrolled because they usually can be ascertained from the medical records and/or most questionnaires.

The main limitation of the current analysis is the large number of subjects excluded from the original cohort because of missing information due to incompletely filled out study questionnaires. It is important to note, however, that quartile cutoffs were based on the entire underlying study population, and yet for all HLS components, except BMI, the final analysis cohort was still divided into four almost equal groups, indicating that the loss of data may have occurred at random.

Ascertainment of screening practices was limited to three years prior to questionnaire administration, a feature that may have led to misclassification of screening status. Subjects who underwent a colonoscopy or sigmoidoscopy within the recommended screening timeframe, but outside the three-year window or were screened after the date of the questionnaire would have been classified as unscreened, likely biasing our $RR_{confounder}$ estimates towards the null. The analysis for FOBT produced results with respect to the role of confounding almost identical to those observed for colonoscopy. On the other hand, the FOBT data in this study do not indicate whether the test was completed at home or at the physician's office (a previously common practice that is now deemed inadequate).

Another important limitation that may have affected the observed association between colonoscopy and CRC mortality in the NIH-AARP study is exclusion of subjects with a history of cancer. This may have exaggerated the protective effect of colonoscopy on CRC mortality. It is unlikely, however, that this limitation affected the estimates of confounding error, the main parameter of interest in this study.

In summary, we conclude that uncontrolled or unmeasured lifestyle-related confounders, either alone or in combination, are unlikely to produce a large spurious association between colonoscopy (or other methods of screening) and CRC mortality. Although limited to a single, albeit large, national cohort, our conclusions about HLS and NSAID are reasonably generalizable because we used the upper and lower 95% limits rather than point estimates in all calculations, and because measures of association in our study are similar to those reported elsewhere.[14, 15, 23, 27, 28, 32–34] It appears that other sources of systematic error such as non-random selection of participants or misclassification of screening status are of greater concern than uncontrolled confounding in observational studies of colonoscopy.

# References

1. Smith RA, Cokkinides V, Brawley OW. Cancer screening in the United States, 2008: a review of current American Cancer Society guidelines and cancer screening issues. CA Cancer J Clin. 2008; 58(3):161–79. [PubMed: 18443206]

2. Whitlock EP, Lin JS, Liles E, Beil TL, Fu R. Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. Ann Intern Med. 2008; 149(9):638–58. [PubMed: 18838718]

3. Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. N Engl J Med. 1993; 328(19):1365–71. [PubMed: 8474513]

4. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet. 1996; 348(9040):1472–7. [PubMed: 8942775]

5. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet. 1996; 348(9040):1467–71. [PubMed: 8942774]

6. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet. 2010; 375(9726):1624–33. [PubMed: 20430429]

7. Schoen RE, Pinsky PF, Weissfeld JL, Yokochi LA, Church T, Laiyemo AO, et al. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. N Engl J Med. 2012; 366(25):2345–57. [PubMed: 22612596]

8. Segnan N, Armaroli P, Bonelli L, Risio M, Sciallero S, Zappa M, et al. Once-only sigmoidoscopy in colorectal cancer screening: follow-up findings of the Italian Randomized Controlled Trial--SCORE. J Natl Cancer Inst. 2011; 103(17):1310–22. [PubMed: 21852264]

9. Dominitz, JA.; Robertson, DJ. Clinicaltrials.gov. 2012. Colonoscopy Versus Fecal Immunochemical Test in Reducing Mortality From Colorectal Cancer (CONFIRM).

10. Kaminski MF, Bretthauer M, Zauber AG, Kuipers EJ, Adami HO, van Ballegooijen M, et al. The NordICC Study: Rationale and design of a randomized trial on colonoscopy screening for colorectal cancer. Endoscopy. 2012; 44(7):695–702. [PubMed: 22723185]

11. Quintero E, Castells A, Bujanda L, Cubiella J, Salas D, Lanas A, et al. Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. N Engl J Med. 2012; 366(8):697–706. [PubMed: 22356323]

