HHS Public Access

Author manuscript

Int J Cancer. Author manuscript; available in PMC 2021 June 15.

Published in final edited form as:

Int J Cancer. 2021 June 15; 148(12): 2947–2953. doi:10.1002/ijc.33489.

A comparison of methods in estimating population attributable risk for colorectal cancer in the United States

Hanseul Kim[1], **Kai Wang**[1], **Mingyang Song**[1,2,3,4], **Edward L. Giovannucci**[1,2,5]

[1]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts

[2]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts

[3]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts

[4]Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts

[5]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts

Abstract

Population attributable risk (PAR) is becoming more widely used for quantifying preventability of cancer. However, its estimations have had a wide range, leading to questions about the true preventability. Our study aimed to compare the two PAR estimation methods (ie, literature-based method and low-risk method) for colorectal cancer (CRC) in the US population based on the same set of modifiable risk factors: physical activity, body mass index, alcoholic drinks, red meat, processed meat, dietary fiber, dietary calcium and cigarette smoking. For the literature-based method, 65% and 53%, and for the low-risk method, 62% and 49% of CRC cases for males and females, respectively, were attributable to the eight dietary and lifestyle risk factors. Additional sensitivity analyses were conducted with respect to the different choices of risk factors, relative risks (RRs) and exposure prevalence estimates used in the literature-based method. The PARs including only the "convincing" factors and excluding "probable" factors defined by the WCRF/AICR were 50% for males and 34% for females. Using RRs derived from different studies changed the PARs considerably (57%−74% for males and 37%−60% for females). Our study assessed the robustness of PAR calculations through a direct comparison between the two methods

Correspondence: Hanseul Kim, Department of Epidemiology, Harvard T.H. Chan School of Public Health, 665 Huntington Avenue, Boston, MA 02115. hanseul.kim@mail.harvard.edu.
Hanseul Kim and Kai Wang contributed equally as co-first authors.
Mingyang Song and Edward L. Giovannucci contributed equally as co-last authors.

CONFLICT OF INTEREST

The authors have declared no conflicts of interest.

DATA AVAILABILITY STATEMENT

The details about the data that support the findings of this study are available at: http://www.nurseshealthstudy.org/ and https://sites.sph.harvard.edu/hpfs/for-collaborators/. These data are not publicly available due to ethical and privacy restrictions. The data that supports the findings of this study are available from the corresponding author upon reasonable request.

SUPPORTING INFORMATION Additional supporting information may be found online in the Supporting Information section at the end of this article.

using different assumptions and data and generally found high concordance. From the additional analyses, we found that the choice of risk factors and RRs could substantially influence the PAR estimates. Given the findings, future studies reporting PAR should consider presenting a range of PAR estimates based on choices of risk factors and RRs.

Keywords

cancer; colorectal cancer; population attributable fraction; population attributable risk; prevention

1 | INTRODUCTION

The proportional reduction in disease incidence that would be achieved in a population if exposures are optimized to minimize risk is termed the “population attributable risk” (PAR). PAR has been a widely used measure in quantifying the preventability of diseases, especially cancer.

Although PAR has become widely used, its estimates have varied widely, leading to some questions about the true preventability. In 1981, Doll and Peto attributed about 35% of cancers in the United States to “diet and nutrition,”[1] which are largely modifiable. Besides, they suggested that 90% of colorectal cancers (CRCs) may be related to diet. However, Blot and Tarone indicated that the 90% estimate appears too high for CRC.[2] In support of this, they referred to the study among the European Prospective Investigation into Cancer and Nutrition cohort, which reported that only 16% of CRCs were attributable to a combination of five healthy lifestyle factors.[3] PAR estimates from studies across the world varied widely between 16% and 90%.[4–6] Such variation could be due to diverse counterfactual scenarios assumed for different studies, differences between countries, and changes in the definition of “diet and nutrition” over time. For instance, recent PAR estimates have included physical activity, body mass index (BMI) and cigarette smoking as nutritional and lifestyle factors related to CRC risk. However, the exact reasons underlying the differences in PAR estimates remain unclear.

