RESEARCH ARTICLE

REVISED Population attributable risk for colorectal and breast cancer in England, Wales, Scotland, Northern Ireland, and the United Kingdom [version 2; peer review: 2 approved]

Shatabdi Goon[1,2*], Hanseul Kim[3*], Edward L. Giovannucci[3,4]

[1]Department of Human Nutrition, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA
[2]Florida Department of Health, Pasco, FL, 34652, USA
[3]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, 02115, USA
[4]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, 02115, USA

* Equal contributors

v2 **First published:** 23 Mar 2021, **3**:11
https://doi.org/10.12688/amrcopenres.12980.1
Latest published: 07 Mar 2022, **3**:11
https://doi.org/10.12688/amrcopenres.12980.2

Abstract

Background: The population attributable risk (PAR) is a statistic commonly used for quantifying preventability of cancer. We report here PAR estimates for the United Kingdom (UK) along with its constituent countries for up-to-date risk factor-attributable colorectal cancer (CRC) and breast cancer (BC), focusing on diet and nutrition related factors and tobacco (CRC) using representative national surveys.

Methods: The PAR was calculated using established, modifiable risk factors by the World Cancer Research Fund/American Institute of Cancer Research (WCRF/AICR): physical activity, body mass index (BMI), alcoholic drinks, red meat, processed meat, dietary fiber, dietary calcium, as well as cigarette smoking for CRC, and physical activity, BMI, alcoholic drinks, and fruits and vegetable consumption for BC. National prevalence estimates and relative risks (RRs) for CRC and BC were obtained from meta-analyses or large pooled analyses.

Results: Based on eight dietary and lifestyle risk factors, the estimates for attributable cases of CRC for males and females, respectively, were as follows: England: 67% and 60%; Scotland: 68% and 59%, Wales: 66% and 61%; Northern Ireland: 67% and 61%; and UK: 67% and 60%. Excluding smoking, the PAR for the UK was 61% for men and 52% for women. Based on four dietary and lifestyle risk factors, the estimates for BC were as follows: England: 26%, Scotland: 27%; Wales: 25%; Northern Ireland: 26%; and UK: 27%.

Conclusion: Up to 67% for CRC and 27% of BC were attributable to modifiable dietary and lifestyle factors in the UK. Moderate differences in PAR are observed between countries due to different prevalence of exposure to risk factors.

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 07 Mar 2022	✓ view	
version 1 23 Mar 2021	✗ view	✓ view

1. **Yibing Ruan**, Alberta Health Services, Calgary, Canada

2. **María Teresa Álvarez Bañuelos**, Universidad Veracruzana, Veracruz, Mexico

Any reports and responses or comments on the article can be found at the end of the article.

Keywords
Population Attributable Risk, Colorectal Cancer, Breast Cancer, Risk Factors, Exposure Prevalence, Relative Risk, Cancer Prevention, United Kingdom.

Corresponding authors: Shatabdi Goon (shatabdi.goon@gmail.com), Edward L. Giovannucci (egiovann@hsph.harvard.edu)

Author roles: Goon S: Formal Analysis, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Kim H**: Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Giovannucci EL**: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the World Cancer Research Fund International (Commissioned to E.L. Giovannucci). 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.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Goon S, Kim H and Giovannucci EL. **Population attributable risk for colorectal and breast cancer in England, Wales, Scotland, Northern Ireland, and the United Kingdom [version 2; peer review: 2 approved]** AMRC Open Research 2022, **3**:11 https://doi.org/10.12688/amrcopenres.12980.2

First published: 23 Mar 2021, **3**:11 https://doi.org/10.12688/amrcopenres.12980.1

REVISED **Amendments from Version 1**

This revised version has fixed some typos (i.e., in the old version it was mentioned as METS-hour/week instead of METS-m/week; it has been fixed in the revised version), clarified some information on the exposure data sources that were used to calculate the population attributable risk (PAR) (i.e., using of UK Biobank and EPIC cohort as a source of some risk factors, for example, meat and calcium were clearly mentioned in the footnotes of the table). Some text has been added to the discussion section to acknowledge some limitations regarding using exposure data sources that are not nationally representative.

Any further responses from the reviewers can be found at the end of the article

Introduction

The population attributable risk (PAR) is a metric commonly utilized to quantify the preventability of a specific disease. Various approaches to calculate the PAR for cancer have been used by researchers. Most typically, the PAR is based on identifying relative risks (RRs) from the scientific literature and prevalence from suitable population sources for specific factors of interest. The process requires several steps. It is critical to first identify the established risk factors that are in principle modifiable. Then it is important to derive a RR estimate from the literature for each risk factor. The RR estimates can either be from the population of interest, or common RRs, such as from meta-analyses or representative studies, that are generalizable. Next, it is important to identify estimates of the prevalence of exposures in the population of interest. With this information, standard formulae to calculate PARs can be used.

Estimates of the PAR for various cancers have often varied widely. Four decades ago, Doll and Peto[1] suggested that 90% of colorectal cancers (CRCs) and 50% of breast cancers (BCs) may be related to diet. Similarly, Parkin *et al.*[2] attributed about 43% of CRC in the UK to five largely modifiable factors including diet and nutrition, and 42% of BC to modifiable risk factors including body fatness and physical activity. PAR estimates for cancer from studies across the world have varied substantially; for example, some estimates were between 16% and 90% for CRC[3–5], and 6.5% and 50% for BC[1–3,6]. Blot and Tarone[7] argued that an estimate of 90% appeared to be too high for CRC. Such variability in the calculated PAR estimates may extend from a number of factors[8] including: 1) the risk factors that were considered; 2) the specific RRs that were utilized; 3) the sources of population prevalence of the risk factors; and 4) the specific calculation methods used to calculate the PAR. Other contributors to variability include variation in the time period and geography, differences in socio-demographic profile of cancer cases, and differences due to screening availability.

We report here PAR estimates for the United Kingdom (UK) along with its constituent countries on the number of risk factor-attributable CRC and BC, the two most common preventable cancers, excluding lung cancer, which is highly related to tobacco use. We focused on diet- and nutrition-related factors (and tobacco for CRC) based on up-to-date criteria from the World Cancer Research Fund/American Institute of Cancer Research (WCRF/AICR)[9]. We used representative national surveys for risk factor prevalence in individual UK countries given that the risk factor prevalence varies[10]. Hence, we estimated the PARs for CRC and BC in the UK, and separately, England, Scotland, Wales, and Northern Ireland.

Methods

Factor selection

Based on criteria from the WCRF/AICR[9], we used only factors classified as having achieved a level of evidence for a causal association as "probable" or "convincing" (displayed in Table 1). This level of evidence is considered "actionable" by the WCRF/AICR. In addition to the seven factors that met this classification for CRC (body mass index (BMI), physical activity, alcoholic drinks, and intakes of dietary fiber, red meat, processed meat, and dietary calcium), we further included cigarette smoking, considered a causal factor for CRC based on the United States Surgeon General Report[11]. We included four factors for BC (body fatness/BMI, physical activity, alcoholic drinks, and intakes of fruits and vegetables). We included "fruits and vegetables" because, although these were considered a limited/suggestive factor for BC, the WCRF/AICR expert panel concluded that that evidence for carotenoid rich foods was stronger for estrogen receptor negative BC than for estrogen receptor positive or total BC. Yet, because estrogen receptor negative BC is an important component of total BC, preventing it would contribute to reducing the burden of BC.