12. Rothman, KJ.; Greenland, S. Modern epidemiology. Philadelphia, PA: Lippincott Williams and Wilkins; 1998.

13. Brenner H, Chang-Claude J, Seiler CM, Rickert A, Hoffmeister M. Protection from colorectal cancer after colonoscopy: a population-based, case-control study. Ann Intern Med. 2011; 154(1): 22–30. [PubMed: 21200035]

14. Doubeni CA, Weinmann S, Adams K, Kamineni A, Buist DS, Ash AS, et al. Screening colonoscopy and risk for incident late-stage colorectal cancer diagnosis in average-risk adults: a nested case-control study. Ann Intern Med. 2013; 158(5 Pt 1):312–20. [PubMed: 23460054]

15. American Institute for Cancer Research., World Cancer Research Fund. Food, nutrition, physical activity and the prevention of cancer: a global perspective : a project of World Cancer Research Fund International. Washington, D.C: American Institute for Cancer Research; 2007.

16. Morrison DS, Batty GD, Kivimaki M, Davey Smith G, Marmot M, Shipley M. Risk factors for colonic and rectal cancer mortality: evidence from 40 years' follow-up in the Whitehall I study. J Epidemiol Community Health. 2011; 65(11):1053–8. [PubMed: 21389009]

17. Thun MJ, Calle EE, Namboodiri MM, Flanders WD, Coates RJ, Byers T, et al. Risk factors for fatal colon cancer in a large prospective study. J Natl Cancer Inst. 1992; 84(19):1491–500. [PubMed: 1433333]

18. Thun MJ, Namboodiri MM, Heath CW. Aspirin use and reduced risk of fatal colon cancer. N Engl J Med. 1991; 325(23):1593–6. [PubMed: 1669840]

19. Kelsall HL, Baglietto L, Muller D, Haydon AM, English DR, Giles GG. The effect of socioeconomic status on survival from colorectal cancer in the Melbourne Collaborative Cohort Study. Soc Sci Med. 2009; 68(2):290–7. [PubMed: 19022550]

20. Klima DA, Brintzenhoff RA, Agee N, Walters A, Heniford BT, Mostafa G. A review of factors that affect mortality following colectomy. J Surg Res. 2012; 174(2):192–9. [PubMed: 22099583]

21. McCaughan E, Prue G, Parahoo K. A systematic review of quantitative studies reporting selected patient experienced outcomes, with a specific focus on gender differences in people with colorectal cancer. Eur J Oncol Nurs. 2009; 13(5):376–85. [PubMed: 19520604]

22. Slattery ML, Edwards SL, Ma KN, Friedman GD. Colon cancer screening, lifestyle, and risk of colon cancer. Cancer Causes Control. 2000; 11(6):555–63. [PubMed: 10880038]

23. Coups EJ, Manne SL, Meropol NJ, Weinberg DS. Multiple behavioral risk factors for colorectal cancer and colorectal cancer screening status. Cancer Epidemiol Biomarkers Prev. 2007; 16(3): 510–6. [PubMed: 17372246]

24. Seeff LC, Nadel MR, Klabunde CN, Thompson T, Shapiro JA, Vernon SW, et al. Patterns and predictors of colorectal cancer test use in the adult U.S. population. Cancer. 2004; 100(10):2093–103. [PubMed: 15139050]

25. Baxter NN, Goldwasser MA, Paszat LF, Saskin R, Urbach DR, Rabeneck L. Association of colonoscopy and death from colorectal cancer. Ann Intern Med. 2009; 150(1):1–8. [PubMed: 19075198]

26. Newcomb PA, Norfleet RG, Storer BE, Surawicz TS, Marcus PM. Screening sigmoidoscopy and colorectal cancer mortality. J Natl Cancer Inst. 1992; 84(20):1572–5. [PubMed: 1404450]

27. Rothwell PM, Fowkes FG, Belch JF, Ogawa H, Warlow CP, Meade TW. Effect of daily aspirin on long-term risk of death due to cancer: analysis of individual patient data from randomised trials. Lancet. 2011; 377(9759):31–41. [PubMed: 21144578]

28. Rothwell PM, Wilson M, Elwin CE, Norrving B, Algra A, Warlow CP, et al. Long-term effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials. Lancet. 2010; 376(9754):1741–50. [PubMed: 20970847]