Two general approaches have been taken to estimate PAR for cancer in the United States. The approach most often used, henceforth termed the “literature-based method,” identifies relative risks (RRs) and prevalence from population sources. The second approach identifies a low-risk group in a cohort study (henceforth termed the “low-risk method”) based on adherence to low risk for identified risk factors and compares the rate of the low-risk group with that in the population.

In the current study, we estimate PAR for CRC within the US population using the literature-based method and compare it with the low-risk method using the same set of risk factors. Through additional sensitivity analyses, we sought to explore the impact of the following components in PAR calculations for the literature-based method: (a) choice of risk factors, (b) different sources of RR estimates, and (c) different sources of nationally representative exposure prevalence estimates.

2 | MATERIALS AND METHODS

2.1 | Literature-based method

For the main analysis, we used the literature-based method to estimate PAR for CRC within the US population. Risk factors, RR estimates and exposure prevalence estimates are the major components of the literature-based method.

2.1.1 | Choice of risk factors—Risk factors for CRC were included in the calculation of PAR if they were classified by the World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR)[7] as having "probable" or "convincing" evidence of causal association (Table 1). In addition to the seven factors (physical activity, body fatness measured by BMI, alcoholic drinks, red meat, processed meat, dietary fiber and dietary calcium), we included cigarette smoking based on the Surgeon General Report.[8] We were focused on modifiable lifestyle factors related to CRC risk instead of nonmodifiable factors that individuals have less control over.

2.1.2 | Choice of RR estimates—RRs were identified through searches in PubMed. Meta-analyses with summary RR estimates and corresponding 95% confidence intervals were the most preferred source of RRs, followed by large pooled analyses. Within meta- and pooled analyses where multiple estimates were reported or where more than one meta- or pooled analysis was available, RRs from the most recent studies were selected and they had to provide the most relevant information to our study. For example, studies that provided RRs on our predefined exposure category levels were preferred. In addition, studies with sex-specific RRs were used for risk factors that showed a sex difference in their relationship with CRC. Details on the selected RRs for each of the factors are included in the Supplementary Methods.

For risk factors (BMI, alcoholic drinks, red meat and processed meat), the lowest category was used as the reference. For protective factors (physical activity, fiber and calcium), the highest category was used as the reference, and PAR was calculated using the reciprocal of the RR. When the reported RRs did not match our predefined category levels, we recalculated the RRs. For instance, for a study that reported an RR associated with A units of exposure (ie, RR_A), the RR of new units of exposure (B units) that matches our predefined exposure category (ie, RR_B) was calculated by this formula[9]: $RR_B = \exp\left(\frac{\log(RR_A)}{A} * B\right)$. The midpoints of categories were used to recalculate the RRs. Details on the specific assumptions made for each of the factors are included in the Supplementary Methods.

2.1.3 | Choice of exposure prevalence estimates—Exposure prevalence estimates had to be from a nationally representative population survey that provides information on each of the exposure category levels. Exposure distribution data for alcohol drinking and cigarette smoking were obtained from the National Health Interview Survey and for other factors from the National Health and Nutrition Examination Survey (NHANES). Detailed methods of deriving weighted prevalence for each exposure were described elsewhere.[9]

2.1.4 | Statistical analysis (literature-based method)—For each of the risk factors with n levels, we estimated PAR using the literature-based method with the following equation:

$$\frac{\sum_{i=1}^{n} P_i^*(\mathrm{RR}_i - 1)}{\sum_{i=1}^{n} P_i^*(\mathrm{RR}_i - 1) + 1},$$

where P_i is the exposure prevalence at the exposure category i and RR_i is the corresponding RR of CRC at exposure category i. Categorization details of exposures are presented in Table 1.

Preventability of CRC attributable to combined lifestyle risk factors was estimated using the following equation:

$$\mathrm{PAR} = 1 - \prod_{j=1}^{J} (1 - \mathrm{PAR}_j),$$

where j is an individual risk factor and J represents the number of risk factors.

All statistical analyses used to calculate PARs were performed using the R statistical software (version 3.6.2; R Foundation for Statistical Computing, Vienna, Austria).

