Mendelian randomization analysis of C-reactive protein on colorectal cancer risk

Authors:

Xiaoliang Wang* [1,2], James Y. Dai[2], Demetrius Albanes[3], Volker Arndt[4], Sonja I. Berndt[3], Stéphane Bézieau[5], Hermann Brenner[4,6,7], Daniel D. Buchanan[8,9,10,11], Katja Butterbach[12], Bette Caan[13], Graham Casey[14], Peter T. Campbell[15], Andrew T. Chan[16], Zhengyi Chen[17], Jenny Chang-Claude[12, 18], Michelle Cotterchio[19, 20], Douglas F. Easton[21], Graham G. Giles[22,23], Edward Giovannucci[24], William M. Grady[25,26], Michael Hoffmeister[4], John L. Hopper[8], Li Hsu[2], Mark A. Jenkins[8], Amit D. Joshi[27], Johanna W. Lampe[1,2], Susanna C. Larsson[28], Flavio Lejbkowicz[29], Li Li[17], Annika Lindblom[30], Loic Le Marchand[31], Vicente Martin[32], Roger L. Milne[22,23], Victor Moreno[33,34], Polly A. Newcomb[1,2], Kenneth Offitt[35], Shuji Ogino[36], Paul D.P. Pharoah[21], Mila Pinchev[29], John D. Potter[1,2,37], Hedy S. Rennert[29], Gad Rennert[29], Walid Saliba[29], Clemens Schafmayer[4], Robert E. Schoen[38], Petra Schrotz-King[6], Martha L. Slattery[39], Mingyang Song[27, 40], Christa Stegmaier[41], Stephanie J. Weinstein[3], Alicja Wolk[28, 42], Michael O Woods[43], Anna H. Wu[10], Stephen B. Gruber[44], Ulrike Peters[1,2], Emily White[1,2]

Affiliations:

[1] Department of Epidemiology, University of Washington, Seattle, WA, USA

[2] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[3] Division of Cancer Epidemiology and Genetics, US National Cancer Institute, NIH, Rockville, MD, USA

[4] Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany

[5] Service de Génétique Médicale, CHU Nantes, Nantes, France

[6] Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany

[7] German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

[8] Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Parkville, Victoria, Australia

[9] Colorectal Oncogenomics Group, Department of Clinical Pathology, The University of Melbourne, Parkville, Victoria, Australia

[10] University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, Victoria, Australia

[11] Genomic Medicine and Family Cancer Clinic, Royal Melbourne Hospital, Melbourne, Victoria, Australia

[12] Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany

[13] Division of Research, Kaiser Permanente Medical Care Program, Oakland, CA, USA

[14] Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

[15] Epidemiology Research Program, American Cancer Society, Atlanta, GA, USA

[16] Division of Gastroenterology, Massachusetts General Hospital, Boston, MA, USA

[17] Department of Family Medicine and Community Health, Mary Ann Swetland Center for Environmental Health, Case Western Reserve University, Cleveland, OH, USA

[18] Genetic Tumour Epidemiology Group, University Cancer Center Hamburg, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[19] Prevention and Cancer Control, Cancer Care Ontario, Toronto, Ontario, Canada

[20] Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada

[21] Centre for Cancer Genetic Epidemiology, Departments of Oncology and Public Health and Primary Care, University of Cambridge, Cambridge, UK

[22] Cancer Epidemiology & Intelligence Division, Cancer Council Victoria, Melbourne, Australia

[23] Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia

[24] Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

[25] Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[26] Department of Medicine, Gastroenterology Division, University of Washington School of Medicine, Seattle, WA, USA

[27] Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

[28] Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

[29] Department of Community Medicine and Epidemiology, Carmel Medical Center and B. Rappaport Faculty of Medicine, Technion, Haifa, Israel

[30] Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden

[31] Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI, USA

[32] Research Group on Gene-Environment Interactions and Health (GIIGAS), University of León and CIBERESP, León, Spain