Prevalence data

The prevalence of exposures to risk factors in the populations was obtained from the nationally representative population surveys: the Health Survey of England, Scotland, Wales, and Northern Ireland[12]. Exposure distribution data on dietary calcium and meat consumption were obtained from EPIC-Oxford cohort[13] and UK biobank cohort[14], respectively. The methods detailing the derivation of prevalence for most exposures have been described in Brown *et al.*[12].

Relative risks

We identified RRs by conducting systematic searches in PubMed. Meta-analyses of cohort studies were the preferred source of RRs, then followed by pooled analyses of cohort studies and individual cohort studies. In some meta- and pooled analyses multiple estimates were reported, and sometimes more than one meta- or pooled analysis was available; in these circumstances, we selected the RRs based on characteristics most relevant to our study. In addition, selected RRs had to provide cut-points for the categories comparable to the exposure data available for UK and its constituents countries. The search string for the risk factors were: Tobacco (tobacco OR cigarette OR smoking OR environmental tobacco smoke OR second-hand smoke), Overweight and obesity (weight OR BMI OR body mass index OR obesity OR obese OR overweight OR adiposity OR body size), Alcohol (alcohol OR alcoholic OR ethanol), Fiber (fibre OR fiber), Processed and red meat (Meat OR bacon OR ham OR sausages OR jerky OR

Table 1. Risk factors associated with increased colorectal cancer and breast cancer incidence considered in this study.

Exposure	Exposure category**	Cancer site (ICD-10)
Colorectal cancer*		
Cigarette smoking	**Never** Former Current	Colorectum (C18-C20)
Body fatness/BMI	**Normal (18.5-< 25 kg/m²)** Overweight (25-29.9 kg/m²) Obese (≥ 30 kg/m²)	Colorectum (C18-C20)
Alcoholic drink	**None** Light (0.1-12.5 g/day) Moderate (12.6-50 g/day) Heavy (> 50 g/day)	Colorectum (C18-C20)
Insufficient physical activity	Inactive, less than 5 days of at least 30 minutes activity per week or < 600 MET-m/week **Active, more than 5 days of at least 30 minutes of activity or ≥ 600 MET-m/week**	Colon (C18)
Processed meat	< 25 g/day vs. ≥ 25 g/day	Colorectum (C18-C20)
Red meat	< 70g/day vs. ≥ 70 g/day	Colorectum (C18-C20)
Insufficient dietary fiber	per 1 gm deficit per day **≥ 30 g/day**	*Colorectum (C18-C20)*
Low dietary calcium	0-524 mg/day 525-699 mg/day 700-899 mg/day 900-1000 mg/day **≥ 1000 mg/day**	*Colorectum (C18-C20)*
Breast cancer*		
Body fatness/BMI	**Normal (18.5-< 25 kg/m²)** Overweight (25-29.9 kg/m²) Obese (≥ 30 kg/m²)	Breast (C50)
Alcoholic drink	**None** Light (0.1-12.5 g/day) Moderate (12.6-50 g/day) Heavy (> 50 g/day)	Breast (C50)
Insufficient physical activity	Inactive, less than 5 days of at least 30 minutes activity per week or < 600 MET-m/week **Active, more than 5 days of at least 30 minutes of activity or ≥ 600 MET-m/week**	Breast (C50)
Insufficient fruits and vegetables	< 5 servings **≥ 5 servings or ≥400 g/day**	Breast (C50)

*For colorectal cancer, exposure selected based on summary of strong evidence on diet, nutrition, physical activity and the prevention of cancer 2018 by WCRF/AICR (probable and convincing; Italicized factors 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. For breast cancer, exposure selected based on summary of strong evidence on diet, nutrition, physical activity, and the prevention of breast cancer 2017 by WCRF/AICR (probable and convincing; Italicized factor is probable, underlined factor is limited/suggestive). Dietary fiber (and breast cancer) excluded because it can be attributed to fruits and vegetables combined.

**Theoretical minimum in bold

salami OR cured OR salted OR Red), Insufficient physical activity (physical OR activity OR exercise OR physically active OR sedentary), Calcium (Calcium OR Dairy), and Fruits and vegetables (Fruits OR Vegetables OR Fruits and Vegetables).

As described previously for CRC[8], we used the formula, $\exp(\frac{ln(X)}{A} * B)$, to calculate RRs of different units for dietary calcium and physical activity. For calcium and CRC, we calculated RRs for each of the following categories: 0–199 mg/day, 200–399 mg/day, 400–599 mg/day, 600–799 mg/day, 800–999 mg/day, and ≥ 1000 mg/day[8]. To match the categories available for the UK prevalence data (Oxford-UK Cohort), we then computed average RRs for overlapping categories to match with the exposure prevalence data available on resulting the RRs for categories: 0–524 mg/day, 525–699 mg/day, 700–899 mg/day, 900–1000 mg/day, and >1000 mg/day (e.g., calculating an average RR for 0–199 mg/day, 200–399 mg/day, and 400–599 mg/day to match 0–524 mg/day). Similarly, for physical activity and CRC, we first calculated RRs for each of the categories: 0–249 MET-m/week, 250–499 MET-m/week, 500–749 MET-m/week, 750–999 MET-m/week, and ≥1000 MET-m/week[8]. We then computed an average RR for the first two categories representing: achieving less than 600 METs-m/week or active less than 5 days per week/ not achieving 30 minutes of physical activity on 5 days per week and active more than 5 days per week[8], that matches well with the exposure prevalence data on the UK national health surveys (prevalence of physical activity was provided as days per week on which at least 30 minutes of moderate physical activity was completed). For physical activity and CRC, we took a weighted average of colon and rectal cancers (70% colon cancer and 30% rectal cancer[15]) to calculate the RR. Such assumptions made for RRs could have resulted in some small differences in the results.

For the factors associated with increased risk (BMI, alcohol, red meat, processed meat), we used the lowest category as the reference group. For the protective factors (physical activity, fiber, fruits and vegetable, and calcium), we chose the highest category as the reference group, and PAR was calculated using the reciprocal of the RR. For studies that provided a linear dose-response relationship, the RRs were first transformed into a log scale, divided by the value, then exponentiated.

Statistical analysis

For each of the risk factors, we identified n levels of exposure categories. We then estimated PAR using the following equation:

$$\frac{\sum_{i=1}^{n} P_i * (RR_i - 1)}{\sum_{i=1}^{n} P_i * (RR_i - 1) + 1}$$

where P_i is the exposure prevalence at the exposure category i and RR_i is the corresponding RR of CRC or BC at exposure category i. The details for the categorizations of exposures are presented in Table 1.

