Lifestyle Factors for the Prevention of Inflammatory Bowel Disease

Emily W. Lopes[1,2], Simon S. M. Chan[3,4], Mingyang Song[1,2,5,6], Jonas F. Ludvigsson[7,8], Niclas Håkansson[9], Paul Lochhead[1], Allan Clark[10], Kristin E Burke[1], Ashwin N Ananthakrishnan[1,2], Amanda Cross[11,12], Domenico Palli, Manuela Bergmann[13], James M Richter[1], Andrew T. Chan[1,2,14], Ola Olen[15,16], Alicja Wolk[17, 18], Hamed Khalili[1,2,19]

Collaborators: EPIC-IBD investigators: Pilar Amian, Aurelio Barricarte, Marie-Christine Boutron-Ruault, Franck Carbonnel, Olof Grip, Marc J Gunter, Rudolf Kaaks, Tim Key, María Dolores Chirlaque López, Robert Luben, Giovanna Masala, Jonas Manjer, Bas Oldenburg, Anja Olsen, Kim Overvad, Elio Riboli, Maria José Sánchez, Carlotta Sacerdote, Anne Tjønneland, Rosario Tumino, Roel Vermeulen, W. M. Monique Verschuren, Nick Wareham

Author Affiliations:

1. Division of Gastroenterology, Massachusetts General Hospital, Boston MA, USA
2. Clinical and Translation Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston MA, USA
3. Department of Medicine, Norwich Medical School, Norwich, UK
4. Department of Gastroenterology, Norfolk and Norwich University Hospitals NHS Trust, Norwich, UK
5. Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, USA
6. Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA
7. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
8. Department of Pediatrics, Örebro University Hospital, Örebro, Sweden
9. Nutritional Epidemiology Unit, Institution of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
10. Norwich Medical School, University of East Anglia, Norwich, UK
11. Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom.
12. Cancer Screening & Prevention Research Group, Department of Surgery & Cancer, Imperial College London, London, United Kingdom.
13. German Institute of Human Nutrition Potsdam-Rehbrücke, Nuthetal, Germany
14. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
15. Clinical Epidemiology Division, Karolinska Institutet, Stockholm, Sweden
16. Pediatric Gastroenterology Unit, Sachs' Children and Youth Hospital, Stockholm, Sweden.
17. Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
18. Department of Surgical Sciences, Uppsala University, Uppsala, Sweden
19. Broad Institute of MIT and Harvard, Cambridge MA 02142

Grant Support
Funded by UM1 CA186107 NHS cohort infrastructure grant, U01 CA176726 NHSII cohort infrastructure grant, and U01 CA167552 HPFS cohort infrastructure grant; the content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work is also funded by the VR 2017-00644 SMC and CoSM cohorts Swedish research infrastructure (SIMPLER) grant. The coordination of EPIC is financially supported by International Agency for Research on Cancer (IARC) and also by the Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London which has additional infrastructure support provided by the NIHR Imperial Biomedical Research Centre (BRC). The national cohorts are supported by: Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), German Institute of Human Nutrition Potsdam-Rehbruecke (DIfE), Federal Ministry of Education and Research (BMBF) (Germany); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy, Compagnia di SanPaolo and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); Health Research Fund (FIS) - Instituto de Salud Carlos III (ISCIII), Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, and the Catalan Institute of Oncology - ICO (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C8221/A29017 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk; MR/M012190/1 to EPIC-Oxford) (United Kingdom). This study was also funded by a senior research award from the Crohn's and Colitis Foundation to HK and a senior research award from the Crohn's and Colitis Foundation to ATC. EL was funded from NIH T32 DK007191 during work on this manuscript. Funding sources did not participate in study design, analysis, interpretation, drafting of manuscript, or submission process.

Abbreviations: Body mass index (BMI), Crohn's disease (CD), Health Professionals Follow-up Study (HPFS), Inflammatory bowel disease (IBD), Nurses' Health Study (NHS), Nurses' Health Study II (NHSII), non-steroidal anti-inflammatory drugs (NSAIDs), semi-quantitative food frequency questionnaire (SFFQ), ulcerative colitis (UC)

Correspondence
Hamed Khalili, MD, MPH, Digestive Healthcare Center- Crohn's and Colitis Center, Massachusetts General Hospital, 165 Cambridge Street, 9th Floor, Boston, MA 02114. Phone: 617 726 7933 Fax: 617 726 3080; hkhalili@mgh.harvard.edu

Disclosures
All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: HK is supported by the American College of Gastroenterology Senior Research Award and the Beker Foundation; HK has received consulting fees from Abbvie and Takeda; HK has also received grant funding from Pfizer and Takeda; ATC is the Stuart and Suzanne MGH Research Scholar; JFL reports funding from Janssen corporation for work unrelated to this manuscript; OO has been PI on projects at Karolinska Institutet, partly financed by investigator-initiated grants from Janssen, Takeda, and Ferring, and Karolinska Institutet has received fees for lectures and participation on advisory boards from Janssen, Ferring, Takeda, and Pfizer; OO also reports a grant to Karolinska Institutet from Pfizer in the context of a national safety monitoring program; ATC has received consulting fees from Bayer Pharma AG, Pfizer

Inc., and Boehringer Ingelheim for work unrelated to this manuscript. SSMC has received travel grants from Abbvie and Takeda. There are no other relationships or activities that could appear to have influenced the submitted work.

Author Contributions

EWL and HK were involved in the study concept and design. EWL, SSMC, KEB, PL, ANA, MMB, JMR, ATC, HK participated in acquisition of data. EWL, SSMC, MS, NH, AC and HK were involved in statistical analysis. EWL and HK had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors participated in interpretation of data. EWL and HK performed drafting of the manuscript. All authors participated in critical revision of the manuscript.

Acknowledgement

We would like to thank the participants and staff of the Nurses' Health Study (NHS), NHSII, Health Professionals Follow-up Study, Swedish Mammography Cohort, Cohort of Swedish Men, and the European Prospective Investigation into Cancer and Nutrition (EPIC) for their valuable contributions.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Several modifiable lifestyle and dietary risk factors have been identified for Crohn's disease (CD) and ulcerative colitis (UC) and are widely thought to contribute to disease pathogenesis.

One approach to the prevention of chronic diseases is via modification of lifestyle and dietary factors.

However, the extent by which adherence to low-risk factors or a healthy lifestyle could decrease the burden of CD and UC is unknown.

WHAT THIS STUDY ADDS

In three prospective US cohorts, adherence to low-risk factors could have prevented 42.9% (95% CI 12.2% to 66.1%) of CD and 44.4% (95% CI 9.0% to 69.8%) of UC cases, while adherence to a healthy lifestyle could have prevented 61.1% (95% CI 16.8% to 84.9%) of CD and 42.2% (95% CI 1.7% to 70.9%) of UC cases. These findings were largely confirmed in three external European cohorts.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

Assuming a causal relationship exists, a substantial proportion of the burden of inflammatory bowel diseases (IBD) may be preventable through lifestyle modification. Lifestyle modification may be an attractive target for future prevention strategies in IBD.

ABSTRACT

Objective To estimate the proportion of cases of Crohn's disease (CD) and ulcerative colitis (UC) that could be prevented by modifiable lifestyle factors.

Design In a prospective cohort study of US adults from the Nurses' Health Study (NHS; n=72 290), NHSII (n=93 909) and Health Professionals Follow-up Study (HPFS; n=41 871), we created modifiable risk scores (MRS; 0–6) for CD and UC based on established lifestyle risk factors, and healthy lifestyle scores (HLS; 0–9) derived from American healthy lifestyle recommendations. We calculated the population attributable risk by comparing the incidence of CD and UC between low-risk (CD-MRS≤1, UC-MRS≤2, HLS≥7) and high-risk groups. We externally validated our findings in three European cohorts: the Swedish Mammography Cohort (n=37 275), Cohort of Swedish Men (n=40 810) and European Prospective Investigation into Cancer and Nutrition (n=404 144).

Results Over 5 117 021 person-years of follow-up (NHS, HPFS: 1986–2016; NHSII: 1991–2017), we documented 346 CD and 456 UC cases. Adherence to a low MRS could have prevented 42.9% (95% CI 12.2% to 66.1%) of CD and 44.4% (95% CI 9.0% to 69.8%) of UC cases. Similarly, adherence to a healthy lifestyle could have prevented 61.1% (95% CI 16.8% to 84.9%) of CD and 42.2% (95% CI 1.7% to 70.9%) of UC cases. In our validation cohorts, adherence to a low MRS and healthy lifestyle could have, respectively, prevented 43.9%–51.2% and 48.8%–60.4% of CD cases and 20.6%–27.8% and 46.8%–56.3% of UC cases.

Conclusions Across six US and European cohorts, a substantial burden of inflammatory bowel diseases risk may be preventable through lifestyle modification.

INTRODUCTION

Crohn's disease (CD) and ulcerative colitis (UC) are chronic inflammatory bowel diseases (IBD) that affect an estimated 3.1 million adults in the US[1] and another 1.3 million in Europe[2]. Globally, the incidence of IBD is increasing, particularly in newly industrialised countries[3]. IBD is associated with significant societal cost, with an estimated annual healthcare cost of US$23 000 per-person in the USA[4]. Thus, strategies to prevent IBD could substantially decrease morbidity associated with disease and healthcare costs. However, to date, no strategies exist to prevent the development of IBD

One approach to prevent chronic diseases is via modification of lifestyle risk factors. Indeed, previous observational studies have identified several lifestyle factors to be associated with IBD[5], but whether modification of these lifestyle factors could be an attractive prevention strategy is unknown. Thus, in this study, we created modifiable risk scores (MRS) based on established risk factors for IBD and estimated the proportion of IBD cases that could have been prevented using population attributable risk (PAR). As some of the established risk factors such as smoking and body mass index (BMI) have opposite associations with CD and UC,[6,7]we also estimated the proportion of cases that could be prevented by adhering to an overall healthy lifestyle, as recommended by the US Department of Health and Human Services (HHS), the US Department of Agriculture (USDA) and the American Heart Association (AHA).

METHODS

Study population

Our primary cohort included participants from three prospective cohorts: the Nurses' Health Study (NHS), NHSII and Health Professionals Follow-up Study (HPFS). Briefly, the NHS enrolled 121 700 female nurses (30–55 years) across 11 US states in 1976, while the NHSII cohort, established in 1989, followed a younger cohort of 116 429 female nurses (25–42 years) from 15 US states.[8]. The HPFS cohort enrolled 51,529 male physicians (ages 40-75 years) across all 50 states in 1986[9]. Participants completed baseline and biennial questionnaires that assessed lifestyle factors, anthropomorphic data, and medical

history. Dietary information was collected every four years via semi-quantitative food frequency questionnaires (SFFQ) beginning in 1986 for NHS and HPFS and 1991 for NHSII (defined as baseline). Follow-up rates for these cohorts have consistently exceeded 85%[8,9].

We excluded participants who had missing baseline SFFQ or implausible daily caloric intake (<600 or >3500 kcal/day for women, <800 or >4200 kcal/day for men; n=67,671 (23%)), those who only completed baseline questionnaire (n=8,177 (2.8%)), those with a diagnosis of IBD at baseline (n=144 (0.05%)), and missing or implausible body mass index (BMI < 10 kg/m^2; n=1,468 (0.5%)).

We also used three large European cohorts to externally validate our results: the Swedish Mammography Cohort (SMC; n=37,275), the Cohort of Swedish Men (CoSM; n=40,810), and the European Prospective Investigation into Cancer and Nutrition (EPIC; n=404,144; **Supplementary Material**). Briefly, the SMC and CoSM are parallel cohorts of Swedish females (age 40-74 years) and males (age 45-79 years)[10], while the EPIC cohort is composed of both males and females (ages 35-70 years) across 10 European countries[11]. Detailed medical, lifestyle, and dietary information was collected at baseline (1997 for SMC and CoSM, 1992-1999 for EPIC) via via self-administered questionnaires in all cohorts (**Supplementary Material)**.

Patient and public involvement

Patients or the public were not involved in the design or interpretation of this study.