[33] Cancer Prevention and Control Program, Catalan Institute of Oncology (ICO), IDIBELL, CIBERESP, Barcelona, Spain

[34] Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain

[35] Clinical Genetics Research Lab, Department of Cancer Biology and Genetics; Clinical Genetics Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[36] Department of Pathology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, USA

[37] Centre for Public Health Research, Massey University, Wellington, New Zealand

[38] Department of Medicine and Epidemiology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

[39] Department of Internal Medicine, University of Utah Health Sciences Center, Salt Lake City, UT, USA

[40] Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[41] Saarland Cancer Registry, Saarland, Germany

[42] Department of Surgical Sciences, Uppsala University, Uppsala, Sweden

[43] Discipline of Genetics, Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland & Labrador, Canada

[44] University of Southern California Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA, USA.

Short Title: Mendelian randomization of CRP on CRC risk

Corresponding Author: Xiaoliang Wang; Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N., M4-B402, Seattle, WA 98109-1024; Email: xwang23@fredhutch.org; Tel: +1(206)667-6503; Fax: +1(206)667-7850.

Abstract

Background: Chronic inflammation is a risk factor for colorectal cancer (CRC). Circulating C-reactive protein (CRP) is also moderately associated with CRC risk. However, observational studies are susceptible to unmeasured confounding or reverse causality. Using genetic risk variants as instrumental variables, we investigated the causal relationship between genetically elevated CRP concentration and CRC risk using a Mendelian randomization approach.

Methods: Individual-level data from 30 480 CRC cases and 22 844 controls from 33 participating studies in three international consortia were used: the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO), the Colorectal Transdisciplinary Study (CORECT) and the Colon Cancer Family Registry (CCFR). As instrumental variables, we included 19 SNPs previously associated with CRP concentration. The SNP-CRC associations were estimated using a logistic regression model adjusted for age, sex, principal components and genotyping phases. An inverse-variance weighted method was applied to estimate the causal effect of CRP on CRC risk.

Results: Among the 19 CRP-associated SNPs, rs1260326 and rs6734238 were significantly associated with CRC risk ($p=7.5\times10^{-4}$, and p=0.003, respectively). A genetically predicted one-unit increase in the log-transformed CRP concentrations (mg/L) was not associated with increased risk of CRC (OR=1.04; 95% CI: 0.97-1.12; p=0.256). No evidence of association was observed in subgroup analyses stratified by other risk factors.

Conclusions: Albeit adequate statistical power to detect moderate association, we found genetically elevated CRP concentration was not associated with increased risk of CRC among individuals of European ancestry. Our findings suggested that circulating CRP is unlikely to be a causal factor in CRC development.

(Word count: 247)

Key Words: C-reactive protein; Colorectal cancer; Mendelian randomization; Epidemiology

Key Messages:

- Meta-analyses of observation studies reported a moderate association between elevated C-reactive protein (CRP) concentration and colorectal cancer (CRC) risk; however, whether the association is causal is unclear.
- In this largest study up to date, we had adequate statistical power to assess the causal relationship between circulating CRP concentration and CRC risk, using Mendelian randomization analysis.
- We found that genetically elevated CRP concentration was not associated with increased risk of CRC among individuals of European ancestry, suggesting CRP is unlikely to play a causal role in CRC development.
- No evidence of genetically elevated CRP-CRC association was observed in subgroup analyses stratified by other risk factors.

Introduction

Chronic inflammation plays a role in the pathogenesis of colorectal cancer (CRC) (1). Meta-analyses of observational studies have shown that a one unit (mg/L) increase in log-transformed high-sensitivity C-reactive protein (CRP), a common biomarker for low-grade chronic inflammation, was associated with 12% higher risk of CRC (2, 3). The association was stronger among men than women, and was stronger in colon cancer than rectal cancer. Although observational studies support a role for CRP in CRC development, they are susceptible to potential bias by unmeasured confounders, such as older age (4), adiposity (5), tobacco smoking (6, 7), lower physical activity (8), and lower use of NSAIDs (9). Observational studies are also susceptible to reverse causality, in which elevated CRP concentrations are due to immune response and inflammation induced by premalignant or preclinical tumor growth (10-12).