29. Guessous I, Dash C, Lapin P, Doroshenk M, Smith RA, Klabunde CN, et al. Colorectal cancer screening barriers and facilitators in older persons. Prev Med. 2010; 50(1–2):3–10. [PubMed: 20006644]

30. Power E, Miles A, von Wagner C, Robb K, Wardle J. Uptake of colorectal cancer screening: system, provider and individual factors and strategies to improve participation. Future Oncol. 2009; 5(9):1371–88. [PubMed: 19903066]

31. Brenner H, Chang-Claude J, Seiler CM, Hoffmeister M. Interval cancers after negative colonoscopy: population-based case-control study. Gut. 2011

32. Breslow NE, Day NE. Statistical methods in cancer research. Volume I - The analysis of case-control studies. IARC Sci Publ. 1980; (32):5–338. [PubMed: 7216345]

33. Schatzkin A, Subar AF, Thompson FE, Harlan LC, Tangrea J, Hollenbeck AR, et al. Design and serendipity in establishing a large cohort with wide dietary intake distributions : the National Institutes of Health-American Association of Retired Persons Diet and Health Study. Am J Epidemiol. 2001; 154(12):1119–25. [PubMed: 11744517]

34. National Cancer Institute. National Cancer Institute Surveillance Research Program, Statistical Methodology and Applications. 2012. DevCan: Probability of Developing or Dying of Cancer Software, Version 6.6.1.

35. Kleinbaum, D.; Klein, M. Survival Analysis: A Self-Learning Text. 2. New York: Springer Science+ Business Media; 2005.

36. Dean, A.; Sullivan, K.; Soe, M. OpenEpi: Open Source Epidemiologic Statistics for Public Health, Version 2.3.1. 2011. Available from: www.OpenEpi.com

37. Centers for Disease Control and Prevention. QuickStats: Outpatient Colonoscopy Procedure Rates,* by Age Group --- National Survey of Ambulatory Surgery, United States, 1996 and 2006. MMWR Morb Mortal Wkly Rep. 2010; 59(33):1075.

38. Samad AK, Taylor RS, Marshall T, Chapman MA. A meta-analysis of the association of physical activity with reduced risk of colorectal cancer. Colorectal Dis. 2005; 7(3):204–13. [PubMed: 15859955]

39. Huxley RR, Ansary-Moghaddam A, Clifton P, Czernichow S, Parr CL, Woodward M. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. Int J Cancer. 2009; 125(1):171–80. [PubMed: 19350627]

40. Jansen MC, Bueno-de-Mesquita HB, Buzina R, Fidanza F, Menotti A, Blackburn H, et al. Dietary fiber and plant foods in relation to colorectal cancer mortality: the Seven Countries Study. Int J Cancer. 1999; 81(2):174–9. [PubMed: 10188715]

**Table 1**

Vital status, demographic factors and medications among NIH-AARP cohort members with and without CRC screening history