2.2 | Low-risk method

We used the Nurses' Health Study (NHS) and Health Professionals Follow-up Study (HPFS) for the low-risk method to compare with the literature-based method. Details on these cohorts are included in the Supplementary Methods.

For the low-risk method, we addressed the same set of lifestyle factors that we included for the literature-based method (ie, physical activity, BMI, alcoholic drinks, red meat, processed meat, dietary fiber, dietary calcium and cigarette smoking). The low risk was defined as moderate-to-vigorous intensity activity for ≥30 min/d, BMI of ≥18.5 and <25.0 kg/m^2, none-to-moderate alcohol intake (≤1 drink [14 g alcohol]/d for women and ≤2 drinks/d for men), red meat intake <0.5 serving/d, processed meat intake <0.2 serving/d, total dietary fiber intake ≥20 g/d for women and ≥30 g/d for men, total calcium intake >1000 mg/d and never smoking or past smoking with pack-years <5. For each of the 8 lifestyle factors, we defined a binary criterion, by which the participants received a score of 1 if they met the criterion and 0 otherwise. An overall healthy lifestyle score (range, 0–8) was then calculated by summing the 8 scores, with a higher score indicating a healthier lifestyle and lower risk of CRC.

2.2.1 | Statistical analysis (low-risk method)—Age- and sex-specific rates for CRC were calculated for the age groups (<45, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, 80–84 and ≥85 years) and then standardized to the age distribution of the 2000 US standard population. We estimated PAR by comparing the age-standardized CRC incidence of our low-risk group with the age-standardized national rates (obtained from the Surveillance,

Epidemiology, and End Results program). We defined the low-risk group as those meeting healthy lifestyle score of ≥7 as an insufficient number of participants met all eight criteria. PAR here could be interpreted as the proportion of CRC cases that would not occur if all individuals followed the lifestyle of the low-risk population.

2.3 | Additional analyses

To test the robustness of the literature-based method that we used to derive PAR estimates, we not only compared the results from the literature-based method to those of the low-risk method from the NHS and HPFS cohorts, but also conducted additional sensitivity analyses. In the additional analyses for the literature-based method, we sought to explore the impact of the following components in the literature-based method: (a) choice of risk factors, (b) different sources of RR estimates and (c) different sources of exposure prevalence estimates. To be specific, we achieved these through (a) excluding WCRF/AICR-defined "probable" factors (red meat, dietary fiber and dietary calcium) from the calculation of PAR, (b) examining how RRs from different sources (ie, different studies that report RRs) change the PAR estimate, and (c) assessing how different sources of nationally representative prevalence estimates change the PAR. Details on these additional analyses are included in the Supplementary Methods.

3 | RESULTS

3.1 | Literature-based method

Summary PAR results are presented in Table 2. Prevalence estimates and RRs used to obtain these PAR estimates are presented in Supplementary Tables 1 and 2. The proportions of CRC cases attributable to lifestyle risk factors were 65% for males and 53% for females in the United States (Table 2). The proportions of CRC cases attributable to lifestyle risk factors excluding cigarette smoking were 60% for males and 47% for females. Dietary fiber contributed to the largest proportion of attributable CRC cases for both males and females, accounting for 18% for males and 17% for females. Body fatness was the second-largest preventable risk factor of CRC for males (17%), but relatively small for females (8%). Large sex differences in PAR were also observed for alcoholic drinks and processed meat, with higher PAR values for males than females.

3.2 | Low-risk method

Results from the low-risk method comparing the cohorts to the national rates are also presented in Table 2. Characteristics throughout follow-up according to healthy lifestyle score of ≥7 are presented in Supplementary Table 3. There were 76 CRC cases in the low-risk group for the NHS and 45 CRC cases for the HPFS. Comparing rates in the low-risk groups for the NHS and HPFS cohorts to the national rates, the proportions of CRC cases attributable to lifestyle factors were 62% for males and 49% for females when defining low-risk group as those adhering to ≥7 healthy factors (Table 2). Using 6 as the cutoff (ie, adhering to ≥6 healthy factors) did not substantially change our results (results not shown, available upon request).