Mendelian randomization analysis, by taking advantage of the random assortment of genetic alleles during gamete formation, is less susceptible to confounding or reverse causality (13). Because genetic variants are distributed randomly at conception, they are generally unrelated to environmental risk factors, and temporally precede both risk factors and the disease process. The heritability of CRP was estimated to range between 25-40%, suggesting a role of genetic factors in baseline CRP concentrations (14). Several studies have used CRP-related genetic variants as a proxy of lifelong CRP concentrations on CRC risk but reported inconsistent findings. A nested case-control study found genetically elevated CRP concentration, based on seven SNPs in the *CRP* gene, was associated with higher CRC risk (15). Another case-control study found a tagSNP in the *CRP* gene to be associated with higher risk of colon cancer, and another SNP associated with lower risk of rectal cancer (16). Other studies using SNPs within the *CRP* gene did not find associations between CRP and CRC risk (17-19). In a prospective cohort study, Prizment et al (20) reported an association between a weighted CRP genetic risk score, based on 20 CRP-associated SNPs identified in a meta-analysis of GWAS studies (14), and CRC risk, corroborating a causal role for CRP in colorectal carcinogenesis. However, the cohort had limited statistical power due to relatively small sample size (7,603 participants with 205 CRC cases). In addition, most previous studies assumed population homogeneity and did not adjust for population stratification, which could bias the results. Furthermore, other SNPs recently found to be associated with CRP concentration were not included (21).

Findings from previous human genetic studies have been inconsistent and have had insufficient power to assess a moderate causal relationship between CRP and CRC risk. In this study, we aimed to investigate whether CRP

plays a causal role in CRC risk using genetic variants that were previously reported to be significantly associated with circulating CRP concentration as instrumental variables (IVs), using the largest study for such analysis to date.

Methods

Study Participants

We used epidemiological and genetic data from 30 480 CRC cases and 22 844 controls from 33 participating studies in three international CRC consortia: the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO), the Colorectal Transdisciplinary Study (CORECT) and the Colon Cancer Family Registry (CCFR). Full details have been published previously (22, 23), and the demographic characteristics of study participants are summarized in Supplemental Table 1. In brief, 10 644 cases and 10 729 controls were included from GECCO from nested case-control studies in 8 cohorts and 6 case-control studies. Further, 19 836 cases and 12 115 controls were included from CORECT from nested case-control studies in 7 cohorts, 9 case-control studies and 3 case-series studies. Nested case-control studies from CCFR participated as individual studies in GECCO and/or CORECT. There was no overlap of participants between studies.

Participants with non-European ancestry were excluded. Informed consent was given by all participants, and studies were approved by their respective Institutional Review Boards.

Assessment of Outcomes and Environmental Variables

Invasive CRC cases (International Classification of Disease for Oncology Code 18.0-18.9, 19.9 and 20.9) were identified by medical record, pathology report, death certificate or record linkage. Age at diagnosis, cancer subsites and stages were obtained from medical records and cancer registries. Patients with Lynch Syndrome and other syndromic causes were excluded. Controls were selected based on study-specific eligibility and matching criteria. Case-series studies only contributed cases to this study.

Demographic and environmental factors were self-reported at either in-person interview or via structured self-administered questionnaires, based on each study. A multistep, iterative data harmonization procedure was applied, and was described previously (22). Age was defined as age at CRC diagnosis for cases, or age at selection for controls. Body mass index (BMI; kg/m^2) was categorized as normal (18.5-24.9), overweight (25-29.9), and obese ($\geq$30). Participants with BMI<18.5 were excluded. Smoking status was defined as never and

ever smokers. Regular use of any NSAIDs, aspirin, or non-aspirin NSAIDs was defined as binary (yes/no). Family history of CRC was defined as CRC occurring in any first degree relative. History of endoscopy included both sigmoidoscopy and colonoscopy.