Ascertainment of IBD diagnosis

Ascertainment of IBD diagnoses in NHS, NHSII and HPFS has been previously described in detail[12]. Briefly, participants first self-reported diagnoses of either CD or UC in biennial questionnaires. Supplementary questionnaires were then mailed to participants requesting detailed information on IBD diagnoses and permission to review medical records. Records were then reviewed by two gastroenterologists blinded to exposure information. IBD cases were confirmed based on endoscopic and

histopathology findings. IBD cases in the validation cohorts were ascertained either through medical record review or using validated definitions used in patient registers (**Supplementary Material).**

Assessment of lifestyle risk factors and other covariates

riefly, non-dietary factors including BMI, family history of IBD, history of appendectomy (self-reported), physical activity, smoking status and non-steroidal anti-inflammatory drug (NSAID) use were assessed from baseline and follow-up questionnaires. Dietary factors, including daily servings of fruit and vegetables and red meat, fibre intake in grams (g), and ratio of n3:n6 polyunsaturated fatty acid (PUFA) intake were ascertained using frequency of intake reported on every 4-year SFFQ and the Harvard Food Composition Database to calculate nutrient-level data.[13]. BMI, smoking status and NSAID use were updated every 2 years, while physical activity and dietary variables were cumulative averaged over the follow-up time to better represent long-term patterns[14]. All variables except family history of IBD were modeled as time-varying covariates. In external cohorts, covariates were ascertained at baseline only. Further details for the variables assessed in the primary and external cohorts are described in the **supplementary material**.

Statistical Analysis

We constructed MRS for each of CD and UC (CD-MRS and UC-MRS) based on established modifiable risk factors, including BMI,[6 15] smoking status,[7] NSAID use,[16] physical activity[17] and daily consumption of fruit and vegetables,[18] fibre,[19 20] n3:n6 PUFAs[21] and red meat.[22 23] The directed acyclic graph for the proposed relationship between risk factors and outcomes is shown in **supplemental figure S1** (created using DAGitty V.3.0).[24] We defined low-risk criteria for each factor based on observed associations from prior literature, some of which had opposite relationships with CD and UC (table 1).[5] For example, never-smoking and non-obese BMI were considered low risk for CD, while current-smoking and obese BMI were considered low risk for UC.[6 7 15] For each participant, we assigned 1 point to each factor not meeting its low-risk criterion (0 otherwise) and summed each category for a total MRS of 0–6 points, so that

higher scores reflected a greater number of disease-specific risk factors. The low-risk group (reference) was defined as a score 0–1 or 0–2 when there were too few cases of CD and UC (defined by non-convergence of the models) in the 0–1 group.

Additionally, we note that adherence to low-risk factors did not necessarily represent healthy habits, particularly for UC, where current smoking and obese BMI are protective. Thus, we additionally constructed healthy lifestyle scores (HLS), to assess adherence to healthy lifestyle recommendations by the US HHS and USDA Dietary Guidelines for Americans and the AHA Guide- lines for Healthy Living[25–27] (supplemental appendix). Healthy criteria were defined as BMI ≥18.5 to <25 kg/m2; never smoking; physical activity ≥7.5 metabolic equivalent of task

RESULTS

In our primary cohort, a total of 208 070 participants were included after exclusions (NHS: n=72 290, NHSII: n=93 909 and HPFS: n=41 871). During 5 117 021 person-years of follow-up, we ascertained 346 CD and 456 UC cases, with an incidence rate of 7 cases of CD and 9 cases of UC per 100 000 person-years. Baseline characteristics for the pooled primary cohort are shown in **Supplementary Table 2**.

Compared with participants with a CD-MRS of 0–1, the aHR (95% CI) of those with a CD-MRS of 6 was 4.15 (1.95 to 8.84; figure 1). Similarly, compared with those with a UC-MRS of 0–2, the aHR (95% CI) of those with a UC-MRS of 6 was 2.78 (1.47 to 5.25). Risk of CD and UC increased with each one-point increase in CD-MRS (p_{trend}<0.0001) and UC-MRS (p_{trend}=0.008), respectively. Our findings were similar for both women and men (all $p_{interaction}$>0.19; **supplemental table S3**). When using binary scores, those with a CD-MRS ≥2 had an aHR (95% CI) of 1.85 (1.12 to 3.06; p=0.02) for CD when compared with those with a score of 0–1. Similarly, those with a UC-MRS≥3 had an aHR (95% CI) of 1.92 (1.08 to 3.40; p=0.03) for UC when compared with those with a score of 0–2.

We estimated that adherence to low CD-MRS (0–1) and UC-MRS (0–2) could have prevented 42.9% (12.2%–66.1%) of CD and 44.4% (9.0%–69.8%) of UC, respectively (PAR; **figure 1**). These findings were similar when incorporating processed meat intake in derivation of CD-MRS and UC-MRS

(supplemental appendix). In a sensitivity analysis using weighted criteria to define MRS, adherence to low CD-MRS and UC-MRS (lowest 15% of scores) could have prevented 41.0% (17.5%–60.0%) of CD and 27.7% (7.5%–45.7%) of UC (**supplemental table S4**).

Falsification analysis yielded anticipated results (**supplemental appendix**). Adherence to low CD-MRS (0–1) could have prevented 32.3% (0.4%–58.3%) of RA, 13.3% (2.3%–23.9%) of CRC and 14.0% (9.6%–18.5%) of CVD. Conversely, adherence to low UC-MRS (0–2) was associated with higher risk of RA, CRC and CVD compared with the UC-MRS >2 group, and therefore PAR for adherence to a low UC-MRS could not be calculated. In other words, adherence to low UC-MRS could not prevent RA, CRC or CVD in our cohorts (**supplemental appendix**).

We also confirmed our primary findings using baseline data. In the pooled NHS, NHSII and HPFS cohort, baseline CD-MRS and UC-MRS remained significantly associated with increased risk of CD and UC, respectively (both $p_{trend} \leq 0.003$; **figures 2 and 3**). For CD, adherence to low baseline CD-MRS (0–1) could have prevented 36.5% (5.3%–61.3%) of CD, while adherence to low baseline UC-MRS (0–2) could have prevented 35.9% (11.2%–56.5%) of UC.

Our findings were similar in the external cohorts. For CD, low baseline CD-MRS (0–1) could have prevented 43.9% (−7.4% to 76.8%) and 51.2% (0.01% to 80.9%) of CD in the pooled SMC and CoSM cohort and EPIC, respectively. Similarly, for UC, low baseline UC-MRS (0–2) could have prevented 20.6% (−14.5% to 51.0%) and 27.8% (0.001% to 51.6%) of UC in the pooled SMC and CoSM cohort and EPIC, respectively.

We also calculated the proportion of IBD cases that could have been prevented by adherence to American healthy lifestyle guidelines. In the pooled NHS, NHSII and HPFS cohort, baseline HLS was associated with decreased risk of CD and UC ($p_{trend} \leq 0.004$ and 0.02, respectively; **table 2**). Adherence to a healthy lifestyle (HLS 7–9) could have prevented 61.1% (16.8%–84.9%) of CD and 42.2% (1.7%–70.9%) of UC cases. These results were consistent in external cohorts (**table 2**). Adherence to a healthy lifestyle could have prevented 48.8% (−37.4% to 89.8%) and 60.4% (4.1% to 87.6%) of CD in the pooled

SMC and CoSM cohort and EPIC, respectively, and 56.3% (1.3%–85.1%) and 46.8% (9.7%–72.5%) of UC in the pooled SMC and CoSM cohort and EPIC, respectively.

Additionally, we explored the contribution of individual lifestyle factors and risk of CD and UC in our primary cohorts (**supplemental tables S5 and S6**). Low fibre intake conferred the largest PAR for CD (27.9%) followed by past or current smoking (14.4%) and low physical activity (12.9%; **supplemental table S7**). Low fruit and vegetable intake contributed the largest PAR for UC (20.1%), followed by past smoking (18.0%) and low n3:n6 PUFA (11.0%; **supplemental table S7**). In comparison, family history of IBD conferred a PAR of 12.2% (8.0%–16.2%) for CD and 8.8% (5.4%–12.1%) for UC.

Finally, we used the E value method to assess for residual confounding in the relationship between binary MRS scores and IBD used in our primary PAR analysis (**supplemental table S8**). To explain an aHR of 1.85 for CD with a CD-MRS ≥2, an unmeasured confounder would need to have a risk ratio of 3.10 with each of the CD-MRS exposure and CD outcome, after controlling for the measured confounders, to fully explain away the observed relationship. Similarly, the observed aHR of 1.92 for UC with a UC-MRS ≥3 would need to be explained by an unmeasured confounder that was associated with a 3.25-fold risk with each of the UC-MRS exposure and UC outcome, after controlling for measured confounders. Weaker confounding could otherwise not explain away the observed relationships.[33]

DISCUSSION

In three large prospective US cohorts, we demonstrate that modifiable lifestyle factors could substantially decrease the burden of IBD. We found that 43% of CD and 44% of UC cases could have been prevented by adhering to low-risk modifiable lifestyle factors, assuming a causal relationship exists. Moreover, adherence to American healthy lifestyle recommendations could have prevented 61% of CD cases and 42% of UC cases. These findings were consistent across three European cohorts. In comparison, in our primary cohorts, family history of IBD had a modest PAR of 12% for CD and 9% for UC.

Few studies have examined the total contribution of lifestyle factors on IBD development at a population level. In an Italian cohort, smoking, oral contraceptive use and lack of breast feeding accounted for roughly 30% of the attributable risk for IBD,[34] while Brant et al estimated that current tobacco use conferred a 47% attributable risk for CD.[35] To our knowledge, our study represents the first to comprehensively assess the contribution of modifiable lifestyle and dietary factors to the risk of CD and UC. Nonetheless, our estimates are similar to those published for other immune-mediated diseases. For example, in two prior studies, modification of lifestyle risk factors could have prevented 41% of RA[36] and 48% of psoriasis cases.[37] Further, similar to our study, family history had only a modest contribution to risk of RA and psoriasis (~20% each).

Importantly, our data suggest that adherence to an overall healthy lifestyle may prevent a modest proportion of cases of CD and UC. While unhealthy factors such as obesity and smoking have been inversely associated with risk of UC,[6 38] we saw that their contribution was outweighed by the total effect of healthy living. That is, more UC cases could have been prevented by adherence to a healthy lifestyle (42%–56%) as compared with adherence to 'traditional' UC risk factors assessed by our UC-MRS scores (21%–44%). Thus, current guidelines for healthy living, which are primarily recommended to reduce CVD risk, may have additional benefits for prevention of other immune-mediated diseases such as IBD.

A key assumption of our findings is that the relationship between lifestyle factors and IBD development is causal. Though this has yet to be established, several lines of evidence support the critical role of environmental and lifestyle factors in development of IBD. First, in genome-wide association studies, genetic factors account for less than 15% of the total variance of IBD.[39] Similarly, in monozygotic twins, concordance for disease is estimated to be around only 15% for UC and 30% for CD.[40 41] Second, the high incidence of IBD in industrialised societies and sharp rise of IBD in developing countries also suggest that Westernisation of diet and environment influences disease development.[3] Further, in immigrants who move from low-incidence to high-incidence countries, risk for IBD is higher in second-generation than first-generation immigrants.[42] Finally, the dietary and lifestyle factors considered here have also been linked with systemic inflammation, microbial dysbiosis and gut

permeability, providing mechanistic plausibility for a causal relationship.[43–46] Thus, although family history of IBD was the single strongest risk factor for IBD in our cohorts (aHR (95% CI)=4.53 (3.38 to 6.07) for CD and 3.24 (2.45 to 4.29) for UC), the collective impact of environmental factors on IBD development is likely greater.

Though there are currently no known disease prevention strategies for CD and UC, dietary and lifestyle modifications may change the immunologic and microbiologic milieu necessary for disease development and therefore could serve as a strategy for IBD prevention. This may be of particular relevance to high-risk groups, such as first-degree relatives of IBD patients, who have an estimated 2%–17% risk of developing the disease over their lifetime.[47] Similar strategies have been applied in other immune-mediated diseases, including type I diabetes[48] and in unaffected first-degree relatives of those with RA.[49]

Our study has several strengths. Exposure data were collected prospectively, minimising the risk of recall or selection bias. Diet and physical activity variables were cumulatively averaged to account for long-term patterns. We used validated methods to assess lifestyle factors across all cohorts,[50 51] and updated them over time to minimise exposure misclassification. Compared with prior studies, we considered a comprehensive list of modifiable lifestyle factors in the quantification of PAR, and avoided use of non-modifiable factors, preclinical markers of disease and surrogates for proximal disease exposures in our MRS.[52] Further, falsification analysis demonstrated that our scores are relatively specific for IBD. For example, though the associations and PAR estimates were similar for RA, a chronic immune-mediated disease with shared risk factors for CD, the corresponding PARs, and therefore preventable cases, were lower in CRC and CVD in spite of similar direction of association. This is largely due to differences in strength of associations and prevalence of risk factors, and presence of other modifiable risk factors such as alcohol and medications or supplements which are strongly associated with these other conditions.[29] 53 We also note that the follow-up period in our cohorts coincided with a significant rise in the incidence of IBD in the Western countries, allowing us to examine relevant secular changes in lifestyle and dietary behaviours.[3] In our primary cohorts and EPIC, cases of IBD were

confirmed through blinded, medical record review by two gastroenterologists, minimising outcome misclassification bias. Additionally, while several PAR values had wide CIs, potentially due to a limited number of cases or a high SE introduced by a broad exposure definition,[52] the large majority did not cross 0%, increasing confidence in potential importance of dietary and lifestyle modifications in preventing CD and UC. Finally, our findings were largely reproducible in three European prospective cohorts, confirming external validity.