3.3 | Additional analyses

After excluding "probable" factors classified by the WCRF/AICR, the PARs including the 5 factors (physical activity, body fatness/BMI, alcoholic drinks, processed meat and cigarette smoking) were 50% for males and 34% for females for the literature-based method (Table 3).

We examined how different sources of RRs changed the PAR estimates. RRs from studies that were used for this sensitivity analysis are presented in Supplementary Table 4. The proportions of CRC cases attributable to lifestyle risk factors ranged from 57% to 74% for males and 37% and 60% for females by varying the RR estimates (Table 4). Such wide variation for PAR could mainly be attributed to differences in RR estimates for physical activity and BMI. To be specific, the lower bounds of PAR derived from different RR sources for physical activity were 3% for both males and females. This low estimate was derived from a meta-analysis by Kyu et al[10] that reported a very wide range of physical activity ranging from <600 MET-m/wk to ≥8000 MET-m/wk (Supplementary Table 4). In addition, the minimum and maximum PARs for BMI varied widely (17% and 36% for males and 4% and 21% for females).

Using available information from studies that reported region-specific RR estimates, the PAR calculated with US-specific RRs and US prevalence data was 35% for the 6 factors: BMI (only the contribution of obesity was estimated), alcoholic drinks, red meat, processed meat, dietary fiber and dietary calcium (Supplementary Table 5). The corresponding PAR for the same factors calculated with worldwide RRs and US prevalence data yielded similar estimate of 37%.

We also assessed how different sources of exposure prevalence estimates change the PAR estimates. The proportions of CRC cases attributable to lifestyle risk factors ranged from 64% to 67% for males and 53% to 56% for females by varying prevalence estimates for body fatness and smoking (Supplementary Table 6). Sources of different prevalence estimates are presented in Supplementary Table 7.

4 | DISCUSSION

We found that for the literature-based method, 65% and 53%, and for the low-risk method, 62% and 49% of CRC cases for males and females, respectively, were attributable to dietary and lifestyle risk factors in adulthood in the United States. Large sex differences in the PAR for the literature-based method were observed for body fatness, alcoholic drinks, and processed meat, mostly driven by sex differences in RRs for body fatness and alcohol consumption, sex differences in prevalence estimates for processed meat consumption and partly by higher alcohol drinking in men.

We reported comparable PAR estimates with two different methods and from additional analyses, using contemporary and nationally representative data on exposure and RRs. The remarkable similarity of the results from the literature-based method and the low-risk method increases the confidence of the overall robustness of the PAR estimation to the specific methods and data sources used. Of note, the two methods are based on entirely

different assumptions; the literature-based method relies on estimated RRs and prevalence of risk factors from the literature, and in the low-risk method, age-standardized incidence rates of a selected low-risk population are compared to that of the general population. Our results were also consistent with the results from the recent paper by Islami et al[9] on the PAR of modifiable risk factors in the United States. Using the literature-based method, they reported that 58% of CRC cases for men and 51% for women were attributable to lifestyle and dietary risk factors among US adults aged 30 years and older, which were similar to our estimates.

The robustness of PAR estimation was assessed through additional analyses. Excluding "probable" factors changed the PAR estimates from 65% to 50% for males and 53% to 34% for females for the literature-based method. As such, selection of factors seems to greatly contribute to the variation in PAR estimates. In addition, varying sources of RRs changed the PAR values considerably (ranges of 57%−74% for males and 37%−60% for females). However, the minimum and maximum ranges were derived from two studies for each factor except for BMI where the result was derived from three studies. Therefore, the range could be regarded as a conservative range of variation based on the best data available. Changing prevalence estimates did not change the PAR values as much, although we could identify different data sources of prevalence estimates only for body fatness and cigarette smoking.