Genotyping

Details on genotyping and imputation have been reported previously (24). In brief, DNA was mostly obtained from blood samples, with some from buccal swabs. Several platforms (the Illumina HumanHap 300k, 240k, 550k and OncoArray 610k BeadChip Array system, or Affymetrix platform) were used for genotyping (25, 26). Samples were excluded on the basis of sample call rate ≤97%, heterozygosity, unexpected duplicates or relative pairs, gender discrepancy and principal component analysis (PCA) outlier of HapMap2 CEU cluster. SNPs were excluded on the basis of inconsistency across platforms, call rate <98%, and out of Hardy-Weinberg equilibrium (HWE) in controls ($p<0.0001$) (25). SNPs were imputed from the 1000 Genome Project reference panel if not directly genotyped, and restricted by imputation accuracy ($R^2>0.3$).

Instrumental Variables

All the selected SNPs came from two resources: 18 SNPs that had been previously used as IVs (14) and 9 SNPs from more recent findings among participants of European ancestry (21) (summarized in Table 1). SNPs in association with CRP concentration at the genome-wide significance threshold of $p<5\times10^{-8}$ were selected. We checked independence between the 27 selected SNPs using linkage disequilibrium (LD) analysis. If two SNPs were in LD ($R^2>0.2$), the SNP with the smaller p-value was included in the final SNP set, and the other excluded. We also conducted a GWAS-catalog search for SNPs that were associated with CRP ($p<5\times10^{-8}$) among participants of European ancestry, had reported the estimated effect sizes and standard errors, and were not in LD with the selected SNPs. No additional SNPs were identified. Altogether, 19 SNPs were included in the final IV set. The allele associated with higher CRP level was coded as risk allele, and the other allele was coded as baseline allele for all SNPs.

Three basic assumptions are made in Mendelian randomization: (i) the genetic marker is robustly associated with the exposure, (ii) the genetic marker is independent of the outcome, given the exposure and confounders of the exposure-outcome association (i.e. the genetic marker has no pleiotropic effect through pathways other than the exposure), and (iii) the genetic marker is independent of factors that confound the exposure-outcome

relation (27). The first assumption was met since we only included SNPs that were significantly associated with CRP concentrations in GWAS. The second assumption could not be tested directly because CRP measures were not available in our study, but sensitivity analyses were performed for global pleiotropic effects. The third assumption was tested by evaluating the association between SNPs and each potential confounder of the CRP-CRC association among controls. No evidence of violation of this assumption was observed. In addition, if multiple IVs are combined into a single estimate by the inverse-variance weighted (IVW) method, a further assumption is made that the variants provide independent information (i.e. not in LD) (28). Furthermore, the statistical association between the risk factor and a valid IV should be strong enough to provide unbiased and precise estimates in finite samples (29). The estimated CRP variance explained by the selected SNPs (R^2) was ~5%. Given the sample size of 53 325 subjects and 19 instruments in our study, the estimated F-statistics was 147.66 (29-31), suggesting strong instruments for the Mendelian randomization analysis.

Statistical Analysis

We performed the Mendelian randomization analysis to estimate the causal effect of CRP on CRC risk using inverse-variance weighted (IVW) method, by summarizing SNP-CRP associations from literature and estimating SNP-CRC associations in our study population. Assuming all the prior assumptions previously stated are met, genetic variant $k, (k = 1 \dots K)$ is associated with an observed X_k mean change in the risk factor per additional variant allele with standard error σ_{Xk} and an observed Y_k log-odds change in the outcome per allele with standard error σ_{Yk}. Assuming additive effects of SNPs on CRP concentrations, an IVW estimate of the causal effect combining the ratio estimates and standard errors of single SNPs can be computed as (28):

$$\hat{\beta}_{IVW} = \frac{\sum_k X_k Y_k \, \sigma_{Yk}^{-2}}{\sum_k X_k^2 \, \sigma_{Yk}^{-2}}$$

and the approximate standard error will be $se(\hat{\beta}_{IVW}) = \sqrt{\frac{1}{\sum_k X_k^2 \sigma_{Yk}^{-2}}}$.