We also acknowledge several limitations. Mean age of IBD diagnosis (~45 years) for our cohort was higher than the typical age of onset of IBD, thus younger onset disease may be under-represented. Given the stronger genetic association with early-onset disease,[54] our PAR figures may overestimate the potential for lifestyle modifications in preventing early-onset IBD. Nonetheless, this finding may remain relevant for older-onset disease, which may be driven more heavily by environmental and lifestyle factors. Early lifestyle factors such as antibiotic exposure and breast feeding, which have not been associated with IBD risk in these cohorts,[55] environmental factors including pollution and socioeconomic factors were also not considered as these may not be readily modifiable.[52] We also acknowledge that we did not have information on several other potentially modifiable factors such as stress in our cohorts. Thus, residual confounding may exist and affect the validity of PAR estimates if all confounders are not modelled.[28 52] However, as most observed relationships between environmental and lifestyle factors and risk of IBD rarely exceed relative risk ratios of 3.00,[5] we feel the E value analysis for residual confounding builds confidence in the validity of our results. We note that longitudinal data were not available for all cohorts thus time-varying exposures could not be used in our external cohorts. PAR is also affected by exposure prevalence, which may differ across non-Western countries, and therefore generalisability may be limited. Finally, because of our limited sample size, we could not independently explore the contribution of modifiable lifestyle factors to risk of IBD in high-risk individuals, defined as those with a first degree relative with IBD.

CONCLUSION:

Across six US and European cohorts, we confirmed that a substantial proportion of CD and UC risk may be preventable through modification of lifestyle risk factors or adherence to a healthy lifestyle. Further prospective interventional studies are needed to determine whether lifestyle modification is effective for the primary prevention of IBD, particularly in high-risk population and younger-onset disease.

References:

1. Dahlhamer JM, Zammitti EP, Ward BW, et al. Prevalence of inflammatory bowel disease among adults aged ≥18 years — United States, 2015. MMWR Morb Mortal Wkly Rep 2016;65:1166–9.

2 Zhao M, Gönczi L, Lakatos PL, et al. The burden of inflammatory bowel disease in Europe in 2020. J Crohns Colitis 2021;15:1573–87.

3 Ng SC, Shi HY, Hamidi N, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. Lancet 2017;390:2769–78.

4 Park KT, Ehrlich OG, Allen JI, et al. The Cost of Inflammatory Bowel Disease: An Initiative From the Crohn's & Colitis Foundation. Inflamm Bowel Dis 2020;26:1–10.

5 Piovani D, Danese S, Peyrin-Biroulet L, et al. Environmental risk factors for inflammatory bowel diseases: an umbrella review of meta-analyses. Gastroenterology 2019;157:647–59.

6 Mendall MA, Jensen CB, Sørensen TIA, et al. Body mass index in young men and risk of inflammatory bowel disease through adult life: a population-based Danish cohort study. Sci Rep 2019;9.

7 Mahid SS, Minor KS, Soto RE, et al. Smoking and inflammatory bowel disease: a meta-analysis. Mayo Clin Proc 2006;81:1462–71.

8 History | Nurses' Health Study. Available: https://www.nurseshealthstudy.org/aboutnhs/history [Accessed March 2, 2020].

9 Health Professionals Follow-Up Study - About the Study. Available: https://sites.sph.harvard.edu/hpfs/about-the-study/ [Accessed March 2, 2020].

10 Harris H, Håkansson N, Olofsson C, et al. The Swedish mammography cohort and the cohort of Swedish men: study design and characteristics of 2 population-based longitudinal cohorts. OA Epidemiology 2013;1:16.

11 Riboli E, Hunt KJ, Slimani N, et al. European prospective investigation into cancer and nutrition (EPIC): study populations and data collection. Public Health Nutr 2002;5:1113–24.

12 Khalili H, Huang ES, Ananthakrishnan AN, et al. Geographical variation and incidence of inflammatory bowel disease among US women. Gut 2012;61:1686–92.

13 Harvard T.H. Chan School of Public Health Nutrition Department. Nutrient tables.

Available: https://regepi.bwh.harvard.edu/health/ [Accessed March 2, 2020].

14 Willett WC. Nutritional epidemiology. Oxford University Press, 1998.

15 Chan SSM, Chen Y, Casey K, et al. Obesity is associated with increased risk of Crohn's disease, but not ulcerative colitis: a pooled analysis of five prospective cohort studies. Clin Gastroenterol Hepatol 2022;20:1048–58.

16 Ananthakrishnan AN, Higuchi LM, Huang ES, et al. Aspirin, nonsteroidal anti-inflammatory drug use, and risk for Crohn disease and ulcerative colitis: a cohort study. Ann Intern Med 2012;156:350–9.

17 Wang Q, Xu K-Q, Qin X-R, et al. Association between physical activity and inflammatory bowel disease risk: a meta-analysis. Dig Liver Dis 2016;48:1425–31.

18 Li F, Liu X, Wang W, et al. Consumption of vegetables and fruit and the risk of inflammatory bowel disease. Eur J Gastroenterol Hepatol 2015;27:623–30.

19 Zeng L, Hu S, Chen P, et al. Macronutrient intake and risk of Crohn's disease: Systematic review and dose–response meta-analysis of epidemiological studies. Nutrients 2017;9:500.

20 Ananthakrishnan AN, Khalili H, Konijeti GG, et al. A prospective study of long-term intake of dietary fiber and risk of Crohn's disease and ulcerative colitis. Gastroenterology 2013;145:970–7.

21. Ananthakrishnan AN, Khalili H, Song M, et al. Genetic Polymorphisms in Fatty Acid Metabolism Modify the Association between Dietary n3:n6 Intake and Risk of Ulcerative Colitis: A Prospective Cohort Study. *Inflamm Bowel Dis*. 2017;23(11):1898-1904.

22. Jantchou P, Morois S, Clavel-Chapelon F, Boutron-Ruault MC, Carbonnel F. Animal protein intake and risk of inflammatory bowel disease: The E3N prospective study. *Am J Gastroenterol*. 2010;105(10):2195-2201. doi:10.1038/ajg.2010.192

23. Ge J, Han TJ, Liu J, et al. Meat intake and risk of inflammatory bowel disease: A meta-analysis. *Turkish J Gastroenterol*. 2015;26(6):492-497. doi:10.5152/tjg.2015.0106

24 Textor J, van der Zander B, Gilthorpe MS, et al. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. Int J Epidemiol 2017;30:dyw341.

25 American Heart Association. Guidelines for healthy living. Available: https://www.heart.org/en/healthy-living [Accessed 21 Apr 2021].

26 U.S. Department of Agriculture US Department of Health and Human Services. Dietary guidelines for Americans, 2010. Available: www.dietaryguidelines.gov [Accessed 21 Apr 2021].

27 Piercy KL, Troiano RP, Ballard RM, et al. The physical activity guidelines for Americans. JAMA 2018;320:2020–8.

28. Spiegelman D, Hertzmark E, Wand HC. Point and interval estimates of partial population attributable risks in cohort studies: Examples and software. *Cancer Causes Control*. 2007;18(5):571-579. doi:10.1007/s10552-006-0090-y

29. 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;148:2947–53.

30 Narula N, Wong ECL, Dehghan M, et al. Association of ultra-processed food intake with risk of I nflammatory bowel disease: prospective cohort study. BMJ 2021;21:n1554.

31 Zaccardelli A, Friedlander HM, Ford JA, et al. Potential of lifestyle changes for reducing the risk of developing rheumatoid arthritis: is an ounce of prevention worth a pound of cure? Clin Ther 2019;41:1323–45.

32 Gioia C, Lucchino B, Tarsitano MG, et al. Dietary habits and nutrition in rheumatoid arthritis: can diet influence disease development and clinical manifestations? Nutrients 2020;12:1456.

33 VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the

E-Value. Ann Intern Med 2017;167:268–74.

34 Corrao G, Tragnone A, Caprilli R, et al. Risk of inflammatory bowel disease attributable to smoking, oral contraception and breastfeeding in Italy: a nationwide case-control study. Cooperative Investigators of the Italian group for the study of the colon and the rectum (GISC). Int J Epidemiol 1998;27:397–404.

35 Brant SR, Wang M-H, Rawsthorne P, et al. A population-based case-control study of CARD15 and other risk factors in Crohn's disease and ulcerative colitis. Am J Gastroenterol 2007;102:313–23.

36 Sparks JA, Chen C-Y, Hiraki LT, et al. Contributions of familial rheumatoid arthritis or lupus and environmental factors to risk of rheumatoid arthritis in women: a prospective cohort study. Arthritis Care Res 2014;66:1438–46.

37 Naldi L, Chatenoud L, Linder D, et al. Cigarette smoking, body mass index, and stressful life events as risk factors for psoriasis: results from an Italian case-control study. J Invest Dermatol 2005;125:61–7.

38 Higuchi LM, Khalili H, Chan AT, et al. A prospective study of cigarette smoking and the risk of inflammatory bowel disease in women. Am J Gastroenterol 2012;107:1399–406.

39 Jostins L, Ripke S, Weersma RK, et al. Host-Microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 2012;491:119–24.

40 Brant SR. Update on the heritability of inflammatory bowel disease: the importance of twin studies. Inflamm Bowel Dis 2011;17:1–5.

41 Pack M. Fishing for missing heritability in IBD. Nat Rev Gastroenterol Hepatol 2015;12:318–20.

42 Agrawal M, Burisch J, Colombel J-F, et al. Viewpoint: inflammatory bowel diseases among immigrants from low- to high-incidence countries: opportunities and considerations. J Crohns Colitis 2020;14:267–73.

43 Jenkins AP, Trew DR, Crump BJ, et al. Do non-steroidal anti-inflammatory drugs increase colonic permeability? Gut 1991;32:66–9.

44 Papoutsopoulou S, Satsangi J, Campbell BJ, et al. Review article: impact of cigarette smoking on intestinal inflammation-direct and indirect mechanisms. Aliment Pharmacol Ther 2020;51:1268–85.

45 Amre DK, D'Souza S, Morgan K, et al. Imbalances in dietary consumption of fatty acids, vegetables, and fruits are associated with risk for Crohn's disease in children. Am J Gastroenterol 2007;102:2016–25.

46 Wu GD, Chen J, Hoffmann C, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science 2011;334:105–8.

47 Santos MPC, Gomes C, Torres J. Familial and ethnic risk in inflammatory bowel disease. Ann Gastroenterol 2018;31:14–23.

48 Knip M, Virtanen SM, Seppä K, et al. Dietary intervention in infancy and later signs of beta-cell autoimmunity. N Engl J Med 2010;363:1900–8.

49 Sparks JA, Iversen MD, Yu Z, et al. Disclosure of personalized rheumatoid arthritis risk using genetics, biomarkers, and lifestyle factors to Motivate health behavior improvements: a randomized controlled trial. Arthritis Care Res 2018;70:823–33.

50 Yuan C, Spiegelman D, Rimm EB, et al. Validity of a dietary questionnaire assessed by comparison with multiple Weighed dietary records or 24-hour recalls. Am J Epidemiol 2017;185:570–84.

51 Rimm EB, Giovannucci EL, Stampfer MJ, et al. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. Am J Epidemiol 1992;135:1114–26.

52 Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. Am J Public Health 1998;88:15–19.

53 Chang M, Hahn RA, Teutsch SM, et al. Multiple risk factors and population attributable risk for ischemic heart disease mortality in the United States, 1971-1992. J Clin Epidemiol 2001;54:634–44.

54 Ananthakrishnan AN, Huang H, Nguyen DD, et al. Differential Effect of Genetic Burden on Disease Phenotypes in Crohn's Disease and Ulcerative Colitis: Analysis of a North American Cohort. Am J Gastroenterol 2014;109:395–400.

55 Khalili H, Ananthakrishnan AN, Higuchi LM, et al. Early life factors and risk of inflammatory bowel disease in adulthood. Inflamm Bowel Dis 2013;19:542–7.

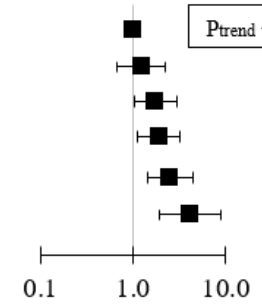
Table 1. Associations between modifiable risk factors and Crohn's disease or ulcerative colitis, and definitions for 'low-risk' criteria used in calculation of modifiable risk scores (MRS).