We then estimated the preventability of CRC or BC that was attributable to the combined dietary and lifestyle risk factors the following equation:

$$\text{PAR} = 1 - \prod_{i=1}^{n} (1 - PAR_i)$$

where i signifies the level of individual risk factors (i=1,…, n).

For fiber intake, the PAR was directly obtained from Brown *et al.* because our estimate was based on the same prevalence data and RR[12]. The PAR estimates were directly computed manually using the formulae.

Sensitivity analysis

We conducted an additional analysis where we exclude probable factors (defined by WCRF/AICR) and kept only the convincing factors from the calculation of the PAR. For CRC, the three probable factors (red meat, dietary fiber, and dietary calcium) were excluded from the analysis. We calculated the proportion of BC attributable to lifestyle factors excluding suggestive (fruits and vegetable consumption) and probable factors (physical activity), resulting in these two factors: body fatness/BMI and alcoholic drinks.

Results

The proportion of CRC cases attributable to lifestyle risk factors for the UK constituents' countries were estimated as follows: 67% for British males and 60% for British females, 68% for Scottish males and 59% for Scottish females, 66% for Welsh males and 61% for Welsh females, 67% for Northern Irish males and 61% for Northern Irish females, and 67% for UK males and 60% for UK females, overall (Table 2). The proportion of CRC cases attributable to lifestyle risk factors excluding cigarette smoking were 62% for British males and 53% for British females, 62% for Scottish males and 51% for Scottish females, 61% for Welsh males and 52% for Welsh females, 62% for Northern Irish males and 52% for Northern Irish females, and 61% for UK males and 52% for UK females (Table 2). The intake of dietary fiber was the major contributor to the attributable CRC cases, accounting for 25% for males and 32% for females in the UK, followed by processed meat intake (14% for men and 10% for women). For alcoholic drinks, the PAR values were substantially higher for men than for women.

The proportions of BC cases attributable to lifestyle risk factors were 26% for British women, 27% for Scottish, 25% for Welsh, 26% for Northern Irish, and 27% for UK women, overall (Table 2). Alcohol was the largest contributor to the estimated attributable BC cases, accounting for 8% for females in UK, followed by the body fatness (7.6%) and insufficient physical activity (7%).

The estimates for the prevalence and the RRs for the various exposures used in our calculation are presented in Table 3 and Table 4.

Sensitivity analysis

The PARs for CRC based on the five "convincing" factors (body fatness/BMI, physical activity, alcohol, processed meat, and cigarette smoking) were 48% for UK males and 34% for UK females, and excluding smoking, these were

Table 2. Preventability estimates for colorectal cancer and breast cancer in England, Scotland, Wales, Northern Ireland, and United Kingdom (UK).

Exposure	Population attributable risk, %									
	England		Scotland		Wales		Northern Ireland		UK	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Colorectal cancer*										
Cigarette smoking	14.1	15.5	14.2	15.7	14.4	18.6	13.3	15.0	14.0	16.2
Physical activity	7.0	4.8	7.5	5.6	6.4	4.2	6.0	5.1	6.7	5.1
Body fatness/BMI	13.5	6.0	13.3	5.7	12.3	5.8	13.9	5.7	13.3	5.6
Alcoholic drinks	13.6	2.2	13.3	1.9	13.4	2.0	13.5	1.9	13.0	1.9
Processed meat	14.5	10.0	14.5	8.5	14.5	7.4	14.5	9.6	14.5	8.5
Red meat	9.0	7.4	9.0	7.4	9.0	7.4	9.0	7.4	9.0	7.4
Dietary fiber	24.6	31.6	26.8	33.3	25.2	32.2	26.9	33.3	24.9	31.8
Dietary calcium	6.7	7.4	6.7	7.4	6.7	7.4	6.7	7.4	6.7	7.4
Total (w/o smoking)	**61.6**	**53.3**	**62.4**	**51.4**	**60.8**	**52.0**	**62.1**	**52.4**	**61.2**	**52.4**
Total estimate	**67.0**	**60.0**	**67.8**	**59.0**	**66.3**	**61.1**	**67.4**	**61.0**	**66.6**	**60.0**
Breast cancer**										
Body fatness/BMI	NA	7.1	NA	6.7	NA	6.1	NA	6.6	NA	7.6
Alcoholic drink	NA	8.5	NA	7.8	NA	8.3	NA	8.2	NA	8.2
Physical activity	NA	7.3	NA	8.2	NA	6.3	NA	7.5	NA	7.5
Fruits and vegetables	NA	6.4	NA	7.2	NA	7.0	NA	6.8	NA	6.8
Total estimate	**NA**	**26.2**	**NA**	**26.7**	**NA**	**25.0**	**NA**	**26.1**	**NA**	**26.9**

*Exposure categories: Physical activity (less than 5 days of at least 30 minutes activity per week or <600_METs-m/week vs. more than 5 days of at least 30 minutes activity per week or ≥ 600 MET-m/week); Body fatness/BMI (Normal (18.5-< 25 kg/m^2), Overweight (25-29.9 kg/m^2), Obese (≥ 30 kg/m^2)); Alcoholic drink (None, light (0.1-12.5 g/day), moderate (12.6-50 g/day), heavy (> 50 g/day)); Red meat (<70 g/day vs. ≥ 70 g/day); Processed meat (<25 g/day vs. ≥ 25 g/day); Dietary fiber (per 1 gm deficit per day, ≥ 30 g/day); Dietary calcium (<700 mg/day vs. ≥ 700 mg/day); Cigarette smoking (Never, Former, Current)

**Exposure categories: Physical activity (less than 5 days of at least 30 minutes activity per week or <600_MET-m/week vs. more than 5 days of at least 30 minutes activity per week or ≥ 600 MET-m/week); Body fatness/BMI (Normal (18.5-< 25 kg/m^2), Overweight (25-29.9 kg/m^2), Obese (≥ 30 kg/m^2)); Alcoholic drink (None, light (0.1-12.5 g/day), moderate (12.6-50 g/day), heavy (> 50 g/day)); fruits and vegetables (<5 servings/day vs. ≥5 servings or 400g/day)

40% and 21%, respectively (Table 5). After excluding the "probable" factors for BC, the PARs including the 2 factors (body fatness/BMI and alcoholic drinks) were 15% for UK females (Table 5).

Discussion

We provided PAR estimates for the UK and its constituent country-level for diet and lifestyle risk factors where evidence for a causal role in CRC and BC development is probable/convincing based on WCRF/AICR systematic reviews. We estimated that 67% of CRC cases in men and 60% of cases for women in the UK, and 27% of BC cases were attributable to dietary and lifestyle risk factors assessed in adulthood. Excluding smoking from the calculation, these estimates for CRC were 61% for males and 52% for females. Significant differences in the CRC PAR by sex were observed for body fatness, alcohol, and fiber intake; these results were mostly driven by sex differences in the RR estimates, and partly by higher levels of alcohol drinking in men. Moderate differences in the PAR estimates were observed between countries due to different prevalence of exposure to risk factors. For instance, prevalence of obesity/body fatness was slightly lower in Wales compared to other countries in the UK, resulting in slightly lower PAR estimates of both CRC and BC for Wales overall. Similarly, the highest prevalence of not achieving 150 minutes of weekly moderate physical activity was observed among Scottish population, resulting in slightly higher PAR estimates of both CRC and BC in Scotland.