We could infer that choosing the appropriate RR is critical in PAR estimation process. In our study, we used the RRs from review, pooled analysis and mostly meta-analyses. As physical activity and BMI are the most studied factors, there was a higher likelihood for greater variability in RRs from reviews or meta-analyses for physical activity and BMI. For example, Calle et al[11] presented RRs of 1.5 and 2.0 for overweight and obese males compared to normal males and 1.2 and 1.5 for overweight and obese females compared to normal females, which were very different from the RRs that we used for the main analysis (1.17 and 1.38 for overweight and obese males; 1.07 and 1.17 for overweight and obese females[12]). RRs could vary even more with different study designs (eg, single cohort study and case-control study).

There were several inherent limitations in studies used to estimate the PARs. First, the PAR formula that we used for the literature-based method can be biased if there is confounding or effect modification.[13] As this is one of the most commonly used PAR formulas in the literature, our aim was to compare this to another method and evaluate which components in the formula contribute to variations in PAR calculations. We found that even with an entirely different set of assumptions, the low-risk method yielded similar estimates compared to the literature-based method. Second, we did not present 95% confidence intervals for the PAR estimates. As the goal of our study was to illustrate variability that stem from systematic differences, we presented variability through a range of PAR estimates from different scenarios rather than presenting 95% confidence intervals that capture sampling variability. Third, when calculating PARs, we assumed that the risk factors are independent. There was no comprehensive information available on the interactions between each risk factors. Our literature-based PAR estimates could therefore be overestimated or underestimated. Nevertheless, the results from the low-risk method, which inherently accounts for interactions, were similar to those from the literature-based method. Fourth, the assumptions

made in the formation of RRs to match the exposure prevalence categories may have impacted the results. Fifth, some risk factors may be more critical early in life, such as adolescent BMI and CRC,[14] but were unaccounted for in the studies of older adults. Sixth, for most risk factors, we used worldwide RRs that were not specific to the United States. If significant heterogeneity existed between countries, worldwide RRs obtained from meta-analyses would provide invalid PAR for the United States. We performed a sensitivity analysis that compared the PAR estimates using US-specific RRs vs worldwide RRs. In this analysis, PARs using US-specific RRs and worldwide RRs were similar (35% and 37%). Moreover, most studies that were included in the meta-analyses were from the United States. Also, for the meta-analyses that reported stratified RRs by region, no significant heterogeneity was observed by geographic area (P for heterogeneity >.05).[15–18] Seventh, for some risk factors (ie, physical activity, red meat and processed meat), sex-specific RRs were not available, which could have impacted the results. We used the best data available and there seemed to be no significant heterogeneity by sex observed for these factors.[7,19] Eighth, the choices of reference groups for exposure categories may have impacted the analysis. Although the reference groups were chosen based on public health recommendations, it may be unrealistic to expect the population to meet the criteria. Finally, our additional sensitivity analyses were based on the best data available. There could be a wider variation in PAR estimation based on different sources of RRs and prevalence estimates.

There were also some limitations for the low-risk method. The low-risk method analysis was based on identifying a low-risk group that meets the identified criteria for “low risk.” Presumably, a higher proportion of our cohorts consisted of health professionals were more health conscious and more likely to meet the low-risk criteria. However, the underlying assumption is that US nonhealth professionals of comparable age, race and sex experience a similar CRC risk. Although the assumptions for the literature-based method and low-risk method are entirely different, these two methods resulted in comparable estimates in our study. In addition, factors forming the risk score in the low-risk method have different effect sizes in the association with CRC. Although it is true that some risk factors may be more important than others, this is irrelevant in our goal of determining the maximal preventable cancers. Moreover, the healthy lifestyle score has been widely used before[20–23] and was created by dichotomizing dietary and lifestyle factors based on public health recommendations and prior knowledge on their associations with CRC.