Crohn's disease		Ulcerative Colitis	
Risk Factor	**"Low-Risk" Criterion for MRS**	**Risk Factor**	**"Low-Risk" Criterion for MRS**
BMI[6,15]	< 30 kg/m^2	BMI[6]	≥ 30 kg/m^2
Smoking[7]	Never smokers	Smoking[7]	Current smokers
NSAIDs[16]	< 2x/week	NSAIDs[16]	< 2x/week
Physical activity[17]	Highest quintile (MET-hrs/wk)	Fruit & vegetables[18]	Highest quintile (servings/d)
Fruit & vegetables[18]	Highest quintile (servings/d)	Red meat[22,23]	Lowest quintile (servings/d)
Fibre[19,20]	Highest quintile (grams/d)	n3:n6 PUFA[21]	Highest quintile (servings/d)

BMI Body mass index. **MET** Metabolic equivalent of task. **NSAIDs** Non-steroidal anti-inflammatory drugs. **PUFA** Polyunsaturated fatty acid.

(a)

CD-MRS					
MRS Score	Person-Years	Cases	aHR (95% CI) [a]	P-value	PAR [b]
0-1 (ref)	424,183	16	1.00 (ref)		42.9% (12.2 - 66.1%)
2	772,547	35	1.24 (0.68-2.24)	0.48	
3	1,729,161	116	1.76 (1.04-2.97)	0.04	
4	1,586,583	115	1.92 (1.13-3.24)	0.02	
5	530,931	52	2.53 (1.44-4.45)	0.001	
6	73,616	12	4.15 (1.95-8.84)	0.0002	



(b)

UC-MRS					
MRS Score	Person-Years	Cases	aHR (95% CI) [a]	P-value	PAR [b]
0-2 (ref)	241,034	12	1.00 (ref)		44.4% (9.0 - 69.8%)
3	737,173	50	1.43 (0.76-2.70)	0.26	
4	1,694,989	158	1.97 (1.10-3.55)	0.02	
5	2,086,559	188	1.90 (1.06-3.42)	0.03	
6	357,266	48	2.78 (1.47-5.25)	0.002	

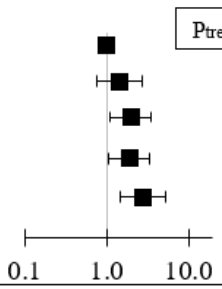
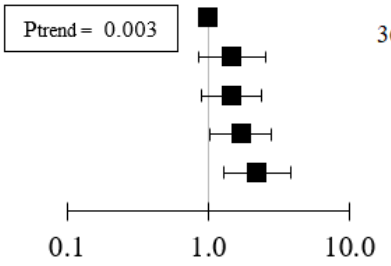


Figure 1. Risk and PAR of (a) Crohn's disease and (b) ulcerative colitis according to modifiable risk score (MRS). Reference level for UC set to 0-2 given low number of scores 0-1. [a] Cox models stratified by age (months), time-period (2-year intervals), and cohort (NHS, NHSII, or HPFS), and adjusted for appendectomy (yes/no), and family history of IBD (yes/no). [b] PAR for binary comparison between high-MRS (2+ risk factors for CD or 3+ risk factors for UC) and low-MRS (reference), adjusted for age (<40, 40 ≤ age <60, ≥ 60 years), cohort (NHS, NHSII, HPFS), appendectomy (yes/no) and family history of IBD (yes/no). **CD** Crohn's Disease. **CI** Confidence Interval. **aHR** Multivariable-adjusted Hazard Ratio. **PAR** Population Attributable Risk. **UC** Ulcerative Colitis.

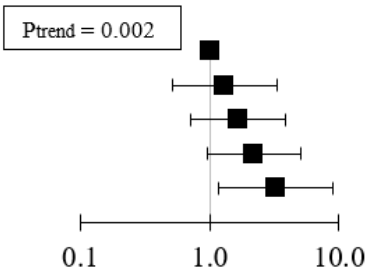
(a)

NHS/NHSII/HPFS				
CD-MRS Score [a]	No. Cases	Prevalence	aHR (95% CI) [b]	PAR [c]
0-1 (ref)	18	8.37%	1.00 (ref)	36.5% (5.3 - 61.3%)
2	52	16.14%	1.48 (0.86-2.53)	
3	116	36.23%	1.47 (0.89-2.41)	
4	114	30.31%	1.72 (1.04-2.82)	
5+	46	8.95%	2.25 (1.30-3.89)	



(b)

SMC/CoSM				
CD-MRS Score [a]	No. Cases	Prevalence	aHR (95% CI) [b]	PAR [c]
0-1 (ref)	6	7.30%	1.00 (ref)	43.9% (-7.4-76.8%)
2	19	17.54%	1.31 (0.52–3.28)	
3	53	38.56%	1.65 (0.71-3.85)	
4	62	32.96%	2.20 (0.95-5.11)	
5	10	3.66%	3.24 (1.17-8.94)	



(c)

EPIC				
CD-MRS Score [a]	No. Cases	Prevalence	aHR (95% CI) [b]	PAR [c]
0-1 (ref)	6	8.00%	1.00 (ref)	51.2% (0.01 – 80.9%)
2	21	17.65%	1.58 (0.64 – 3.91)	
3	47	36.87%	1.61 (0.69 – 3.77)	
4	68	33.16%	2.62 (1.14 – 6.04)	
5	6	4.32%	1.79 (0.57 – 5.55)	

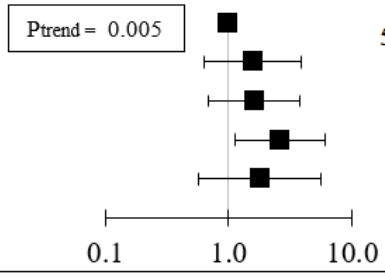
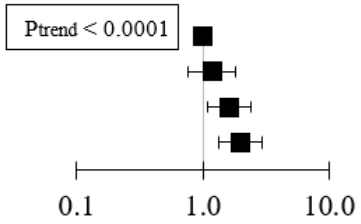


Figure 2. Risk and preventable fraction of Crohn's disease according to baseline modifiable risk score (MRS) for (a) pooled NHS/NHSII/HPFS cohort, (b) pooled SCM/CoSM cohort, and (c) EPIC cohort. [a]NSAID data missing from external cohorts, thus maximum MRS = 5. [b]Cox models adjusted for baseline age (years) and cohort. [c]PAR for 2+ risk factors compared to reference (0-1), adjusted for age (<40, 40 ≤ age <60, ≥ 60 years) and cohort. **CD** Crohn's Disease. **CI** Confidence Interval. **CoSM** Cohort of Swedish Men. **EPIC** European Prospective Investigation into Cancer and Nutrition. **HPFS** Health Professional's Follow-up Study. **aHR** Multivariable-adjusted Hazard Ratio. **MRS** Modifiable Risk Score. **NHS** Nurses' Health Study. **PAR** Population Attributable Risk. **SMC** Swedish Mammography Cohort.

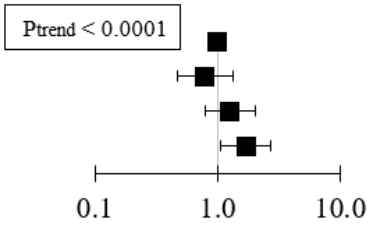
(a)

NHS/NHSII/HPFS				
UC-MRS Score [a]	No. Cases	Prevalence	aHR (95% CI) [b]	PAR [c]
0-2 (ref)	28	9.55%	1.00 (ref)	35.9% (11.2 - 56.5%)
3	75	21.67%	1.19 (0.77-1.83)	
4	185	39.57%	1.60 (1.08-2.39)	
5+	168	29.21%	1.98 (1.33-2.96)	



(b)

SMC/CoSM				
UC-MRS Score [a]	No. Cases	Prevalence	aHR (95% CI) [b]	PAR [c]
0-2 (ref)	20	8.04%	1.00 (ref)	20.6% (-14.5 - 51.0%)
3	50	23.83%	0.79 (0.47-1.33)	
4	152	41.99%	1.27 (0.80–2.03)	
5	137	26.14%	1.71 (1.06–2.75)	



(c)

EPIC				
UC-MRS Score [a]	No. Cases	Prevalence	aHR (95% CI) [b]	PAR [c]
0-2 (ref)	34	12.30%	1.00 (ref)	27.8% (0.001 – 51.6%)
3	95	30.93%	1.08 (0.73 – 1.60)	
4	175	43.34%	1.41 (0.98 – 2.04)	
5	74	13.27%	1.88 (1.25 – 2.82)	

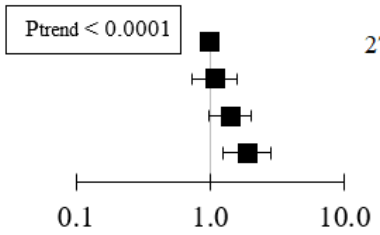


Figure 3. Risk and preventable fraction of ulcerative colitis according to baseline modifiable risk score (MRS) for (a) pooled NHS/NHSII/HPFS cohort, (b) pooled SCM/CoSM cohort, and (c) EPIC cohort. [a]NSAID data missing from external cohorts, thus maximum MRS = 5. UC-MRS adapted such that "low-risk" criterion for smoking defined as never-smoking only. Given the low incidence of UC-MRS 0-1, reference level set to 0-2. [b]Cox models adjusted for baseline age (years) and cohort. [c]PAR for 3+ risk factors compared to reference, adjusted for age (<40, 40 ≤ age <60, ≥ 60 years) and cohort. **CI** Confidence Interval. **CoSM** Cohort of Swedish Men. **EPIC** European Prospective Investigation into Cancer and Nutrition. **HPFS** Health Professional's Follow-up Study. **aHR** Multivariable-adjusted Hazard Ratio. **MRS** Modifiable Risk Score. **NHS** Nurses' Health Study. **PAR** Population Attributable Risk. **SMC** Swedish Mammography Cohort. **UC** Ulcerative colitis.

Table 2. Preventable fraction of CD and UC cases due to American healthy lifestyle recommendations

NHS/NHSII/HPFS				
Crohn's Disease				
Healthy lifestyle score [a]	**No. Cases**	**Prevalence**	**aHR (95% CI)** [b]	**PAR (95% CI)** [c]
7-9	5	4.24%	1.00 (ref)	61.1% (16.8 - 84.9%)
6	29	7.44%	3.13 (1.21-8.09)	
5	37	16.50%	1.67 (0.65-4.26)	
4	92	26.06%	2.58 (1.05-6.37)	
3	91	25.59%	2.57 (1.04-6.34)	
2	71	15.18%	3.42 (1.38-8.51)	
0-1	21	4.99%	3.01 (1.13-8.02)	
P-trend			0.004	
Ulcerative Colitis				
Healthy lifestyle score [a]	**No. Cases**	**Prevalence**	**aHR (95% CI)** [b]	**PAR (95% CI)** [c]
7-9	10	4.24%	1.00 (ref)	42.2% (1.7 - 70.9%)
6	30	7.44%	1.58 (0.77-3.23)	
5	71	16.50%	1.61 (0.83-3.13)	
4	109	26.06%	1.50 (0.78-2.87)	
3	132	25.59%	1.85 (0.97-3.53)	
2	72	15.18%	1.72 (0.88-3.34)	
0-1	32	4.99%	2.39 (1.17-4.87)	
P-trend			0.02	

SMC/CoSM				
Crohn's Disease				
Healthy lifestyle score [a]	**No. Cases**	**Prevalence**	**aHR (95% CI)** [b]	**PAR (95% CI)** [c]
7-9	2	2.62%	1.00 (ref)	48.8% (-37.4 – 89.8%)
6	16	10.19%	1.99 (0.46-8.66)	
5	34	23.32%	1.86 (0.45-7.74)	
4	36	30.88%	1.51 (0.36-6.27)	
3	39	22.31%	2.27 (0.55-9.41)	
2	19	8.79%	2.85 (0.66-12.23)	
0-1	4	1.89%	2.77 (0.51-15.14)	
P-trend			0.08	
Ulcerative Colitis				
Healthy lifestyle score [a]	**No. Cases**	**Prevalence**	**aHR (95% CI)** [b]	**PAR (95% CI)** [c]
7-9	4	2.62%	1.00 (ref)	56.3% (1.3 – 85.1%)
6	21	10.19%	1.27 (0.44-3.70)	
5	79	23.32%	2.10 (0.77-5.72)	
4	123	30.88%	2.52 (0.93-6.83)	
3	82	22.31%	2.35 (0.86-6.41)	
2	41	8.79%	3.03 (1.09-8.46)	
0-1	9	1.89%	3.05 (0.94-9.89)	
P-trend			0.0004	