| Variables of interest | Colorectal cancer screening in past 3 years | | |
| --- | --- | --- | --- |
| | No screening (N=45,751) | Colonoscopy (N=22,780) | Any screening (N=109,689) |
| | N (column %) | N (column %) | N (column %) |
| **CRC Death** | | | |
| No | 45,517 (99.5) | 22,726 (99.8) | 109,321 (99.7) |
| Yes | 234 (0.5) | 54 (0.2) | 368 (0.3) |
| **Age** | | | |
| <55 | 6,917 (15.1) | 1,933 (8.5) | 11,715 (10.7) |
| 55–59 | 10,879 (23.8) | 4,281 (18.8) | 22,775 (20.8) |
| 60–64 | 12,886 (28.2) | 6,533 (28.7) | 31,968 (29.1) |
| 65–69 | 13,603 (29.7) | 8,976 (39.4) | 38,946 (35.5) |
| 70+ | 1,466 (3.2) | 1,057 (4.6) | 4,285 (3.9) |
| **Education** [*] | | | |
| <8 years | 2,159 (4.7) | 820 (3.6) | 3,631 (3.3) |
| 8–11 years | 9,286 (20.3) | 3,477 (15.3) | 16,184 (14.8) |
| High school | 5,175 (11.3) | 2,090 (9.2) | 10,331 (9.4) |
| Some college | 11,056 (24.2) | 5,150 (22.6) | 24,514 (22.3) |
| College and post graduate | 17,081 (37.3) | 10,788 (47.4) | 53,027 (48.3) |
| **Race** [*] | | | |
| White | 43,174 (94.4) | 21,545 (94.6) | 103,915 (94.7) |
| Black | 922 (2.0) | 503 (2.2) | 2,247 (2.0) |
| Hispanic | 700 (1.5) | 302 (1.3) | 1,451 (1.3) |
| Asian | 459 (1.0) | 195 (0.9) | 1,037 (0.9) |
| Pacific Islander | 42 (0.1) | 12 (0.1) | 69 (0.1) |
| American Indian/Alaskan | 95 (0.2) | 43 (0.2) | 192 (0.2) |
| **Diabetes** | | | |
| No | 42,968 (93.9) | 21,088 (92.6) | 101,700 (92.7) |
| Yes | 2,783 (6.1) | 1,692 (7.4) | 7,989 (7.3) |
| **Sex and HRT use** | | | |
| Male | 24,430 (53.4) | 15,834 (69.5) | 72,507 (66.1) |
| Female never used HRT | 11,009 (24.1) | 2,496 (11.0) | 13,507 (12.3) |
| Female used HRT | 10,312 (22.5) | 4,450 (19.5) | 23,675 (21.6) |
| **Aspirin or NSAID use** | | | |
| Never | 6,470 (14.1) | 3,072 (13.5) | 12,910 (11.8) |
| Monthly | 16,370 (35.8) | 6,747 (29.6) | 34,276 (31.2) |
| Weekly | 10,559 (23.1) | 4,939 (21.7) | 25,252 (23.0) |
| Daily | 12,352 (27.0) | 8,022 (35.2) | 37,251 (34.0) |

[*] Numbers do not add up to total due to missing values

**Table 2**

Lifestyle Score (LS) components expected to be associated with increased CRC risk among NIH-AARP cohort members with and without colonoscopy screening history

| Score Variables (sex-specific cutoffs)[*] | Colonoscopy in past 3 years | |
|---|---|---|
| | No (N=45,751) | Yes (N=22,780) |
| | N (row %) | N (row %) |
| **BMI (kg/m$^2$)** | | |
| Upper 25% (>29.19 M, >29.29 F) | 10,150 (65.9) | 5,258 (34.1) |
| 50–75% (26.58–29.19 M, 25.64–29.29 F) | 10,751 (65.4) | 5,695 (34.6) |
| 25–50%(24.42–26.57 M, 22.74–25.63 F) | 12,313 (67.1) | 6,036 (32.9) |
| Lower 25% (  24.41 M,  22.73 F) | 12,537 (68.4) | 5,791 (31.6) |
| **Waist-to-hip ratio** | | |
| Upper 25% (>0.99 M, >0.86 F) | 11,877 (66.1) | 6,100 (33.9) |
| 50–75% (0.96–0.99 M, 0.82–0.86 F) | 11,457 (66.4) | 5,792 (33.6) |
| 25–50% (0.92–0.95 M, 0.77–0.81 F) | 11,379 (67.1) | 5,579 (32.9) |
| Lower 25% (  0.91 M,  0.76 F) | 11,038 (67.5) | 5,309 (32.9) |
| **Red meat consumption (g/day)** | | |
| Upper 25% (>109.3 M, >67.3 F) | 12,590 (70.6) | 5,229 (29.4) |
| 50–75% (71.5–109.3 M, 42.0–67.3 F) | 11,907 (67.9) | 5,626 (32.0) |
| 25–50% (43.1–71.4 M, 23.7–41.9 F) | 11,132 (64.6) | 6,088 (35.4) |
| Lower 25% (  43.0 M,  23.6 F) | 10,122 (63.4) | 5,837 (36.6) |
| **Processed meat consumption (g/day)** | | |
| Upper 25% (>34.8 M, >18.3 F) | 11,921 (68.2) | 5,545 (31.8) |
| 50–75% (19.2–34.8 M, 9.4–18.3 F) | 11,678 (67.0) | 5,759 (33.0) |
| 25–50% (9.8–19.1 M, 4.4–9.3 F) | 11,488 (66.6) | 5,754 (33.4) |
| Lower 25% (  9.7 M,  4.3 F) | 10,664 (65.1) | 5,722 (34.9) |
| **Alcohol consumption (g/day)** | | |
| Upper 25% (>18.3 M, >5.7 F) | 11,772 (66.2) | 6,004 (33.8) |
| 50–75% (4.8–18.3 M, 1.2–5.7 F) | 11,041 (65.4) | 5,833 (34.6) |
| 25–50% (0.7–4.7 M, 0.1–1.1 F) | 11,098 (66.5) | 5,585 (33.5) |
| Lower 25% (  0.6 M, 0.0 F) | 11,840 (68.8) | 5,358 (31.2) |