In conclusion, our study addressed the robustness of PAR calculations by comparing one method (ie, literature-based method) to another (ie, low-risk method) and through multiple additional sensitivity analyses. Limitations of PAR calculations often stem from subjective choices of methods used and data sources for RRs and prevalence estimates. To examine the influence of these sources of variability, we used PAR for CRC in the US population as an example. The PAR estimates from two different methods yielded similar results that ~60% of CRC cases for males and ~50% of CRC cases for females are potentially preventable with healthy diet and lifestyle, which are modifiable. From the additional analyses, we found that for the literature-based method, choices of risk factors and RRs could substantially influence the PAR estimation. Given the findings, future PAR studies that use the literature-based method should consider presenting a range of estimates that could be obtained with different choices of risk factors and RRs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors would like to thank the participants and staff of the Nurses' Health Study and Health Professionals Follow-Up Study for their continued contributions, as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA and WY. The authors assume full responsibility for analyses and interpretation of these data.

Funding for grant 2018/1818 was obtained from World Cancer Research Fund UK (WCRF), as part of the World Cancer Research Fund International grant programme [to E.L. Giovannucci]. This work was also supported by the American Cancer Society Mentored Research Scholar Grant [MRSG-17-220-01 to M Song]; and by the U.S. National Institutes of Health grants [P01 CA87969 to M.J. Stampfer, UM1 CA186107 to M.J. Stampfer, P01 CA55075 to W.C. Willett, UM1 CA167552 to W.C. Willett, U01 CA167552 to L.A. Mucci and W.C. Willett, R00 CA215314 to M Song].

Funding information American Cancer Society, Grant/Award Number: MRSG-17-220-01; National Institutes of Health, Grant/Award Numbers: P01 CA55075, P01 CA87969, R00 CA215314, U01 CA167552, UM1 CA167552, UM1 CA186107; World Cancer Research Fund UK, Grant/Award Number: Grant 2018/1818

Abbreviations:

BMI	body mass index
CRC	colorectal cancer
HPFS	Health Professionals Follow-up Study
NHANES	National Health and Nutrition Examination Survey
NHS	Nurses' Health Study
PAR	population attributable risk
RR	relative risk
WCRF/AICR	World Cancer Research Fund/American Institute for Cancer Research

REFERENCES

1. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. J Natl Cancer Inst. 1981;66:1191–1308. [PubMed: 7017215]
2. Blot W, Tarone R. Doll and Peto's quantitative estimates of cancer risks: holding generally true for 35 years. J Natl Cancer Inst. 2015;107: djv044. [PubMed: 25739419]
3. Aleksandrova K, Pischon T, Jenab M, et al. Combined impact of healthy lifestyle factors on colorectal cancer: a large European cohort study. BMC Med. 2014;12:168. [PubMed: 25319089]
4. Gu MJ, Huang QC, Bao CZ, et al. Attributable causes of colorectal cancer in China. BMC Cancer. 2018;18:38. [PubMed: 29304763]
5. IARC (International Agency for Research on Cancer). IARC working group reports. Attributable Causes of Cancer in France in the Year 2000. Vol 3. Lyon: IARC Press; 2007.