EPIC				
Crohn's Disease				
Healthy lifestyle score [a]	**No. Cases**	**Prevalence**	**aHR (95% CI)** [b]	**PAR (95% CI)** [c]
7-9	4	6.50%	1.00 (ref)	60.4% (4.1 - 87.6%)
6	17	12.48%	2.13 (0.71 – 6.34)	
5	18	22.62%	1.21 (0.41 – 3.57)	
4	52	27.40%	2.82 (1.02 – 7.81)	
3	43	20.45%	3.09 (1.11 – 8.63)	
2	12	8.86%	1.97 (0.64 - 6.12)	
0-1	4	1.69%	3.43 (0.86 – 13.76)	
P-trend			0.005	
Ulcerative Colitis				

Healthy lifestyle score [a]	No. Cases	Prevalence	aHR (95% CI) [b]	PAR (95% CI) [c]
7-9	13	6.50%	1.00 (ref)	46.8% (9.7 - 72.5%)
6	29	12.48%	1.21 (0.62 – 2.38)	
5	65	22.62%	1.45 (0.78 – 2.68)	
4	123	27.40%	2.20 (1.22 – 3.98)	
3	103	20.45%	2.43 (1.34 – 4.42)	
2	44	8.86%	2.37 (1.25 – 4.48)	
0-1	7	1.69%	1.97 (0.78 – 4.99)	
P-trend			<0.001	

[a]Maximum value set to 7+ given few participants with scores of 8-9 in primary cohort. [b]Cox models adjusted for baseline age (years) and cohort (NHS, NHSII, HPFS, SMC, CoSM). [c]PAR for 3+ risk factors compared to reference category, adjusted for age (<40, 40 ≤ age <60, ≥ 60 years) and cohort. **CD** Crohn's Disease. **CI** Confidence Interval. **aHR** Multivariable-adjusted Hazard Ratio. **PAR** Population Attributable Risk. **UC** Ulcerative Colitis.

Supplementary Appendix

Table of Contents

Supplementary Methods

Validation Study Populations

The Swedish Mammography Cohort (SMC) was established in 1987-1990, when 66,651 Swedish females in Uppsala and Västmanland Counties were enrolled in a mammography screening program (74% of general female population in the study area in central Sweden) that included completion of an initial questionnaire assessing medical, lifestyle and dietary information[1,2]. An expanded survey was sent to participants in 1997 (70% response rate) that updated this information and also assessed physical activity and smoking habits. The 1997 questionnaire was also administered to men in Västmanland and Örebro Counties that same year, establishing the Cohort of Swedish Men (CoSM, n=45,906) with a response rate of approximately 49%[2,3].

The European Prospective Investigation into Cancer and Nutrition (EPIC) cohort was established in 1992-1999, when 521,324 participants were enrolled from 23 different centers across Spain, Italy, France, the United Kingdom, Greece, Germany, the Netherlands, Denmark, Sweden, and Norway[4,5]. Participants completed a combination of lifestyle and dietary/food frequency questionnaires, and anthropomorphic data was gathered either by participant report or direct measurement by trained staff. For our study, the Greek cohort was excluded (n=28,572) as were several centers/countries for which IBD information is not available: Norway (n=37,200), Asturias in Spain (n=8,542), and Naples and Milan in Italy (n=17,141). Finally, Umea was excluded as physical activity in METs was not available (n=25,725).

After exclusions, there were 37,275; 40,810; and 404,144 participants from each of the SMC, CoSM, and EPIC cohorts, respectively, with EPIC participants from Spain, Italy, France, United Kingdom, Germany, the Netherlands, Denmark, and Southern Sweden included in this study.

IBD Ascertainment

In our primary cohorts (NHS, NHSII, and HPFS), date of IBD diagnosis was used as date of index colonoscopy and histopathology results as determined during medical record review. We did not include participants who denied record review or whose diagnoses were unable to be confirmed (n=1209).

For SMC and CoSM, IBD was ascertained through linkage with the Swedish Patient Register and presence of at least two encounters that used an IBD-diagnosis code, through end of follow-up in 2017. The positive predictive value of this definition has been previously reported as 93% (95% CI: 87-97) for any IBD, 79% (66-88) for UC and 72% (60-82) for CD[6]. In the EPIC cohort, one or more of the following was used for disease confirmation through 2009: linkage with national or regional registries, follow-up questionnaires, and/or medical records reviewed by 1-2 physicians[7].

Rheumatoid Arthritis (RA), Colorectal Cancer (CRC), and Cardiovascular Disease (CVD) Ascertainment

RA cases were first self-reported on biennial questionnaires in NHS and NHSII cohorts. Participants who reported a diagnosis of RA were then sent a 6-question connective tissue disease screening questionnaire (CSQ) assessing symptoms of RA, as well as permission to request and review medical records. A positive CSQ screen was considered as 2 or more symptoms of connective tissue disease. We excluded those who denied a history of RA, were diagnosed with RA prior to baseline, denied permission for record review, or had a negative CSQ screen. Two rheumatologists blinded to exposure information then reviewed medical records. Cases were confirmed according to American College of Rheumatology criteria for RA and date of diagnosis was defined as date of first RA symptom.

Diagnoses of CRC were ascertained as in previously described methods[8]. Briefly, participants were queried on biennial questionnaires regarding new cancer diagnoses, and permission was requested to obtain and review medical records. Study physicians blinded to exposure information then reviewed records and pathology to confirm diagnoses. Additionally, for participants who died from colorectal cancer, next of kin, National Death Index, death certificates, and medical records were used to confirm the diagnosis in participants who had not reported their diagnosis on biennial questionnaires.

Finally, CVD was defined as a combined endpoint of coronary heart disease (CHD), coronary revascularization (including coronary bypass (CABG) or angioplasty), or stroke as has previously been described[9]. Briefly, participants self-reported a history of myocardial infarction (MI), angina, CABG, transient ischemic attack (TIA) or stroke on biennial questionnaires. Permission to obtain medical records was requested and reviewed by physicians. Non-fatal MI was diagnosed according to World Health Organization and the updated European Society of Cardiology and American College of Cardiology criteria[10,11]. CABG was self-reported. Non-fatal stroke was diagnosed according to Nation Survey of Stroke criteria[12]. Fatal CHD or stroke were identified by next of kin or by search of the National Death Index and confirmed on death certificates, medical records during hospitalization, or autopsy records.

Assessment of non-dietary factors

NHS, NHSII and HPFS cohorts

Non-dietary factors were ascertained from baseline and biennial questionnaires. Those with missing baseline dietary and lifestyle factors were excluded as described in the Methods. Missing data on follow up questionnaires were imputed using the last value from prior questionnaire. Body mass index (BMI) was calculated in kilograms of weight divided by meter squared baseline height, using updated weight every 2-years. Participants reported time spent in various physical activities per week, which was then converted to metabolic equivalent-hours (MET-hours) per week using expected METs per exercise[13–15], and cumulative-averaged to represent long-term patterns. Physical activity was categorized in quintiles of MET-hours/week. Smoking status was determined as previously described[16], using current and past smoking habits reported on baseline and subsequent questionnaires, and categorized as never, past, or current. Assessment of non-steroidal anti-inflammatory drug (NSAID) use has been previously reported[17]. Briefly, participants were asked about frequency and quantity of NSAID use on follow-up questionnaires. Regular NSAID use was defined as twice or more weekly use, and missing NSAID data was set to non-regular use. Family history of IBD in a first-degree relative was documented in 2012 in NHS and HPFS and in 2013 for NHSII. History of and date of appendectomy was asked in 1992 (NHS), 1995 (NHSII) and 1986 (HPFS). Family history of IBD and appendectomy were categorized as "yes/no," with missing data set to "no". Oral contraceptives were not included due to pooled cohorting of both males and females. Longitudinal antibiotic use was not available and was thus excluded from this analysis. For use in sensitivity analysis, family history of RA or systemic lupus erythematosus (SLE) in a first-degree relative was assessed in NHS in 2008, while in NHSII family history of RA in a first-degree relative was assessed in 2013. Family history of a first-degree relative with colorectal cancer was specifically asked and updated in NHS (2004, 2008), NHSII (2005, 2009) and HPFS (1996, 2008, 2012). Family history of a first degree relative with myocardial infarction and/or stroke were specifically queried in NHS (1996, 2008), NHSII (1993, 1997, 2001, 2005), and HPFS (1986).

Replication cohorts (SMC, CoSM, EPIC)

In validation studies, baseline data was used as longitudinal data was not uniformly available across all cohorts. In all cohorts, information on BMI and smoking were available at baseline. When using baseline

data, both past and current smoking status conferred increased risk for UC, potentially due to lack of a time-varying exposure. To account for this, UC-MRS was adapted such that the "low-risk" criterion for smoking was defined as never-smoking only. For physical activity, in SMC and CoSM, we included walking, cycling, or other exercise as physical activity[1,3,15]. In the EPIC cohort, physical activity (occupational, walking, cycling, gardening, housework, physical exercise, climbing stairs) was reported on standardized questionnaires for: France, Italy, Spain, the United Kingdom, the Netherlands, Greece, and Germany[4]. Denmark and Sweden, which joined the cohort later, developed independent scales of assessment that were then recoded to standardize physical activity to the original survey tool. For all cohorts, time spent in activities was multiplied by expected METs to achieve MET-hours per week, then modeled in quintiles[14]. Family history of IBD, history of appendectomy, and NSAID use were not uniformly available and thus were not included in derivation of lifestyle scores.

Assessment of dietary factors:

NHS, NHSII and HPFS cohorts

Participants were sent semi-quantitative food-frequency questionnaires (SFFQs) every 4 years and were asked to report food intake patterns over the previous year. The SFFQ assessed frequency of intake of standard-sized food items, ranging from never or less than once a month to 6 or more times per day. The Harvard Food Composition Database was used to calculate nutrient-level intake data from SFFQ data [18]. Derivation of nutrient intake data using SFFQ has previously been shown to correlate with 7-day dietary records and 24-hour dietary recalls [19,20]. To reduce the effect of extraneous variation in nutrient reporting, all nutrient values were adjusted to total energy intake using the residuals method[21,22]. We used cumulative average of daily servings to better represent long-term patterns of dietary intake[23], and modeled dietary variables in quintiles.

Red meat was defined as pork, beef or lamb products, including processed meat. As we chose to include processed red meat in our red meat definition, we did not include a separate score for degree of food processing in our risk scores to avoid overlap between variables.

Replication cohorts (SMC, CoSM, EPIC)

Dietary and nutrient data were assessed at baseline. For the SMC and CoSM cohorts, participants reported how frequently they consumed age-sized portions (servings) of common food items, ranging from never to 3 or more time per day. Nutrient values were then calculated based on the expected content of foods. The reproducibility and validity of this method for nutrient values has been previously established[2]. In the EPIC cohort, one of several methods was used to assess dietary intake across the various centers and countries, including quantitative dietary questionnaires (either self-reported or face to face interviews), SFFQs, and combined dietary methods (a combination of SFFQ and 7- or 14-day dietary recall)[4]. The use of SFFQ for this cohort has previously been validated against 24-hour recall questionnaires and nutrient biomarkers[24]. Ultimately, all servings and nutrient data from each of these methods was converted to grams per day to standardize reporting across these centers before central pooling of data. For use in our analysis, we converted grams per day to servings per day for fruit & vegetables and red meat intake using expected portion sizes reported in prior studies (one serving fruit and vegetables = 80 grams[25,26], one serving red meat = 90 grams [27,28]). All dietary factors were modeled in quintiles.

Statistical analysis

The Cox proportionality assumption was tested for by creating interaction terms between follow-up time and each of CD-MRS and UC-MRS and comparison the models with these interaction terms with the main models (without interaction terms) using the log likelihood ratio test. We observed no evidence for

an interaction between follow-up time and MRS for either CD (P=0.83) or UC (P=0.08), suggesting that the proportional hazards assumption was valid.

Sensitivity analyses

We constructed MRS using weighted criteria to account for known linear relationships between risk factors and IBD risk. For fruit and vegetables, physical activity, fiber, and n3:n6 PUFAs, lowest quintile of intake was assigned a maximum value of 5 points, followed by 2nd quintile: 4 points, 3rd quintile: 3 points, 2nd quintile: 2 points, and highest quintile of intake: 1 point. Conversely, red meat intake was scored 5 points for highest quintile intake, down to 1 point for lowest quintile. Regular NSAID use was given 5 points, with 1 point for non-use. For smoking status, never-smokers received 1 point, past-smokers 3 points, and current smokers 5 points for CD, while for UC, past smokers received 5 points, never smokers received 3 points, and current smokers received 1 point. For BMI, 1 point was assigned to BMI <30 kg/m^2 and 5 points to BMI ≥30 kg/m^2 for CD, and vice-versa for UC. Similar to primary analysis, weighted CD-MRS summed scores for smoking, NSAID use, BMI, physical activity, fruit & vegetable intake, and fiber intake, while weighted UC-MRS was totaled based on scores from smoking, NSAID use, BMI, fruit & vegetable intake, n3:n6 PUFA intake, and red meat intake (final MRS range 6-30 for each).