[*]M = males, F = females

**Table 3**

Lifestyle Score (LS) components expected to be associated with decreased CRC risk among NIH-AARP cohort members with colonoscopy screening history

| Score Variables (sex-specific cutoffs)[*] | Colonoscopy in past 3 years | |
| --- | --- | --- |
| | No (N=45,751) | Yes (N=22,780) |
| | N (row %) | N (row %) |
| **Physical activity**[**] | | |
| Never or Rarely | 10,895 (69.2) | 4,840 (30.8) |
| 1–3 Hours per Week | 11,528 (67.2) | 5,637 (32.8) |
| 4–7 Hours per Week | 11,643 (64.7) | 6,350 (35.3) |
| >7 Hours per Week | 11,685 (66.2) | 5,953 (33.8) |
| **Milk consumption (g/day)** | | |
| Lower 25% ( 69 M, 49 F) | 11,867 (68.8) | 5,375 (31.2) |
| 25–50% (70–173 M, 49–150 F) | 11,282 (65.3) | 6,000 (34.7) |
| 50–75% (174–367 M, 151–332 F) | 11,294 (65.7) | 5,896 (34.3) |
| Upper 25% (>367 M, >332 F) | 11,308 (67.2) | 5,509 (32.8) |
| **Calcium intake (mg/day)** | | |
| Lower 25% ( 585 M, 514 F) | 12,114 (68.5) | 5,557 (31.5) |
| 25–50% (586–769 M, 515–690 F) | 11,548 (66.0) | 5,935 (34.0) |
| 50–75% (770–1036 M, 691–943 F) | 11,321 (65.9) | 5,861 (31.1) |
| Upper 25% (>1036 M, >943 F) | 10,768 (66.5) | 5,427 (33.5) |
| **Fiber intake (g/day)** | | |
| Lower 25% ( 15.8 M, 13.4 F) | 12,455 (71.0) | 5,098 (29.0) |
| 25–50% (20.3–15.9 M, 17.2–13.5 F) | 11,299 (66.5) | 5,681 (33.5) |
| 50–75% (20.4–26.0 M, 21.8–17.3 F) | 10,969 (64.9) | 5,923 (35.1) |
| Upper 25% (>26.0 M, >21.8 F) | 11,028 (64.5) | 6,078 (35.5) |

[*] M = males, F = females

[**] The same cutoffs used for males and females

**Table 4**

RR$_{confounder}$ estimates reflecting the relation of confounding factors of interest (Lifestyle Score and aspirin/ NSAID use) to CRC mortality

| Confounders of interest | Deaths | Person-Years | RR$_{confounding}$ (95% CI) |
|---|---|---|---|
| **Unweighted HLS** | | | |
| <20 | 188 | 414,127 | 1.0 (ref) |
| 20–22 | 166 | 411,725 | 0.88 (0.71–1.09) |
| 23–25 | 126 | 417,459 | 0.66 (0.52–0.82) |
| 26+ | 122 | 457,881 | 0.57 (0.46–0.72) |
| **Weighted HLS** | | | |
| <6 | 126 | 255,722 | 1.0 (ref) |
| 6–7 | 189 | 480,307 | 0.81 (0.65–1.01) |
| 8–9 | 190 | 566,170 | 0.69 (0.55–0.87) |
| 10+ | 97 | 398,994 | 0.50 (0.38–0.65) |
| **Education** | | | |
| High School or less (<12 yrs.) | 199 | 511,933 | 1.0 (ref) |
| At least some college (>12 yrs.) | 386 | 1,157,041 | 0.91 (0.77–1.08) |
| **At least weekly Aspirin/NSAID use** | | | |
| No | 301 | 770,847 | 1.0 (ref) |
| Yes | 301 | 930,346 | 0.80 (0.68–0.94) [**] |