Table 3. Distribution of colorectal cancer and breast cancer exposures in the UK by country and sex*.

Exposure	Exposure category/unit	Prevalence, %									
		England		Scotland		Wales		Northern Ireland		UK	
		Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
BMI	Normal (18.5-< 25 kg/m^2)	31	40	37	35	35	41	38	38	36.9	46.5
	Overweight (25-29.9 kg/m^2)	43	33	41	31	42	32	39	30	41.3	31.5
	Obese (≥ 30 kg/m^2)	22	24	22	23	18	18	25	23	21.8	22
Cigarette smoking	Never	4	56	45	55	45	46	50	62	46	54.7
	Former	28	20	27	21	26	24	23	13	26	19.5
	Current	27	24	28	24	29	30	27	25	27.8	25.8
Alcohol	None	16	20	27	41	12	20	24	30	9.4	17.4
	Light (0.1-12.5 g/day)	44	54	42	58	45	57	44	56	43.8	56
	Moderate (12.6-50 g/day)	34	26	35	24	38	24	36	25	35.8	24.8
	Heavy (> 50 g/day)	12	2	11	1	10	2	11	2	11	1.8
Physical activity	% not achieving 150+ minutes moderate physical activity per week or active 5 days/week	39	27	43	31	36	23	33	28	38	28
	% achieving 150+ minutes moderate physical activity per week or active 5 days/week	61	73	57	69	64	77	67	72	62	73
Dietary fiber	g per day	20	16	19	15	19	16	18	15	19	15.5
Processed meat	% eating processed meat ≥25 g per day	74	50	74	50	74	50	74	50	74	50
Red meat	% eating red meat ≥70 g per day	55	47	55	47	55	47	55	47	55	47
Dietary calcium	0-524 mg/day	5.9	6.4	5.9	6.4	5.9	6.4	5.9	6.4	5.9	6.4
	525-699 mg/day	11.5	12.0	11.5	12.0	11.5	12.0	11.5	12.0	11.5	12.0
	700-899 mg/day	19.1	22.3	19.1	22.3	19.1	22.3	19.1	22.3	19.1	22.3
	900-1000 mg/day	11.0	11.6	11.0	11.6	11.0	11.6	11.0	11.6	11.0	11.6
	≥ 1000 mg/day	52.4	47.8	52.4	47.8	52.4	47.8	52.4	47.8	52.4	47.8
Fruits and vegetables (g/day)	% not eating five or more portions or 400g of fruit and vegetables per day	72	68	83	78	78	75	85	73	80	73

*These estimates were prevalence calculated by Brown *et.al.* 2015[12]. Exposure distribution data was obtained from the Health survey of England, Scotland, Wales, and Northern Ireland, and National Diet and Nutrition Survey[16] except for dietary calcium and meat. Exposure distribution data on dietary calcium was obtained from the EPIC-oxford cohort[13]. Exposure distribution data on meat consumption was obtained from the UK biobank cohort[14].

Previous studies have considered the PAR in the UK and its individual countries. The results we report here show overall consistency with those from similar studies, despite some differences in the risk factors considered, the time periods encompassed, and the RR sources. In one study to determine preventability in the UK (2018) adults aged 30 years and older, the PAR of CRC attributable to tobacco smoking, alcohol, intakes of meat and fiber, overweight and obesity, physical exercise, and ionizing radiation was 57% for men and 51% for women[12]. In the same study, the PAR of BC attributable to alcohol drinking, intakes of fruits and vegetables, overweight and obesity was 27% in 2010[2] and 23% in 2018[12]. These results were similar but slightly lower compared with our estimates. This other analysis[12] did not account for red meat or dietary calcium, considered as probable risk factors by WCRF/AICR, and included radiation and oral contraceptive use, which we did not consider as our analysis focused on the most up-to-date evidence from the WCRF/AICR for modifiable cancers under the domain of diet and nutrition.

Table 4. **Relative risks for the associations between risk factors and colorectal cancer and breast cancer.**

	In original reports				In this analysis when our exposure categories were different from those in original reports		
			Colorectal Cancer				
Exposure	**Exposure category/unit**	**Region; Study design**	**Relative risk (95% CI)**		**Exposure category/unit**	**Relative risk (95% CI)**	
			Men	**Women**		**Men**	**Women**
Cigarette smoking[17]	Never Former Current	Worldwide, Meta-analysis	Reference 1.20 (1.10, 1.30) 1.40 (1.20, 1.70)	Reference 1.20 (1.10, 1.30) 1.60 (1.40, 1.90)	same	same	same
BMI[18]	Normal (18.5-< 25 kg/m^2) Overweight (25-29.9 kg/m^2) Obese (≥ 30 kg/m^2)	Worldwide, Meta-analysis	Reference 1.17 (1.12, 1.22) 1.38 (1.32, 1.44)	Reference 1.07 (1.01, 1.14) 1.17 (1.06, 1.30)	same	same	same
Alcohol[19]	None Light (0.1-12.5 g/day) Moderate (12.6-50 g/day) Heavy (> 50 g/day)	Worldwide, Meta-analysis	Reference 1.05 (0.95, 1.16) 1.21 (1.11, 1.32) 1.53 (1.30, 1.80)	Reference 1.00 (0.89, 1.01) 1.07 (0.99, 1.16) 1.24 (0.68, 2.25)	same	same	same
Physical activity[20]	Higher vs. lower (90th percentile vs. 10th percentile)	Worldwide, Meta-analysis	0.84 (0.77, 0.91)	0.84 (0.77, 0.91)	Inactive, less than 5 days of at least 30 minutes activity per week or < 600 MET-m/week Active, more than 5 days of at least 30 minutes of activity or ≥ 600 MET-m/week	1.19 Reference	1.19 Reference
Dietary fiber[3]	Per 1 gm deficit per day	Worldwide, Meta-analysis	1.03 (1.01,1.06)	1.03 (1.01,1.06)	same	same	same
Processed meat[21]	Per 50 g per day/increase	Worldwide, report	Colon: 1.23 (1.11, 1.35) Rectum: 1.08 (1.00, 1.18)	Colon: 1.23 (1.11, 1.35) Rectum: 1.08 (1.00, 1.18)	< 25 g/day ≥ 25 g/day	Reference 1.23	Reference 1.23
Red meat[21]	Per 100 g per day/increase	Worldwide, report	Colon: 1.22 (1.06, 1.39) Rectum: 1.13 (0.96, 1.34)	Colon: 1.22 (1.06, 1.39) Rectum: 1.13 (0.96, 1.34)	< 70 g/day ≥ 70 g/day	References 1.18	References 1.18
Dietary calcium[22]	Per 200 mg/day increase	Worldwide, Meta-analysis	0.93 (0.88, 0.99)	0.93 (0.91, 0.95)	0-524 mg/day 525-699 mg/day 700-899 mg/day 900-999 mg/day ≥ 1000 mg/day	1.34 1.20 1.12 1.04 Reference	1.33 1.20 1.12 1.06 Reference