Healthy lifestyle scores

Healthy lifestyle criteria were created using standardized health recommendations from the US Department of Health and Human Services and US Department of Agriculture (USDA) Dietary Guidelines for Americans, and the American Heart Association (AHA) Guidelines for Healthy Living[29–31]. We defined healthy physical activity as ≥7.5 MET-hours per week based on recommendations for 150 minutes per week of moderate (MET=3) physical activity[31]. Never-smoking and 18.5 ≤ BMI < 25 kg/m^2 were considered healthy[29,30]. Healthy dietary patterns were chosen to reflect a Mediterranean style diet, which is supported by both the AHA and USDA guidelines. Though recommended intake for fruit and vegetables varies between the USDA and AHA guidelines, we chose to use the higher threshold of ≥ 8 servings/day recommended by the AHA given the overall higher intake of fruits and vegetables in our cohort. We chose a minimum of 25 grams/day of fiber intake as "healthy" based on the minimum amount of intake recommended for US females [32]. Using a standard serving size for meat of 3 oz[29] and a recommended maximum meat intake of 1.8 oz daily[32], we defined healthy red meat intake as <0.5 servings/day. Based on a recommendation for 8 oz fish per week and a standard serving size of 3.5-4 oz fish per serving, we defined healthy fish (including shellfish) intake as a minimum of 2 servings per week. Using a standard serving size (1 oz equivalents) of 0.5 oz for nuts and seeds and 1 tablespoon for nut butters[29,30], and a recommended minimum intake of 0.5 oz equivalents per day of nuts, seeds, and nut butters[32], we defined healthy nut and seed intake as a minimum of 0.5 servings/day. We did not consider beans/legumes as a separate category, as beans/peas/lentils as vegetables were included in our daily servings of vegetables assessed, and recommended intake of protein from beans/legumes are not explicitly defined in these guidelines. Finally, healthy alcohol use was defined as a maximum of 1 drink per day for women or 2 per day for men[29,30], with a standard serving size defined as 14 g alcohol content per drink, or 12 fluid oz of beer (1 glass, can, or bottle), 4-5 fluid oz wine, or 1.5 fluid oz (1 shot or drink) distilled spirits[30]. Missing dietary data was set to baseline cohort-specific median. Baseline data were used to be consistent across primary and validation cohorts. A reference group of 7-9 was used for healthy lifestyle scores and scores of 0-1 were grouped due to a small percentage of scores equal to 0.

Assessing residual confounding using the E-value

To assess for residual confounding, we applied the E-value method to our primary analysis to estimate the minimum strength of association that an unmeasured confounder would need to have on both the exposure and outcome to fully explain the observed aHR (95% CIs). As our outcome was relatively rare, we employed the following formula as described by VanderWeele et al[33]:

E-value for HR = HR + sqrt{HR×(HR− 1)}
E-value for LL (for a HR > 1) = LL+ sqrt{LL×(LL− 1)}

Where HR = the adjusted hazard ratio for CD- and UC-MRS with CD and UC, respectively, conditional on age, family history of IBD, and history of appendectomy; and LL = the lower limit of the confidence interval for these aHRs. We applied this to the binary CD- and UC-MRS used in the primary PAR analysis, as well as to individual levels of the CD- and UC-MRS.

Supplementary Results

Inclusion of a term for processed meat in MRS

We included a term for processed meat in the construction of CD-MRS (range 0-7) and replaced the term for red meat with processed meat when calculating UC-MRS (range 0-6), similar to our primary analysis. Compared to those with a CD-MRS of 0-1, those with a CD-MRS of 2, 3, 4, 5, 6, and 7 had an aHR (95% CI) of: 1.03 (0.46-2.29), 1.18 (0.57-2.42), 1.59 (0.80-3.15), 1.86 (0.94-3.68), 2.12 (1.03-4.35) and 4.04 (1.69-9.67) for CD, respectively (P_{trend}<0.0001). Adherence to low CD-MRS (0-1) could have prevented 37.7% (-7.2-69.9%) cases of CD (PAR). Similarly, compared to those with a UC-MRS of 0-2, those with a UC-MRS of 3, 4, 5, and 6 had an aHR (95% CI) of: 1.60 (0.84-3.07), 2.14 (1.16-3.94), 2.11 (1.15-3.89), and 2.92 (1.50-5.65) for UC, respectively (P_{trend}=0.0007). Adherence to low UC-MRS (0-2) could have prevented 49.0% (14.4-73.0%) of UC cases (PAR).

Falsification Analyses

Compared to those with a CD-MRS of 0-1, the aHR (95%CI) for RA of those with a CD-MRS of 2, 3, 4, 5, and 6 was 1.09 (0.82-0.45), 1.02 (0.79-1.33), 1.24 (0.96-1.60), 1.59 (1.20-2.11), and 1.77 (1.13-2.77), respectively (P_{trend} <0.0001). Compared to those with a UC-MRS of 0-2, the aHR (95%CI) for RA of those with a UC-MRS of 3, 4, 5, and 6 was 0.80 (0.58-1.80), 0.88 (0.67-1.17), 0.91 (0.69-1.20), and 1.15 (0.82-1.60), respectively (P_{trend}=0.12). Adherence to low CD-MRS (0-1) could have prevented 32.3% (0.4-58.3%) of RA. Conversely, because a UC-MRS > 2 (when compared to a reference of 0-2) was associated with an aHR < 1, PAR for adherence to low UC-MRS (0-2) could not be calculated. In other words, adherence to low UC-MRS could not prevent RA in our cohorts.

Compared to those with a CD-MRS of 0-1, the aHR (95%CI) for CRC of those with a CD-MRS of 2, 3, 4, 5, and 6 was 1.15 (0.98-1.35), 1.09 (0.94-1.26), 1.36 (1.18-1.57), 1.44 (1.22-1.71), and 1.47 (1.06-2.03), respectively (P_{trend} = <0.0001). Adherence to low CD-MRS (0-1) could have prevented 13.3% (2.3-23.9%) of CRC. Compared to those with a UC-MRS of 0-2, the aHR (95%CI) for CRC of those with a UC-MRS of 3, 4, 5, and 6 was 1.11 (0.92-0.34), 1.12 (0.93-1.33), 1.12 (0.94-1.33), and 0.89 (0.71-1.12), respectively (P_{trend} = 0.60). Adherence to low UC-MRS (0-2) was associated with higher risk for CRC compared to the UC-MRS > 2 group, and therefore PAR for adherence to low UC-MRS could not be calculated. In other words, adherence to low UC-MRS could not prevent CRC in our cohorts.

In the CVD analysis, compared to those with a CD-MRS of 0-1, the aHR (95%CI) for CVD of those with a CD-MRS of 2, 3, 4, 5, and 6 was 1.05 (0.98-0.13), 1.10 (1.03-1.17), 1.25 (1.18-1.33), 1.54 (1.43-1.65),

and 1.68 (1.47-1.92), respectively (P_{trend} = < 0.0001). Adherence to low CD-MRS (0-1) could have prevented 14.0% (9.6-18.5%) of CVD. Compared to those with a UC-MRS of 0-2, the aHR (95%CI) for CVD of those with a UC-MRS of 3, 4, 5, and 6 was 0.93 (0.86-1.00), 0.88 (0.82-0.95), 0.79 (0.74-0.85), and 0.80 (0.73-0.87), respectively (P_{trend} = < 0.0001). Adherence to low UC-MRS (0-2) was associated with higher risk for CVD compared to the UC-MRS > 2 group, and therefore PAR for adherence to low UC-MRS could not be calculated. In other words, adherence to low UC-MRS could not prevent CVD in our cohorts.

Table S1. Definitions for healthy criteria used in construction of Healthy Lifestyle Scores (HLS).

Modifiable Risk Factor	Healthy Criterion
BMI	18.5 $kg/m^2 \leq$ BMI < 25 kg/m^2
Smoking status	Never smoking
Physical activity	$\geq$ 7.5 MET-hours/week
Fruit & vegetables	$\geq$ 8 servings/day
Fiber	$\geq$ 25 grams/day
Red meat intake	$<$ 0.5 servings/day
Fish intake	$\geq$ 2 servings/week
Nuts & Seeds	$\geq$ 0.5 serving/day
Alcohol[a]	Females: $\leq$ 1 drink/day Males: $\leq$ 2 drinks/day

BMI Body mass index. **MET** Metabolic equivalent of task. [a] One drink is equivalent to 14g of alcohol.

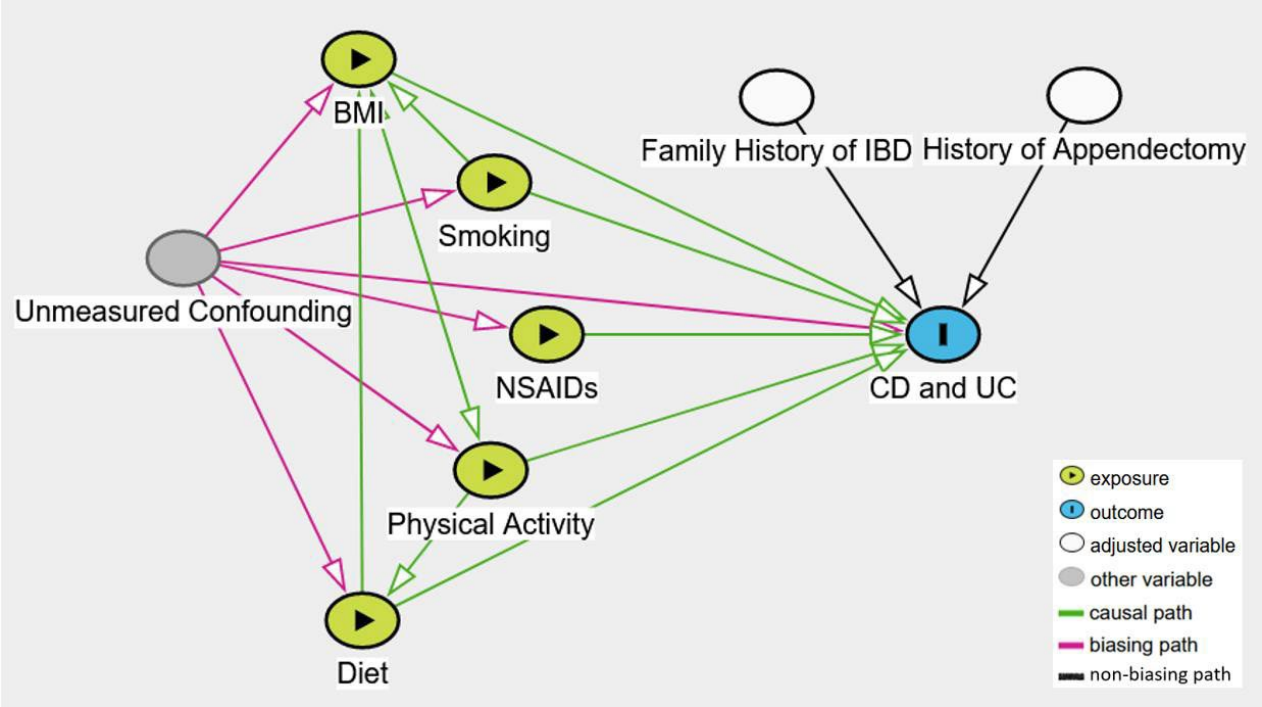


Figure S1. Directed acyclic graph (DAG) for proposed relationships between risk factors, outcomes, and potential confounders. **CD** Crohn's disease. **NSAID** Non-steroidal anti-inflammatory drug. **UC** Ulcerative colitis.

Table S2. Baseline characteristics of pooled primary cohort according to modifiable risk score (MRS).