[*] Adjusted for age, sex/HRT, education, race, diabetes, family history of CRC, history of screening and aspirin or NSAID use

[**] Adjusted for age, sex/HRT, education, race, diabetes, family history of CRC, history of screening and HLS

**Table 5**

Range of $P_0$, $P_1$, and confounding error estimates in the association between CRC screening and mortality

| Factor | No CRC screening ("NO") N=45,751 | | Colonoscopy ("COLO") N=22,780 | | Any CRC screening ("ANY") N=109,689 | | Fecal occult blood test (FOBT) N=94,418 | | Estimated range of confounding error: $RR_{biased}/RR_{true}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | $P_0$ (95% CI) | N | $P_1$ (95% CI) | N | $P_1$ (95% CI) | N | $P_1$ (95% CI) | NO vs. COLO | NO vs. ANY | NO vs. FOBT |
| **Unweighted HLS** | | | | | | | | | | | |
| <20 | 12,408 | .271 (.267–.275) | 5,562 | .244 (.239–.250) | 25,785 | .235 (.233–.238) | 5,562 | .231 (.229–.234) | | | |
| 20–22 | 11,248 | .246 (.242–.250) | 5,501 | .241 (.236–.247) | 26,483 | .241 (.242–.247) | 5,501 | .240 (.237–.243) | | | |
| 23–25 | 10,906 | .238 (.235–.242) | 5,680 | .249 (.244–.255) | 27,224 | .248 (.246–.251) | 5,680 | .250 (.247–.253) | | | |
| 26+ | 11,189 | .245 (.241–.249) | 6,037 | .265 (.259–.271) | 30,197 | .275 (.273–.278) | 6,037 | .279 (.276–.282) | 0.960–0.995 | 0.958–0.991 | 0.955–0.989 |
| **Weighted HLS** | | | | | | | | | | | |
| <6 | 7,844 | .171 (.168–.175) | 3,497 | .154 (.149–.158) | 15,902 | .145 (.143–.147) | 13,428 | .142 (.140–.145) | | | |
| 6–7 | 13,440 | .294 (.290–.298) | 6,565 | .288 (.282–.294) | 30,813 | .281 (.278–.284) | 26,254 | .278 (.275–.281) | | | |
| 8–9 | 14,868 | .325 (.321–.329) | 7,544 | .331 (.325–.337) | 36,680 | .334 (.332–.337) | 31,750 | .336 (.333–.339) | | | |
| 10+ | 9,599 | .210 (.205–.213) | 5,174 | .227 (.222–.233) | 26,294 | .240 (.237–.242) | 22,986 | .243 (.241–.246) | 0.959–0.997 | 0.955–0.990 | 0.952–0.988 |
| **Education** | | | | | | | | | | | |
| 12 yrs. | 16,620 | .371 (.367–.376) | 6,387 | .286 (.280–.292) | 30,146 | .280 (.277–.283) | 25,564 | .276 (.273–.279) | | | |
| 12 yrs | 28,137 | .629 (.624–.633) | 15,938 | .714 (.708–.720) | 77,541 | .720 (.717–.723) | 67,176 | .724 (.721–.727) | 0.974–1.006 | 0.973–1.006 | 0.972–1.007 |
| **At least weekly aspirin or NSAID use** | | | | | | | | | | | |
| No | 22,840 | .499 (.495–.504) | 9,819 | .431 (.425–.438) | 47,186 | .430 (.427–.433) | 40,160 | .425 (.422–.429) | | | |
| Yes | 22,911 | .501 (.496–.505) | 12,961 | .569 (.563–.575) | 62,503 | .570 (.567–.573) | 54,258 | .575 (.572–.578) | 0.970–0.996 | 0.971–0.996 | 0.969–0.996 |