6. Parkin DM, Boyd L, Walker LC. 16. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. Br J Cancer. 2011;105(Suppl 2):S77–S81. [PubMed: 22158327]
7. World Cancer Research Fund International/American Institute for Cancer Research, Continuous Update Project Report: Diet, Nutrition, Physical Activity and Colorectal Cancer, 2017.
8. National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. The Health Consequences of Smoking—50Years of Progress: A Report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2014.
9. Islami F, Goding Sauer A, Miller KD, et al. Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. CA Cancer J Clin. 2018;68:31–54. [PubMed: 29160902]
10. Kyu HH, Bachman VF, Alexander LT, et al. Physical activity and risk of breast cancer, colon cancer, diabetes, ischemic heart disease, and ischemic stroke events: systematic review and dose-response meta-analysis for the global burden of disease study 2013. BMJ. 2016;354:i3857. [PubMed: 27510511]
11. Calle EE, Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. Nat Rev Cancer. 2004;4:579–591. [PubMed: 15286738]
12. Xue K, Li FF, Chen YW, Zhou YH, He J. Body mass index and the risk of cancer in women compared with men: a meta-analysis of prospective cohort studies. Eur J Cancer Prevent. 2017;26:94–105.
13. Steenland K, Armstrong B. An overview of methods for calculating the burden of disease due to specific risk factors. Epidemiology. 2006; 17:512–519. [PubMed: 16804473]
14. Levi Z, Kark JD, Katz LH, et al. Adolescent body mass index and risk of colon and rectal cancer in a cohort of 1.79 million Israeli men and women: a population-based study. Cancer. 2017;123: 4022–4030. [PubMed: 28736986]
15. Keum N, Aune D, Greenwood DC, Ju W, Giovannucci EL. Calcium intake and colorectal cancer risk: dose-response meta-analysis of prospective observational studies. Int J Cancer. 2014;135:1940–1948. [PubMed: 24623471]
16. Bagnardi V, Rota M, Botteri E, et al. Alcohol consumption and site-specific cancer risk: a comprehensive dose-response meta-analysis. Br J Cancer. 2015;112:580–593. [PubMed: 25422909]
17. Zhao Z, Feng Q, Yin Z, et al. Red and processed meat consumption and colorectal cancer risk: a systematic review and meta-analysis. Oncotarget. 2017;8:83306–83314. [PubMed: 29137344]
18. Aune D, Chan DS, Lau R, et al. Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. BMJ. 2011;343:d6617. [PubMed: 22074852]
19. Moore SC, Lee IM, Weiderpass E, et al. Association of leisure-time physical activity with risk of 26 types of cancer in 1.44 million adults. JAMA Intern Med. 2016;176:816–825. [PubMed: 27183032]
20. Song M, Giovannucci E. Preventable incidence and mortality of carci-noma associated with lifestyle factors among white adults in the United States. JAMA Oncol. 2016;2:1154–1161. [PubMed: 27196525]
21. Li Y, Pan A, Wang DD, et al. Impact of healthy lifestyle factors on life expectancies in the US population. Circulation. 2018;138: 345–355. [PubMed: 29712712]
22. Kirkegaard H, Johnsen NF, Christensen J, Frederiksen K, Overvad K, Tjønneland A. Association of adherence to lifestyle recommendations and risk of colorectal cancer: a prospective Danish cohort study. BMJ. 2010;341:c5504. [PubMed: 20978063]
23. Carr PR, Weigl K, Jansen L, et al. Healthy lifestyle factors associated with lower risk of colorectal cancer irrespective of genetic risk. Gas-troenterology. 2018;155:1805–15.e5.

What's new?

Population attributable risk (PAR) is increasingly used for quantifying the preventability of cancer. However, estimations vary widely. This study assessed the robustness of PAR calculations for colorectal cancer in the US population through a direct comparison between the literature-based method and the low-risk method using the same risk factors, and generally found high concordance. Additional sensitivity analyses, however, showed that for the literature-based method, the choice of risk factors and relative risks could substantially influence PAR estimations. Future PAR studies using the literature-based method should consider presenting a range of estimates from different choices of risk factors and relative risks.

TABLE 1

Literature-based method: risk factors associated with increased colorectal cancer incidence considered in this study[a]

Exposure	Exposure category[b]	Cancer site (ICD-10)
Physical inactivity	0 to 249 MET-m/wk	Colon (C18)
	250 to 499 MET-m/wk	
	500 to 749 MET-m/wk	
	750 to 999 MET-m/wk	
	≥1000 MET-m/wk	
Body fatness/BMI	**Normal (18.5 to <25 kg/m ^{2})**	Colorectum (C18-C20)
	Overweight (25 to 29.9 kg/m^2)	
	Obese (≥30 kg/m2)	
Alcoholic drinks	**None**	Colorectum (C18-C20)
	<1 drink/d	
	1 to <4 drinks/d	
	≥4 drinks/d	
Red meat	**0 to 9 g/d**	Colorectum (C18-C20)
	10 to 24 g/d	
	25 to 49 g/d	
	50 to 74 g/d	
	75 to 99 g/d	
	≥100 g/d	
Processed meat	**0 to 4 g/d**	Colorectum (C18-C20)
	5 to 24 g/d	
	25 to 49 g/d	
	50 to 74 g/d	
	≥75 g/d	
Low dietary fiber	0 to 9 g/d	Colorectum (C18-C20)
	10 to 19 g/d	
	20 to 29 g/d	
	≥30 g/d	
Low dietary calcium	0 to 199 mg/d	Colorectum (C18-C20)
	200 to 399 mg/d	
	400 to 599 mg/d	
	600 to 799 mg/d	
	800 to 999 mg/d	
	≥1000 mg/d	
Cigarette smoking	**Never**	Colorectum (C18-C20)
	Former	
	Current	