	In original reports				In this analysis when our exposure categories were different from those in original reports		
Breast cancer							
BMI[23] Females	Normal (18.5-< 25 kg/m^2) Overweight (25-29.9 kg/m^2) Obese (≥ 30 kg/m^2)	Worldwide, Meta-analysis	NA	Reference 1.10 (1.06, 1.13) 1.18 (1.12, 1.25)	same	same	same
Alcohol[19] Females	None Light (0.1-12.5 g/day) Moderate (12.6-50 g/day) Heavy (> 50 g/day)	Worldwide, Meta-analysis	NA	Reference 1.04 (0.99, 1.09) 1.23 (1.16, 1.32) 1.60 (1.23, 2.19)	same	same	same
Physical activity[20]	Higher vs. lower (90th percentile vs. 10th percentile)	US and Europe; Pooled analysis	0.94 (0.90, 0.98)	0.94 (0.90, 0.98)	Inactive, less than 5 days of at least 30 minutes activity per week or < 600 MET-m/week Active, more than 5 days of at least 30 minutes of activity or ≥ 600 MET-m/week	1.29 Reference	1.29 Reference
Fruits and vegetables combined[24]	Highest vs. Lowest; ≥ 5 servings or 400 g/d vs. 100 g/d (no studies with zero intake)	Worldwide, Meta-analysis	0.89 (0.80, 0.99)	0.89 (0.80, 0.99)	same	same	same

Table 5. Additional analysis: Preventability estimates (PAR) for CRC and BC in the UK, excluding probable factors based on the WCRF/AICR.

Exposure*	Standard PAR, %									
	Colorectal Cancer									
	England		Scotland		Wales		Northern Ireland		UK	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Cigarette smoking	14.1	15.5	14.2	15.7	14.4	18.6	13.3	15.0	14.0	16.2
Physical activity	7.0	4.8	7.5	5.6	6.4	4.2	6.0	5.1	6.7	5.1
Body fatness/BMI	13.5	6.0	13.3	5.7	12.3	5.8	13.9	5.7	13.3	5.6
Alcoholic drink	13.6	2.2	13.3	1.9	13.4	2.0	13.5	1.9	13.0	1.9
Processed meat	14.5	10.0	14.5	10.0	14.5	10.0	14.5	10.0	14.5	10.0
Total (w/o smoking)	40.5	21.1	40.1	21.4	40.0	20.4	40.1	20.4	40.0	21.0
Total estimate	***49.0***	***33.4***	***48.5***	***33.7***	***47.8***	***35.0***	***48.1***	***32.4***	***48.3***	***33.7***
	Breast Cancer									
Body fatness/BMI	7.1		6.7		6.1		6.6		7.6	
Alcoholic drink	8.5		7.8		8.3		8.2		8.2	
Total estimate	***14.9***		***13.9***		***13.9***		***14.3***		***15.2***	

*Exposure categories same as Table 1

Wide variations in PAR estimates were observed in some previous studies across the world. For example, in a study that examined overweight and obesity, physical inactivity, and low consumption of fruits and vegetables in relation to CRC risk worldwide, the estimated PAR was 13%[25]. A report of alcohol, obesity and overweight, and physical inactivity as modifiable risk factors associated with CRC risk found that the PAR was 19% for the French population, and 21% for men and 16% for women[5]. Although our estimates were for the UK rather than worldwide, they suggest that the PAR of 13% probably underestimates the true preventability because the overall CRC rates are not markedly different between France and the UK.

Our estimates of the PAR here were based on our judgement of the best available and most relevant data, and thus cautious interpretation is warranted. These limitations could bias either toward underestimation or overestimation of PARs. We have previously shown that the choice of risk factors and selection criteria for the sources of RRs could influence the PAR estimates, although in general, the methods appear relatively robust[8]. We further performed sensitivity analysis by excluding probable factors to assess the influence of risk factor selection on PAR estimation. Yet this analysis is likely to considerably underestimate the true PAR because the strength of evidence for "probable" risk factors is high enough to be considered actionable by the WCRF/AICR. The associations classified as "probable" by the WCRF/AICR are robust, but, nevertheless, because they are based on observational studies, confounding cannot be discounted with certainty. Another limitation is, even if the considered factors are truly causal, it is unclear when in life they need to be altered to mitigate risk.

On the other hand, some considerations indicate that we may underestimate the true preventability. Most studies utilize a single measure based on a dietary or physical activity assessment with considerable measurement error; thus, any true associations would tend to be underestimated because random error typically causes a true association to tend towards the null value. Although BMI is a useful measure and generally measured well, it is not a perfect measure of the most relevant component of adiposity (e.g., visceral fat). The PARs for alcohol are prone to be underestimates because non-drinkers (the presumed low-risk group) may include past drinkers who may have consumed heavily before stopping to drink[19]. Finally, the overall preventability over the life course is likely to be underestimated because we only considered adulthood diet. In fact, adolescence is emerging as an important time period of increased susceptibility to carcinogenic exposures, including diet, especially for BC but also for CRC.

In conclusion, our study reported the PAR of CRC and BC attributable to modifiable risk factors in England, Scotland, Wales, Northern Ireland, and UK. Up to 67% for CRC and 27% of BC were attributable to known established modifiable risk factors in adulthood in the UK. These results reinforce the importance for diet and lifestyle for the prevention of major cancers in the UK.

Data availability

All data underlying the results are available as part of the article and no additional source data are required.

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**(6): 1192–308.
PubMed Abstract | Publisher Full Text

2. 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): S77–81.
PubMed Abstract | Publisher Full Text | Free Full Text

3. Parkin DM, Boyd L: **6. Cancers attributable to dietary factors in the UK in 2010. III. Low consumption of fibre.** *Br J Cancer.* 2011; **105 Suppl 2**(Suppl 2): S27–30.
PubMed Abstract | Publisher Full Text | Free Full Text

4. 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**(1): 168.
PubMed Abstract | Publisher Full Text | Free Full Text

5. **IARC Publications Website - Attributable Causes of Cancer in France in the Year 2000.**
Reference Source

6. Mezzetti M, La Vecchia C, Decarli A, *et al.*: **Population attributable risk for breast cancer: Diet, nutrition, and physical exercise.** *J Natl Cancer Inst.* 1998; **90**(5): 389–94.
PubMed Abstract | Publisher Full Text

7. Blot WJ, Tarone RE: **Doll and Peto's quantitative estimates of cancer risks: Holding generally true for 35 years.** *J Natl Cancer Inst.* Oxford University Press; 2015; **107**(4): djv044.
PubMed Abstract | Publisher Full Text

8. Kim H, Wang K, Song M, *et al.*: **A comparison of methods in estimating population attributable risk for colorectal cancer in the United States.** *Int J Cancer.* 2021; ijc.33489.
PubMed Abstract | Publisher Full Text