	CD-MRS		UC-MRS	
	MRS 0-1 (n=18,110)	**MRS 6 (n=1,207)**	**MRS 0-2 [a] (n=10,083)**	**MRS 6 (n=8,567)**
Age [b] (years)	45.95 (11.4)	42.87 (9.4)	47.6 (10.8)	40.2 (9.3)
Sex (% female)	79	84	85	75
Body mass index (kg/m^2)	23.5 (3.1)	35.04 (4.8)	26.9 (6.2)	24.2 (2.9)
Physical activity (MET-hrs/wk)	41.0 (38.0)	7.2 (7.3)	27.3 (33.8)	16.8 (21.0)
Smoking status				
- Never (%)	85	0	39	63
- Past (%)	12	65	26	37
- Current (%)	3	35	35	0
Regular NSAID use [c] (%)	12	55	15	61
History of appendectomy (%)	17	29	20	23
Family history of IBD (%)	4	4	4	4
Fruits (servings/day)	2.9 (1.7)	1.0 (0.8)	2.6 (1.8)	1.2 (0.8)
Vegetables (servings/day)	6.6 (2.9)	2.9 (1.3)	6.7 (3.3)	2.9 (1.2)
Fiber (grams/day)	27.1 (6.5)	15.5 (3.4)	25.1 (8.2)	17.1 (4.5)
Red meat (servings/day)	0.5 (0.5)	0.9 (0.6)	0.3 (0.4)	0.8 (0.5)
n3:n6 PUFA ratio	0.15 (0.05)	0.12 (0.03)	0.18 (0.06)	0.11 (0.02)

[a] Reference level for UC set to 0-2 given low number of scores 0-1. [b] All values other than age are standardized to the age distribution of the study population. Values are mean (standard deviation) unless stated otherwise. Values of polytomous variables may not sum to 100% due to rounding. [c] NSAIDs use in year 1986 for NHS and 1991 for NHSII (regular use equals ≥2x/week). **CD** Crohn's disease. **IBD** Inflammatory bowel disease. **MET** Metabolic equivalent of task. **MRS** Modifiable Risk Score. **NSAID** Non-steroidal anti-inflammatory drug. **PUFA** Polyunsaturated fatty acid. **UC** Ulcerative colitis.

Table S3. Risk and PAR of CD and UC according to CD-MRS and UC-MRS, stratified by sex.

Females					Males					
	Cases	**Person-Years**	**aHR (95% CI)**	**PAR (95% CI)**		**Cases**	**Person-Years**	**aHR (95% CI)**	**PAR (95% CI)**	**$P_{interaction}$**
CD-MRS					**CD-MRS**					
0-1	14	327,753	1.00 (ref)	43.0% (10.1-67.4%)	0-1	2	96,430	1.00 (ref)	43.4% (-44.6-88.7%)	0.79
2	30	601,991	1.21 (0.64-2.29)		2	5	170,556	1.38 (0.27-7.17)		
3	105	1,366,130	1.81 (1.03-3.16)		3	11	363,031	1.33 (0.29-6.05)		
4	99	1,261,718	1.85 (1.05-3.24)		4	16	324,865	2.37 (0.54-10.42)		
5+	59	515,414	2.74 (1.52-4.93)		5+	5	89,134	2.49 (0.47-13.09)		
P_{trend}	--	--	<0.0001		P_{trend}	--	--	0.09		
UC-MRS					**UC-MRS**					
0-2	11	197,213	1.00 (ref)	42.3% (4.1-69.7%)	0-2	1	43,822	1.00 (ref)	61.1% (-46.1-95.8%)	0.19
3	46	604,700	1.44 (0.74-2.78)		3	4	132,473	1.42 (0.15-12.92)		
4	139	1,377,580	1.92 (1.04-3.54)		4	19	317,409	2.69 (0.35-20.45)		
5	161	1,604,875	1.86 (1.01-3.44)		5	27	481,684	2.40 (0.32-17.98)		
6	38	288,640	2.45 (1.25-4.81)		6	10	68,626	6.03 (0.74-49.00)		
P_{trend}	--	--	0.006		P_{trend}	--	--	0.03		

Cox models adjusted for age (months), appendectomy (yes/no), and family history IBD (yes/no). PAR calculations adjusted for age (< 60, ≥ 60), appendectomy (yes/no) and family history IBD (yes/no). UC-MRS reference level set to 0-2 given low number of UC-MRS of 0-1. CD-MRS set to maximum value 5+ given n=0 for a CD-MRS of 6 in men. **aHR** Adjusted hazard ratio. **CD** Crohn's disease. **CI** Confidence interval. **MRS** Modifiable risk score. **PAR** Population attributable Risk. **UC** Ulcerative colitis.

Table S4. Risk of CD and UC using weighted CD-MRS and UC-MRS criteria.

	Cases	Person-Years	aHR (95% CI)	PAR (95% CI)
CD-MRS				
Low	27	691,467	1.00 (ref)	41.0% (17.5-60.0%)
High	319	4,425,554	1.78 (1.20-2.63)	
P-value			0.004	
UC-MRS				
Low	48	734,969	1.00 (ref)	27.7% (7.5-45.7%)
High	408	4,382,053	1.46 (1.08-1.97)	
P-value			0.01	

MRS score ranges 6-30 based on scaled criteria for known risk factors of CD and UC. Low MRS defined as lowest 15% compared to high MRS (remainder). Cox models stratified by age (months), time-period (2-year intervals), and cohort (NHS, NHSII, or HPFS), and adjusted for appendectomy (yes/no), and family history IBD (yes/no). PAR calculations adjusted for age (<40, 40 ≤ age <60, ≥ 60), cohort (NHS, NHSII, HPFS), appendectomy (yes/no) and family history IBD (yes/no). **aHR** Adjusted hazard ratio. **CD** Crohn's disease. **CI** Confidence interval. **MRS** Modifiable risk score. **PAR** Population attributable Risk. **UC** Ulcerative colitis.

Table S5. Risk of CD according to known risk factors for disease.

	Mean (sd) or %	Cases	Person-Years	aHR (95% CI)	P-value [a]
Smoking status					
Never smoking	55	164	2,829,085	1.00 (ref)	
Past smoking	30	131	1,832,503	1.26 (1.00-1.60)	0.05
Current smoking	15	51	455,432	1.85 (1.33-2.56)	0.0002
Regular NSAID use					
< 2 times/week	82	271	4,195,245	1.00 (ref)	
≥ 2 times/week	18	75	921,776	1.17 (0.90-1.52)	0.25
Appendectomy					
No	81	264	4,109,334	1.00 (ref)	
Yes	19	82	1,007,688	1.21 (0.93-1.56)	0.15
Family history of IBD					
None	96	291	4,905,946	1.00 (ref)	
Positive family history	4	55	211,075	4.53 (3.38-6.07)	<.0001
BMI					
BMI < 30 kg/m^2	80	260	4,089,944	1.00 (ref)	
BMI ≥ 30 kg/m^2	20	86	1,027,077	1.20 (0.93-1.55)	0.16
Physical activity (MET-hrs/wk)					
1st quintile	2.8 (1.8)	64	864,365	1.00 (ref)	
2nd quintile	7.9 (3.0)	88	1,030,065	1.19 (0.86-1.65)	
3rd quintile	14.5 (4.5)	73	1,080,409	0.97 (0.69-1.37)	
4th quintile	24.3 (6.8)	63	1,095,906	0.81 (0.57-1.16)	
5th quintile	52.2 (27.0)	58	1,046,277	0.81 (0.55-1.18)	0.04
Fruit & vegetable intake					
1st quintile	2.4 (0.7)	74	989,114	1.00 (ref)	
2nd quintile	3.8 (0.4)	76	1,028,130	1.05 (0.75-1.47)	
3rd quintile	5.0 (0.5)	70	1,037,979	0.96 (0.67-1.40)	
4th quintile	6.3 (0.6)	68	1,039,140	0.97 (0.65-1.45)	
5th quintile	9.4 (2.4)	58	1,022,659	0.97 (0.62-1.52)	0.78
Fiber intake (g/day)					
1st quintile	13.1 (2.0)	74	954,654	1.00 (ref)	
2nd quintile	16.5 (1.4)	66	1,037,921	0.88 (0.62-1.25)	
3rd quintile	18.9 (1.6)	83	1,057,753	1.14 (0.79-1.63)	
4th quintile	21.6 (2.0)	79	1,055,670	1.14 (0.77-1.70)	
5th quintile	27.4 (4.9)	44	1,011,023	0.67 (0.41-1.11)	0.61
Red meat intake (servings/day)					
1st quintile	0.2 (0.1)	62	963,342	1.00 (ref)	
2nd quintile	0.4 (0.1)	73	1,012,905	0.98 (0.69-1.38)	
3rd quintile	0.6 (0.1)	65	1,072,761	0.81 (0.56-1.16)	
4th quintile	0.8 (0.1)	73	1,055,649	0.89 (0.62-1.27)	
5th quintile	1.3 (0.4)	73	1,012,365	0.92 (0.63-1.33)	0.74

Stratified by age (months), time-period (2-year intervals), and cohort (NHS, NHSII, or HPFS). Adjusted for BMI (< 30, or ≥ 30 kg/m^2); family history of IBD (yes/no); appendectomy (yes/no); physical activity (quintiles); smoking status (never, past, or current); regular NSAID use (yes/no); fruits, vegetables and red meat intake (quintiles servings/day), and fiber (quintiles g/day). [a] P-values for quintile data are P_{trend} values. **aHR** Adjusted hazard ratio. **BMI** Body mass index. **CD** Crohn's disease. **CI** Confidence interval. **IBD** Inflammatory bowel disease. **MET** Metabolic equivalent of task. **NSAID** Non-steroidal anti-inflammatory drug.

Table S6. Risk of UC according to known risk factors for disease.

	Mean (sd) or %	Cases	Person-Years	aHR (95% CI)	P-value [a]
Smoking status					
Current smoking	55	36	455,432	1.00 (ref)	
Never smoking	30	213	2,829,085	1.00 (0.70-1.43)	0.98
Past smoking	15	207	1,832,503	1.66 (1.16-2.38)	0.006
Regular NSAID use					
< 2 times/week	82	344	4,195,245	1.00 (ref)	
≥ 2 times/week	18	112	921,776	1.44 (1.15-1.79)	0.001
Appendectomy					
No	81	376	4,109,334	1.00 (ref)	
Yes	19	80	1,007,688	0.87 (0.68-1.12)	0.27
Family history of IBD					
None	96	398	4,905,946	1.00 (ref)	
Positive family history	4	58	211,075	3.24 (2.45-4.29)	<.0001
BMI					
BMI < 30 kg/m^2	80	367	4,089,944	1.00 (ref)	
BMI ≥ 30 kg/m^2	20	89	1,027,077	0.90 (0.71-1.14)	0.39
Physical activity (MET-hrs/wk)					
1st quintile	2.8 (1.8)	84	864,365	1.00 (ref)	
2nd quintile	7.9 (3.0)	83	1,030,065	0.87 (0.64-1.18)	
3rd quintile	14.5 (4.5)	98	1,080,409	0.99 (0.73-1.33)	
4th quintile	24.3 (6.8)	107	1,095,906	1.08 (0.80-1.45)	
5th quintile	52.2 (27.0)	84	1,046,277	0.92 (0.67-1.27)	0.81
Fruit & vegetable intake (servings/day)					
1st quintile	2.4 (0.7)	90	989,114	1.00 (ref)	
2nd quintile	3.8 (0.4)	119	1,028,130	1.31 (0.99-1.73)	
3rd quintile	5.0 (0.5)	76	1,037,979	0.83 (0.60-1.14)	
4th quintile	6.3 (0.6)	101	1,039,140	1.13 (0.83-1.53)	
5th quintile	9.4 (2.4)	70	1,022,659	0.82 (0.58-1.15)	0.15
Red meat intake (servings/day)					
1st quintile	0.2 (0.1)	83	963,342	1.00 (ref)	
2nd quintile	0.4 (0.1)	96	1,012,905	1.05 (0.78-1.41)	
3rd quintile	0.6 (0.1)	91	1,072,761	0.93 (0.68-1.26)	
4th quintile	0.8 (0.1)	97	1,055,649	0.99 (0.73-1.34)	
5th quintile	1.3 (0.4)	89	1,012,365	0.96 (0.70-1.31)	0.73
n3:n6 PUFA (ratio/day)					
1st quintile	0.09 (0.01)	101	960,831	1.00 (ref)	
2nd quintile	0.11 (0.01)	91	1,041,392	0.83 (0.63-1.11)	
3rd quintile	0.12 (0.01)	104	1,051,904	0.96 (0.72-1.27)	
4th quintile	0.14 (0.01)	79	1,050,103	0.73 (0.53-0.99)	
5th quintile	0.18 (0.04)	81	1,012,791	0.79 (0.57-1.09)	0.10

Stratified by age (months), time-period (2-year intervals), and cohort (NHS, NHSII, or HPFS). Adjusted for BMI (< 30, or ≥ 30 kg/m^2); family history of IBD (yes/no); appendectomy (yes/no); physical activity quintiles); smoking status (never, past, or current); regular NSAID use (yes/no); fruits, vegetables and red meat intake (quintiles servings/day), and n3:n6 PUFA (quintiles). [a] P-values for quintile data are P_{trend} values. **aHR** Adjusted hazard ratio. **BMI** Body mass index. **CI** Confidence interval. **IBD** Inflammatory bowel disease. **MET** Metabolic equivalent of task. **NSAID** Non-steroidal anti-inflammatory drug. **PUFA** Polyunsaturated fatty acid. **UC** Ulcerative colitis.