[a]Based on summary of strong evidence on diet, nutrition, physical activity and the prevention of cancer 2018 by the WCRF/AICR (physical activity, body fatness/BMI, alcoholic drinks, and processed meat are “convincing” factors; red meat, dietary fiber, and dietary calcium are

"probable" factors) plus cigarette smoking; whole grain (and colorectal cancer) excluded because it can be attributed to fiber; dairy (and colorectal cancer) excluded because it can be attributed to calcium.

[b] Theoretical minimum is mentioned in bold.

TABLE 2

Population attributable risk (PAR) for colorectal cancer in the United States

Exposure[a]	**PAR, %**	
Literature-based method	**Men**	**Women**
Physical activity	10	11
Bodyfatness/BMI	17	8
Alcoholic drinks	13	2
Red meat	9	5
Processed meat	12	6
Dietary fiber	18	17
Dietary calcium	6	10
Cigarette smoking	12	12
Total w/o smoking	60	47
Total estimate	65	53
	PAR, %	
Low-risk method[b]	**Men (HPFS)**	**Women (NHS)**
Healthy lifestyle score ≥7	62	49

[a]Exposure categories: physical activity (0–249 MET-m/wk, 250–499 MET-m/wk, 500–749 MET-m/wk, 750–999 MET-m/wk, ≥ 1000 MET-m/wk); body fatness/BMI (normal (18.5 to <25 kg/m^2), overweight (25–29.9 kg/m^2), obese (≥ 30 kg/m^2)); alcoholic drinks (none, <1 drink/d, 1 to <4 drinks/d, ≥ 4 drinks/d); red meat (0–9 g/d, 10–24 g/d, 25–49 g/d, 50–74 g/d, 75–99 g/d, ≥100 g/d); processed meat (0–4 g/d, 5–24 g/d, 25–49 g/d, 50–74 g/d, ≥75 g/d); dietary fiber (0–9 g/d, 10–19 g/d, 20–29 g/d, ≥30 g/d); dietary calcium (0–199 mg/d, 200–399 mg/d, 400–599 mg/d, 600–799 mg/d, 800–999 mg/d, ≥1000 mg/d); cigarette smoking (never, former, current).

[b]Considering physical activity, BMI, alcohol intake, red meat intake, processed meat intake, total dietary fiber intake, total calcium intake and cigarette smoking.

TABLE 3

Additional analysis: population attributable risk (PAR) for colorectal cancer in the United States, excluding probable factors based on the WCRF/AICR report

Exposure[a]	**PAR, %**	
Literature-based method	**Men**	**Women**
Physical activity	10	11
Body fatness/BMI	17	8
Alcoholic drinks	13	2
Processed meat	12	6
Cigarette smoking	12	12
Total w/o smoking	43	25
Total estimate	50	34

[a]Exposure categories same as in Table 2.

TABLE 4

Additional analysis: minimum and maximum population attributable risk (PAR) for colorectal cancer in the United States, derived from different sources of relative risks

Exposure[a]	**(Minimum PAR, maximum PAR), %**	
Literature-based method	**Men**	**Women**
Physical activity	(3, 10)	(3, 11)
Body fatness/BMI	(17, 36)	(4, 21)
Alcoholic drinks	(13)	(2)
Red meat	(8, 9)	(4, 5)
Processed meat	(12, 14)	(6, 7)
Dietary fiber	(13, 18)	(12, 17)
Dietary calcium	(5, 6)	(7, 10)
Cigarette smoking	(8, 12)	(6, 12)
Total w/o smoking	(53, 70)	(33, 55)
Total estimate	(57, 74)	(37, 60)

[a]Exposure categories same as in Table 2.