9. World Cancer Research Fund International. In: The Grants Register 2019. 2019; 983–983.
Reference Source

10. **Scottish Health Survey - topic report: UK comparisons.** gov.scot.
Reference Source

11. National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health: **The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General.** *A Rep Surg Gen.* 2014; 1081.
PubMed Abstract

12. Brown KF, Rumgay H, Dunlop C, *et al.*: **The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015.** *Br J Cancer.* 2018; **118**(8): 1130–41.
PubMed Abstract | Publisher Full Text | Free Full Text

13. Key TJ, Appleby PN, Spencer EA, *et al.*: **Calcium, diet and fracture risk: a prospective study of 1898 incident fractures among 34 696 British women and men.** *Public Health Nutr.* 2007; **10**(11): 1314–20.
PubMed Abstract | Publisher Full Text

14. Bradbury KE, Murphy N, Key TJ: **Diet and colorectal cancer in UK Biobank: A prospective study.** *Int J Epidemiol.* 2020; **49**(1): 246–58.
PubMed Abstract | Publisher Full Text | Free Full Text

15. Siegel RL, Miller KD, Sauer AG, *et al.*: **Colorectal cancer statistics, 2020.** *CA Cancer J Clin.* 2020; **70**(3): 145–64.
PubMed Abstract | Publisher Full Text

16. Whitton C, Nicholson SK, Roberts C, *et al.*: **National Diet and Nutrition Survey: UK food consumption and nutrient intakes from the first year of the rolling programme and comparisons with previous surveys.** *Br J Nutr.* Europe PMC Funders; 2011; **106**(12): 1899–914.
PubMed Abstract | Publisher Full Text | Free Full Text

17. Carter BD, Abnet CC, Feskanich D, *et al.*: **Smoking and mortality--beyond established causes.** *N Engl J Med.* 2015; **372**(7): 631–40.
PubMed Abstract | Publisher Full Text

18. Xue K, Li FF, Chen YW, *et al.*: **Body mass index and the risk of cancer in women compared with men: A meta-analysis of prospective cohort studies.** *Eur J Cancer Prev.* 2017; **26**(1): 94–105.
PubMed Abstract | Publisher Full Text

19. 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**(3): 580–93.
PubMed Abstract | Publisher Full Text | Free Full Text

20. 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.* BMJ Publishing Group; 2016; **354**: i3857.
PubMed Abstract | Publisher Full Text | Free Full Text

21. **Diet, nutrition, physical activity and cancer: a global perspective.** A summary of the third ER 2018. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. Continuous Update Project Expert Report 2018. 2018; 1–53.
Reference Source

22. WCRFI: **Analysing research on cancer prevention and survival.** *World Cancer Res Fund Int Contin Updat Proj Summ Rep.* 2017; **24**.

23. Munsell MF, Sprague BL, Berry DA, *et al.*: **Body mass index and breast cancer risk according to postmenopausal estrogen-progestin use and hormone receptor status.** *Epidemiol Rev.* 2014; **36**(1): 114–36.
PubMed Abstract | Publisher Full Text | Free Full Text

24. Aune D, Chan DSM, Vieira AR, *et al.*: **Fruits, vegetables and breast cancer risk: A systematic review and meta-analysis of prospective studies.** *Breast Cancer Res Treat.* 2012; **134**(2): 479–93.
PubMed Abstract | Publisher Full Text

25. Danaei G, Hoorn SV, Lopez AD, *et al.*: **Causes of cancer in the world: Comparative risk assessment of nine behavioural and environmental risk factors.** *Lancet.* 2005; **366**(9499): 1784–93.
PubMed Abstract | Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 14 March 2022

https://doi.org/10.21956/amrcopenres.14143.r26951

✔ **Yibing Ruan**
Department of Cancer Epidemiology and Prevention Research, Alberta Health Services, Calgary, AB, Canada

The authors have made some revisions and have largely acknowledged the potential biases of their PAR estimates and the limitations of their study. I have no further comments.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Cancer epidemiology; Population attributable risk; Biostatistics; Molecular epidemiology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 24 September 2021

https://doi.org/10.21956/amrcopenres.14053.r26795

✔ **María Teresa Álvarez Bañuelos**
Instituto de Salud Pública, Universidad Veracruzana, Veracruz, Mexico

This reports here the population attributable risk (PAR) estimates for the United Kingdom along

with its constituent countries for up-to-date risk factor-attributable colorectal cancer (CRC) and breast cancer (BC), focusing on diet and nutrition related factors and tobacco (CRC) using representative national surveys.

The manuscript is well-written and nicely presented, with a good balance of descriptive text and speech and practical tables of the use of the package.

The methods used seem relatively robust but limited in their causality, they did not include incidence, although the study provides interesting data on the two important neoplasms in public health. However, it was limited to estimates of national prevalence only and the relative risks (RR) failed to include important socio-demographic risk factors such as age group for both CRC and CC.

Discussion. The limitations of the study were mentioned, which are cautious in making strong statements regarding the results.

An important question would be about the validity of the study, the text does not say anything and about the robustness of the study, or if the research question was important and if the study it was original.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 20 July 2021

https://doi.org/10.21956/amrcopenres.14053.r26735

Yibing Ruan
Department of Cancer Epidemiology and Prevention Research, Alberta Health Services, Calgary, AB, Canada

The authors provided some PAR estimates for the dietary and lifestyle risk factors associated with CRC and BC in UK along with its constituent countries. I couldn't help wondering about the rationale underlying this study and several methodologic issues after reading this manuscript.

1. The rationale of carrying out this study eludes me, especially a large chunk of the result was just a repetition of Brown et al. (Br J Cancer, 2018, 118(8): 1130–41[1]). The PAR for individual and combined risk factors of CRC and BC can be found in the supplementary materials of Brown's paper, except for dietary calcium and red meat. Was the rationale of this study to update the PAR estimates in UK by including these two risk factors, which were excluded in Brown's paper?

2. The authors obtained exposure distribution data on dietary calcium and red meat consumption from EPIC-Oxford cohort and UK biobank cohort (page 3). However, these two cohorts are not population representative of UK. For example, the EPIC-Oxford cohort clearly stated on their website that the sample is not representative of the general population, as 50% of their sample are vegetarians. (http://www.epic-oxford.org/faq/1262/frequently-asked-questions). Similarly, UK Biobank participants were more likely to be older, to be female, and to live in less socioeconomically deprived areas than nonparticipants (Fry A. et al., Am J Epidemiol. 2017, 186(9):1026-34[2]). Therefore, the PAR estimates of calcium and red meat are bound to be biased using the data.

3. I wonder why the authors did not obtain cancer incidence data from the cancer registries in UK. Without the incidence data, there are at least three limitations regarding the PAR estimates from this study. First, the authors could not estimate the attributable cancer cases, which is a constituent part of PAR estimation. Second, the authors could not do an age-stratified analysis. The summation of attributable cases of the age groups, divided by the cancer incidence of all age groups, provides a more accurate all-age PAR estimate than directly using an all-age prevalence to estimate PAR (the latter of which I suspect is what the authors did). This is because some risk factors are more prevalent among older age groups (e.g., physical inactivity, body fatness) and most cancer incidences are disproportionally higher among older age groups. Second, the authors had to assume a weighted average of 70% colon and 30% rectal cancers for physical activity and CRC (page 5), which wouldn't be necessary if they had the incidence data by colon and rectum.