The mean change in log-transformed CRP concentration (mg/L) per variant allele and its standard error (X_k and σ_{Xk}) were obtained from prior studies (14, 21), and the effect size of genetic variants on CRC risk were estimated within our study populations. We used logistic regression models to estimate the association between each genetic variant and CRC risk in GECCO and CORECT separately, adjusting for age, sex, genotyping phase, and principle components. The estimates from GECCO and CORECT were then combined into a summary

causal estimate using fixed-effect meta-analysis if there is no heterogeneity. Random-effect meta-analysis would be used otherwise.

Exploratory stratified analyses were carried out using an *a priori* list of CRC risk factors and the same regression models, including sex, BMI, smoking, NSAID use, aspirin use, family history of CRC and history of endoscopy. In addition, we evaluated differences by cancer subsites and stages. We also performed sensitivity analysis using Egger regression (32) for global pleiotropic effect.

Power Calculation

Based on the methods described by Burgess (33), our sample size of 30 480 CRC cases and 22 844 controls has an estimated 99.4% power to detect the previously estimated causal effect size of CRP (OR=1.19) (20) at a significance level of 0.05, assuming the SNPs explain a total of 5% variance of CRP based on previous estimates (14). Alternatively, we have 82.5% power to detect a minimal odds ratio of 1.12 (2, 3) at a significance level of 0.05, given our sample size.

Results

The mean age of participants was 63.4 years (SD=10), and 50.7% were male (Supplemental Table 1). A total of 27 SNPs were identified and their associations with CRP concentration are summarized in Table 1. The imputation accuracy (R^2) ranged between 0.84 and 1.0. The estimated associations between the 19 SNPs and CRC risk are shown in Figure 1. In pooled analysis combining GECCO and CORECT estimates, rs1260326 (T/C) was associated with higher risk of CRC ($p=7.5\times10^{-4}$), and rs6734238 (G/A) was associated with lower CRC risk (p=0.003). No other SNP was statistically significantly associated with CRC.

Using the 19 SNPs as IVs, we found that one unit increase in the log-transformed genetically elevated CRP concentration (mg/L) was associated with a 4% higher risk of CRC (OR=1.04; 95% CI: 0.97, 1.12; Table 2); however, the association was not statistically significant (p=0.256). No heterogeneity was observed between the two consortia (p-heterogeneity=0.509).

Genetically elevated CRP concentration was not associated with CRC risk in any of the subgroups defined by sex, BMI, smoking, NSAID use, aspirin use, family history of CRC or history of endoscopy (Table 3). The strength of associations between genetically elevated CRP concentration and CRC risk was similar between subgroups.

We also stratified by CRC subsites and stages. Genetically elevated CRP concentration was not associated with any subsite of CRC. There was a association between genetically elevated CRP concentration (mg/L) and distant CRC (OR=1.19; 95% CI: 1.00, 1.42; p=0.049), but not for local or regional CRC.

In sensitivity analysis, we observed no association between genetically elevated CRP and CRC risk using the two SNPs from the CRP gene only. Our results persisted using other Mendelian randomization methods (Supplemental Figure 1 and Supplemental Table 2). We also tested for global pleiotropic effect using Egger regression (Figure 2). None of the intercepts was significant (p>0.05), suggesting no global violation of pleiotropic assumptions.

Discussion

In this large multi-consortium study, we did not find evidence for an association between genetically elevated CRP concentrations and CRC risk among participants of European ancestry. No association was found in subgroups stratified by CRC risk factors. Our results suggest that circulating CRP does not play a causal role in colorectal carcinogenesis.