Table S7. Risk and PAR for Crohn's disease and ulcerative colitis according to individual risk factors.

	Reference level	Cases (ref)	Person-Years (ref)	aHR (95% CI)	P-value	PAR (95% CI)
Crohn's Disease [a]						
Family history	None	291	4,905,946	4.49 (3.35-6.01)	<0.0001	12.2% (8.0-16.2%)
Appendectomy	None	264	4,109,334	1.21 (0.93-1.56)	0.15	3.8% (-1.9-9.4%)
All Modifiable Risk Factors						
• BMI	<30 kg/m^2	260	4,089,944	1.18 (0.91-1.51)	0.21	4.3% (-2.3-10.7%)
• Past or current smoking	Never smokers	164	2,829,085	1.39 (1.12-1.73)	0.003	14.4% (1.5-26.9%)
• Regular NSAID use	<2x/week	271	4,195,245	1.16 (0.89-1.51)	0.27	3.2% (-2.4-8.8%)
• Physical activity [c]	Highest quintile	58	1,046,277	1.25 (0.86-1.81)	0.04	12.9% (-4.9-29.9%)
• Fruits & Vegetables [c]	Highest quintile	58	1,022,659	1.07 (0.69-1.64)	0.69	4.0% (-27.9-35.1%)
• Fiber [c]	Highest quintile	44	1,011,023	1.47 (0.93-2.34)	0.52	27.9% (-1.5-52.9%)
Ulcerative Colitis [b]						
Family history	None	398	4,905,946	3.24 (2.45-4.29)	<0.0001	8.8% (5.4-12.1%)
No Appendectomy	+Appendectomy	80	1,007,688	1.15 (0.90-1.47)	0.27	9.8% (-8.6-27.6%)
All Modifiable Risk Factors						
• BMI	≥30 kg/m^2	89	1,027,077	1.11 (0.87-1.41)	0.39	8.4% (-8.8-25.2%)
• Past smoking	Current smoking	36	455,432	1.66 (1.16-2.38)	0.006	18.0% (11.5-24.4%)
• Never smoking	Current smoking	36	455,432	1.00 (0.70-1.43)	0.98	N/A (HR ≤ 1)
• Regular NSAID use	<2x/week	344	4,195,245	1.44 (1.15-1.79)	0.001	7.1% (2.1-12.1%)
• Fruits & Vegetables [c]	Highest quintile	70	1,022,659	1.22 (0.87-1.73)	0.15	20.1% (-2.5-40.8%)
• Red meat [d]	Lowest quintile	83	963,342	0.96 (0.70-1.31)	0.73	2.9% (-12.4-18.1%)
• n3:n6 PUFA [c]	Highest quintile	81	1,012,791	1.27 (0.92-1.75)	0.10	11.0% (-4.8-26.4%)

All Cox models stratified by age (months), time-period (2-year intervals), and cohort (NHS, NHSII, or HPFS), and adjusted for BMI (<30, or ≥ 30 kg/m^2), family history of IBD (yes/no), appendectomy (yes/no), physical activity (quintiles MET-hrs/wk), smoking status (never, past, or current), regular NSAID use (yes/no), and fruits & vegetables (quintiles servings/day). [a]Additionally adjusted for fiber intake (quintiles g/day). [b]Additionally adjusted for red meat intake (quintiles servings/day) and n3:n6 PUFA (quintiles). [c]Protective factor; HR for lowest quintile and P_{trend} provided. [d]Risk factor; HR for highest quintile and P_{trend} provided. PAR calculations adjusted for age (<40, 40 ≤ age <60, ≥ 60 years), cohort (NHS, NHSII, HPFS), and other covariates as per Cox models. Indicator variables with HR<1 excluded from PAR analysis. **aHR** Multivariable-adjusted Hazard Ratio. **BMI** Body mass index. **CI** Confidence Interval. **NSAID** Non-steroidal anti-inflammatory drug. **PAR** Population Attributable Risk. **PUFA** Polyunsaturated fatty acid.

Table S8. E-values to assess minimum aHR that an unmeasured confounder would need to have with both the exposure and outcome to fully explain the observed relationships between CD- and UC-MRS and CD and UC, respectively.

CD-MRS	aHR	LL 95% CI	E-Value for aHR	E-Value for LL
CD-MRS 0-1	1.00	--	Ref	--
CD-MRS ≥ 2	1.85	1.12	3.10	1.49
CD-MRS = 2	1.24	0.68	1.79	1.00
CD-MRS = 3	1.76	1.04	2.92	1.24
CD-MRS = 4	1.92	1.13	3.25	1.51
CD-MRS = 5	2.53	1.44	4.50	2.24
CD-MRS = 6	4.15	1.95	7.77	3.31
UC-MRS				
UC-MRS 0-2	1.00	--	Ref	--
UC-MRS ≥ 3	1.92	1.08	3.25	1.37
UC-MRS = 3	1.43	0.76	2.21	1.00
UC-MRS = 4	1.97	1.10	3.35	1.43
UC-MRS = 5	1.90	1.06	3.21	1.31
UC-MRS = 6	2.78	1.47	5.00	2.30

Supplementary References:

1. The Swedish Mammography Cohort (SMC) | Karolinska Institutet. Accessed April 22, 2021. https://ki.se/en/imm/the-swedish-mammography-cohort-smc

2. Harris H, Håkansson N, Olofsson C, Stackelberg O, Julin B, Åkesson A. The Swedish mammography cohort and the cohort of Swedish men: study design and characteristics of two population-based longitudinal cohorts. *OA Epidemiology*. 2013;1(2):16. http://ki.se/imm/

3. COSM - a cohort of 50,000 Swedish men | Karolinska Institutet. Accessed April 22, 2021. https://ki.se/en/imm/cosm-a-cohort-of-50000-swedish-men

4. European Prospective Investigation into Cancer and Nutrition (EPIC) - Study Resources. Accessed May 10, 2021. https://epic.iarc.fr/about/studyresources.php

5. Riboli E, Hunt K, Slimani N, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr*. 2002;5(6b):1113-1124. doi:10.1079/PHN2002394

6. Jakobsson GL, Sternegård E, Olén O, et al. Validating inflammatory bowel disease (IBD) in the Swedish National Patient Register and the Swedish Quality Register for IBD (SWIBREG). *Scand J Gastroenterol*. 2017;52(2):216-221. doi:10.1080/00365521.2016.1246605

7. Bergmann MM, Hernandez V, Bernigau W, et al. No association of alcohol use and the risk of ulcerative colitis or Crohn's disease: data from a European Prospective cohort study (EPIC). *European Journal of Clinical Nutrition 2017 71:4*. 2017;71(4):512-518. doi:10.1038/ejcn.2016.271

8. Wang L, Du M, Wang K, et al. Association of ultra-processed food consumption with colorectal cancer risk among men and women: Results from three prospective US cohort studies. *The BMJ*. 2022;378. doi:10.1136/bmj-2021-068921

9. Wang YX, Li Y, Rich-Edwards JW, et al. Associations of birth weight and later life lifestyle factors with risk of cardiovascular disease in the USA: A prospective cohort study. *EClinicalMedicine*. 2022;51:101570. doi:10.1016/j.eclinm.2022.101570

10. World Health Organization. *Ischemic Heart Disease Registers: Report of the Fifth Working Group.*; 1971. Accessed October 18, 2022. https://books.google.com/books/about/Ischemic_Heart_Disease_Registers.html?id=frb3MwEACAAJ

11. Alpert J, Thygesen K, Antman E, Bassand J. Myocardial infarction redefined--a consensus document of The Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction. *J Am Coll Cardiol*. 2000;36(3):959-969. doi:10.1016/S0735-1097(00)00804-4

12. Walker A, Robins M, Weinfeld F. The National Survey of Stroke. Clinical findings. . *Stroke*. 1981;12(2):13-44. doi:10.1007/BF01186982

13. Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of Physical Activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc*. 1993;25(1):71-80.

14. Orsini N, Bellocco R, Bottai M, Pagano M, Wolk A. Age and temporal trends of total physical activity among Swedish women. *Med Sci Sports Exerc*. 2006;38(2):240-245. doi:10.1249/01.mss.0000185086.19220.b3

15. Orsini N, Bellocco R, Bottai M, et al. Validity of self-reported total physical activity questionnaire among older women. *Eur J Epidemiol*. 2008;23(10):661-667. doi:10.1007/s10654-008-9273-z

16. Higuchi LM, Khalili H, Chan AT, Richter JM, Bousvaros A, Fuchs CS. A prospective study of cigarette smoking and the risk of inflammatory bowel disease in women. *American Journal of Gastroenterology*. 2012;107(9):1399-1406. doi:10.1038/ajg.2012.196

17. Ananthakrishnan AN, Higuchi LM, Huang ES, et al. Aspirin, Nonsteroidal Anti-inflammatory Drug Use, and Risk for Crohn Disease and Ulcerative Colitis A Cohort Study. *Ann Intern Med*. 2012;156:350-359. https://annals.org

18. Harvard T.H. Chan School of Public Health Nutrition Department. Nutrient Tables. Accessed March 2, 2020. https://regepi.bwh.harvard.edu/health/

19. Yuan C, Spiegelman D, Rimm EB, et al. Validity of a Dietary Questionnaire Assessed by Comparison With Multiple Weighed Dietary Records or 24-Hour Recalls. *Am J Epidemiol*. 2017;185(7):570-584. doi:10.1093/aje/kww104

20. Rimm EB, Giovannucci EL, Stampfer MJ, et al. Reproducibility and Validity of an Expanded Self-Administered Semiquantitative Food Frequency Questionnaire among Male Health Professionals. *Am J Epidemiol*. 1992;135(10). https://academic.oup.com/aje/article-abstract/135/10/1114/67168

21. Willett W, Stampfer MJ, Hegsted M, et al. Total Energy Intake: Implications for Epidemiologic Analyses. *Am J Epidemiol*. 1986;124(1).

22. Rhee JJ, Cho E, Willett WC. Energy adjustment of nutrient intakes is preferable to adjustment using body weight and physical activity in epidemiological analyses. *Public Health Nutr*. 2014;17(5):1054-1060. doi:10.1017/S1368980013001390

23. Willett WC. Food-Frequency Methods. In: *Nutritional Epidemiology*. Oxford University Press; 1998. doi:10.1093/acprof:oso/9780195122978.003.05

24. Kaaks R, Slimani N, Riboli E. Pilot Phase Studies on the Accuracy of Dietary Intake Measurements in the EPIC Project: Overall Evaluation of Results. *Int J Epidemiol*. 1997;26(1).

25. Amiano P, Barcos A, Barricarte A, et al. Dietary intake of vegetables and fruits among adults in five regions of Spain. EPIC Group of Spain. European Prospective Investigation into Cancer and Nutrition. *Eur J Clin Nutr*. 1999;53(3):174-180.

26. Ashfield-Watt P, Welch A, Day N, Bingham S. Is 'five-a-day' an effective way of increasing fruit and vegetable intakes? *Public Health Nutr*. 2004;7(2):257-261. doi:10.1079/phn2003524

27. Crawley H, Mills A, Patel S. *Food Portion Sizes*. 2nd ed. HM Stationery Office; 1993.

28. Pan A, Sun Q, Bernstein AM, et al. Red meat consumption and mortality: Results from 2 prospective cohort studies. *Arch Intern Med*. 2012;172(7):555-563. doi:10.1001/archinternmed.2011.2287

29. American Heart Association. Guidelines for Healthy Living. Accessed April 22, 2021. https://www.heart.org/en/healthy-living

30. U.S. Department of Agriculture US Department of Health and Human Services. 2020-2025 Dietary Guidelines for Americans. https://www.

31. Piercy KL, Troiano RP, Ballard RM, et al. The physical activity guidelines for Americans. *JAMA - Journal of the American Medical Association*. 2018;320(19):2020-2028. doi:10.1001/jama.2018.14854

32. U.S. Department of Agriculture US Department of Health and Human Services. 2010 Dietary Guidelines for Americans. Accessed April 21, 2021. www.dietaryguidelines.gov

33. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: Introducing the E-Value. *Ann Intern Med*. 2017;167(4):268-274. doi:10.7326/M16-2607