4. The authors didn't correctly estimate the PAR of body fatness and physical activity for BC. These two risk factors are only associated with postmenopausal breast cancers, whereas the authors included premenopausal breast cancers as well.

5. The way the authors calculated RRs for calcium and physical activity is difficult to understand. They used a formula that assumes a loglinear dose-response relationship and

calculates RR based on per unit RR, which is a usual practice in this field. Then for dietary calcium, they calculated an average RR for 0–199 mg/day, 200–399 mg/day, and 400–599 mg/day from their IJC paper (Kim H. et al., Int J Cancer, 2021, 148(12):2947-53[3]) to match 0–524 mg/day of the lowest category of exposure in this study (page 5, first paragraph). It was unclear to me how they averaged the RR and how they resolved the issue that the highest category (400-599mg/day) exceeds the upper limit (524 mg/day). Why not just take the midpoint value of this category (262 mg/day) and combine it with the risk per unit value? For example, for a reference calcium level of 1000 mg/day and per mg/day unit RR of r_mg, the RR for the category of 0-524 mg/day is r_mg^(1000-262).

Some other concerns:

1. The authors used a formula, exp(lnX/A *B) to calculate RR, while neither A nor B was defined. The authors referenced their recent publication in IJC for this formula (Kim H. et al., Int J Cancer, 2021, 148(12):2947-53[3]). But in their IJC paper, the formula was laid out as "RR_B = exp(log(RR_B)/A *B)", which is simply wrong. "log(RR_B)" should be "log(RR_A)". The authors should write an erratum to IJC to have it corrected.

2. In the second paragraph of the Method section, the author stated, "the prevalence of exposures to risk factors in the populations was obtained from the nationally representative population surveys" (page 3). However, which risk factors exactly were not stated. This caused some confusions to me, as I tried to figure out from which data source they obtained the prevalence of processed meat consumption.

3. In Table 2, there is no variation in PAR by countries for red meat and dietary calcium (page 6). It is unclear whether there is indeed no country-level variation for these two risk factors, or there were no prevalence data to support this analysis, or there was a data-entry mistake. The authors didn't have any discussions around it. If there were no prevalence data to support this analysis, then the discussion around differences in PAR between countries should acknowledge this limitation.

4. At the end of the second paragraph of introduction (page 3), the authors stated that other contributors to the variability in calculated PAR estimates by different studies differences include "socio-demographic profile of cancer cases, and differences due to screening availability". As the authors didn't provide citations to support this statement, could the authors further elaborate on how these two factors contribute to the variability? In particular, it is difficult to fathom how screening availability, a factor related to the secondary intervention of CRC and BC, contributes to the PAR of diet and nutrition.

5. The authors used an equation that has been widely adopted for estimating combined risk factors (page 5, paragraph 4). However, many researchers used this formula without disclaiming the two important assumptions associated with it (Steenland and Armstrong, 2006, Epidemiology, 17(5):512-9[4]): independent risk exposure distributions and no statistical interactions between any two risks. These two assumptions should not only be explicitly stated in the method section, but also be worthy of several sentences in the discussion section. Also on page 5 after the formula, the author said "where i signifies the level of individual categories (i = 1,..., n)." "Categories" here should be "risk factors". Again for this paragraph, the first sentence "We then estimated the preventability of CRC or BC that was attributable to the combined dietary and lifestyle risk factors the following equation" is missing a preposition.

6. Page 5 line 10, “600 METs-hours/week” should be “600 MET-m/week”.

In summary, I find little value being added by this paper beyond the study by Brown et al. (2018). The authors should justify the validity of their PAR estimates for dietary calcium and red meat and address all my concerns above before the paper could be considered having reasonable quality.

References

1. Brown KF, Rumgay H, Dunlop C, Ryan M, et al.: The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015.*Br J Cancer*. **118** (8): 1130-1141 PubMed Abstract | Publisher Full Text
2. Fry A, Littlejohns TJ, Sudlow C, Doherty N, et al.: Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population.*Am J Epidemiol*. 2017; **186** (9): 1026-1034 PubMed Abstract | Publisher Full Text
3. Kim H, Wang K, Song M, Giovannucci EL: A comparison of methods in estimating population attributable risk for colorectal cancer in the United States.*Int J Cancer*. 2021; **148** (12): 2947-2953 PubMed Abstract | Publisher Full Text
4. Steenland K, Armstrong B: An overview of methods for calculating the burden of disease due to specific risk factors.*Epidemiology*. 2006; **17** (5): 512-9 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
No

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Cancer epidemiology; Population attributable risk; Biostatistics; Molecular epidemiology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Author Response 02 Mar 2022

Shatabdi Goon, University of Illinois at Urbana-Champaign, Urbana, USA

The authors provided some PAR estimates for the dietary and lifestyle risk factors associated with CRC and BC in the UK along with its constituent countries. I couldn't help wondering about the rationale underlying this study and several methodologic issues after reading this manuscript.

1. The rationale of carrying out this study eludes me, especially a large chunk of the result was just a repetition of Brown et al. (Br J Cancer, 2018, 118(8): 1130–41). The PAR for individual and combined risk factors of CRC and BC can be found in the supplementary materials of Brown's paper, except for dietary calcium and red meat. Was the rationale of this study to update the PAR estimates in the UK by including these two risk factors, which were excluded in Brown's paper?

Author's response: The rationale was to conduct the PAR based on up-to-date criteria from the World Cancer Research Fund/American Institute of Cancer Research (WCRF/AICR). This other analysis by Brown et al did not account for red meat or dietary calcium, now considered as probable risk factors by WCRF/AICR, and included radiation and oral contraceptive use, which we did not consider. While technically modifiable, radiation on OC use is not considered "lifestyle" in the same way as diet. In addition, we showed results including and excluding smoking, which is an important risk factor but not in the WCRF/AICR purview of nutrition, physical activity, and body weight. While there are similarities with the Brown et al paper, our main goal was to calculate the PAR estimate for the most up-to-date data from the WCRF/AICR for of nutrition, physical activity, and body weight.

1. The authors obtained exposure distribution data on dietary calcium and red meat consumption from EPIC-Oxford cohort and UK biobank cohort (page 3). However, these two cohorts are not population representative of UK. For example, the EPIC-Oxford cohort clearly stated on their website that the sample is not representative of the general population, as 50% of their sample are vegetarians. (http://www.epic-oxford.org/faq/1262/frequently-asked-questions). Similarly, UK Biobank participants were more likely to be older, to be female, and to live in less socioeconomically deprived areas than nonparticipants (Fry A. et al., Am J Epidemiol. 2017, 186(9):1026-34[2]). Therefore, the PAR estimates of calcium and red meat are bound to be biased using the data.

Author's response: We agree that both EPIC and UK biobank is not population-representative like UK survey data. Due to prevalence data unavailability on red/proceed meat and calcium intake- we planned to use other large UK-based cohorts (i.e., UK biobank and EPIC). Diet in both UK biobank and EPIC is somewhat healthier side on average; we have probably slightly underestimated PAR.

1. I wonder why the authors did not obtain cancer incidence data from the cancer registries in the UK. Without the incidence data, there are at least three limitations regarding the PAR estimates from this study. First, the authors could not estimate the attributable cancer cases, which is a constituent part of PAR estimation. Second, the authors could not do an age-stratified analysis. The summation of attributable cases of the age groups, divided by the cancer incidence of all age groups, provides a more accurate all-age PAR estimate than directly using an all-age prevalence to estimate

PAR (the latter of which I suspect is what the authors did). This is because some risk factors are more prevalent among older age groups (e.g., physical inactivity, body fatness) and most cancer incidences are disproportionally higher among older age groups. Second, the authors had to assume a weighted average of 70% colon and 30% rectal cancers for physical activity and CRC (page 5), which wouldn't be necessary if they had the incidence data by colon and rectum.

Author's response: The relative risk estimates are based on the literature and not given in age stratification typically. The method relies on two parameters, the relative risk and the prevalence of the exposure. The proportion is what is calculated, not the absolute number of cancers estimated. Of course, the proportion can be multiplied by the incidence to get the absolute numbers. The 70-30 assumption was made based on the paper was based on the RR estimated for colon cancer is 0.84 and that for rectal cancer 0.87. We assumed that for colorectal cancer, the RR would be a weighted average of these, which is around 0.85 (i.e., if the RR for colon cancer is 0.84 and that for rectal cancer around 0.87, the RR for CRC would be around 0.85). Because these values are so close anyway, this assumption appears reasonable. The 70-30 prevalence was assumed based on published data. (Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2020).

1. The authors didn't correctly estimate the PAR of body fatness and physical activity for BC. These two risk factors are only associated with postmenopausal breast cancers, whereas the authors included premenopausal breast cancers as well.

Author's response: Incidence data do not distinguish pre- and post-menopausal breast cancer, so these cannot be separated directly. Physical activity is considered a (protective) risk factor by the WCRF/AICR for both pre- and post-menopausal breast cancer. The only factor that is different is body fatness, which is considered a risk factor only for postmenopausal women. Assuming that body fatness is only a risk factor for postmenopausal breast cancer, they account for about 75% of cases. Since one of the four factors is not relevant for 25% of the cases, our estimate may be overestimated by around 5%.

1. The way the authors calculated RRs for calcium and physical activity is difficult to understand. They used a formula that assumes a loglinear dose-response relationship and calculates RR based on per unit RR, which is a usual practice in this field. Then for dietary calcium, they calculated an average RR for 0–199 mg/day, 200–399 mg/day, and 400–599 mg/day from their IJC paper (Kim H. et al., Int J Cancer, 2021, 148(12):2947-53[3]) to match 0–524 mg/day of the lowest category of exposure in this study (page 5, first paragraph). It was unclear to me how they averaged the RR and how they resolved the issue that the highest category (400-599mg/day) exceeds the upper limit (524 mg/day). Why not just take the midpoint value of this category (262 mg/day) and combine it with the risk per unit value? For example, for a reference calcium level of 1000 mg/day and per mg/day unit RR of r_mg, the RR for the category of 0-524 mg/day is r_mg^(1000-262).

Author's response: Because of the exposure data unavailability that matches with the RR- we had to average RR to match with a specific category of intake. Exposure data wasn't specified in the cohort we used data from the way we categorized to match with RR.

Some other concerns:

1. The authors used a formula, exp(lnX/A *B) to calculate RR, while neither A nor B was

defined. The authors referenced their recent publication in IJC for this formula (Kim H. et al., Int J Cancer, 2021, 148(12):2947-53[3]). But in their IJC paper, the formula was laid out as "RR_B = exp(log(RR_B)/A *B)", which is simply wrong. "log(RR_B)" should be "log(RR_A)". The authors should write an erratum to IJC to have it corrected.

Author's response: This was a typo in the PMC version. Now it is fixed.

1. In the second paragraph of the Method section, the author stated, "the prevalence of exposures to risk factors in the populations was obtained from the nationally representative population surveys" (page 3). However, which risk factors exactly were not stated. This caused some confusion to me, as I tried to figure out from which data source they obtained the prevalence of processed meat consumption.

Author's response: Meat and calcium were not from the nationally representative survey, data sources were mentioned in the table as a footnote.

1. In Table 2, there is no variation in PAR by countries for red meat and dietary calcium (page 6). It is unclear whether there is indeed no country-level variation for these two risk factors, or there were no prevalence data to support this analysis, or there was a data-entry mistake. The authors didn't have any discussions around it. If there were no prevalence data to support this analysis, then the discussion around differences in PAR between countries should acknowledge this limitation.

Author's response: Prevalence data were only for the UK- not for country-specific.

1. At the end of the second paragraph of the introduction (page 3), the authors stated that other contributors to the variability in calculated PAR estimates by different studies differences include "socio-demographic profile of cancer cases and differences due to screening availability". As the authors didn't provide citations to support this statement, could the authors further elaborate on how these two factors contribute to the variability? In particular, it is difficult to fathom how screening availability, a factor related to the secondary intervention of CRC and BC, contributes to the PAR of diet and nutrition.

Author's response: Screening is typically associated with the socio-demographic profile, and thus often with the dietary and lifestyle factors that we assessed. The broad effect of colonoscopy is to lower colorectal cancer risk (by removing precursors) and mammography tends to lead to more cases diagnosed ("overdiagnosis"). How these affect our results is complex; for example, they may affect the initial RR estimates (confounding).

1. The authors used an equation that has been widely adopted for estimating combined risk factors (page 5, paragraph 4). However, many researchers used this formula without disclaiming the two important assumptions associated with it (Steenland and Armstrong, 2006, Epidemiology, 17(5):512-9[4]): independent risk exposure distributions and no statistical interactions between any two risks. These two assumptions should not only be explicitly stated in the method section but also be worthy of several sentences in the discussion section. Also on page 5 after the formula, the author said "where i signifies the level of individual categories (i = 1,..., n)." "Categories" here should be "risk factors". Again for this paragraph, the first sentence "We then estimated the preventability of CRC or BC that was attributable to the combined dietary and lifestyle risk factors the following equation" is missing a

preposition.

Author's response: We agree that these are important assumptions for the formula. It has been fixed in the revised one.

1. Page 5 line 10, "600 METs-hours/week" should be "600 MET-m/week". (fixed)
2. In summary, I find little value being added by this paper beyond the study by Brown et al. (2018). The authors should justify the validity of their PAR estimates for dietary calcium and red meat and address all my concerns above before the paper could be considered having reasonable quality.
   Author's response: Because nationally representative exposure data were not available for meat and calcium – we had to use a large UK-based cohort (Biobank and EPIC). This may have caused a slight underestimate in the PAR estimate.

Competing Interests: No competing interests were disclosed.