Our estimate of the CRP-CRC association is smaller than the 12% found in meta-analyses of prospective observational studies that used measured CRP concentrations (2, 3), suggesting that the association between measured CRP concentrations and CRC risk may be partially due to confounding. Our findings are different from the only previous study that used GWAS-identified SNPs as IVs for assessing the relationship between CRP and CRC risk (20). Prizment et al found a statistically significant 19% higher risk in CRC with a one unit (mg/L) increment of the log-transformed CRP concentration, while our analysis suggested only a modest non-significant effect size of 4%. The sample size of the previous study was small, with 205 CRC cases diagnosed in 7,603 participants. We used a much larger sample size of 30 480 CRC cases and 22 844 controls for adequate statistical power to test for moderate causal association. In addition, two SNPs that were not statistically significantly associated with CRP concentrations in GWAS (p-value > 5×10^{-8}) (14) were included in previous analysis (20). In comparison, we did not include these two SNPs, but rather included one additional SNP that was recently found to be associated with CRP concentration in a large consortium (p-value < 5×10^{-8}) (21) and was independent of the previous 18 SNPs. Lastly, there is possibility that the SNPs associated with CRP were

also associated with other inflammation-related traits, and led to a spurious positive association between CRP and CRC in the previous analyses.

Our findings are consistent with most prospective cohort studies that reported no causality using multiple SNPs as IVs (17-19). However, a nested case-control study found a two-fold higher genetically determined CRP concentration (mg/L), based on seven SNPs in the *CRP* gene, was associated with higher CRC risk (15). But the effect of genetically elevated CRP was not significantly attenuated after adjusting for measured CRP concentrations, indicating a potentially pleiotropic effect of selected SNPs which could lead to biased estimates of the SNP-CRP relationship in the first stage.

Chronic inflammation is a key predisposing factor in colorectal neoplasia (1). It has been suggested that chronic inflammation creates a microenvironment that promotes inflammatory cells to release reactive oxygen and nitrogen species which could lead to malignant DNA alteration (34), and increase the production of inflammatory cytokines that promote tumor growth (35). As a biomarker of low-grade inflammation, CRP has been proposed to play a role in colorectal carcinogenesis. CRP was found to be a major serum leptin-interacting protein that directly inhibited the binding of leptin to its receptors and its ability to signal *in vitro*, which resulted in leptin resistance and obesity *in vivo* (36). Lower concentrations of leptin and circulating adiponectin were also found in patients with CRC and adenomas as compared to controls (37), suggesting the possibility of interaction between CRP and leptin in colorectal carcinogenesis. However, a case-control study reported no association between circulating CRP concentration and pathologic measures of colonic inflammation (38). Mendelian randomization analysis also found that CRP concentration itself was unlikely to be a causal factor in coronary heart disease (39), although persistent inflammation was found to be a contributor (40). Similar to coronary heart disease, it is possible that chronic inflammation promotes colorectal carcinogenesis through other inflammatory mediators than CRP.

Our study has several strengths. It is the largest study to investigate causality between CRP and CRC risk using genetic variants, and has adequate statistical power to detect a moderate association. It is also the first to explore whether effects differed between subgroups stratified by other CRC risk factors, and cancer subsites and stages. Since we have environmental factors measured in most of the participating studies, we were able to test the assumption that the genetic variants were not associated with confounders of CRP and CRC. In addition, we used a comprehensive set of GWAS-identified genetic variants. Taking advantage of the random assortment of

alleles, our results from Mendelian randomization should be less susceptible to confounding and reverse causality compared with observational studies. Therefore, our results provide stronger evidence of non-causality on this topic. Furthermore, we adjusted for principal components, which accounted for potential confounding by population stratification.

There are some limitations. First, two Mendelian randomization assumptions could not be fully tested. Therefore, potential violations of the assumptions cannot be ruled out. Although we tested the assumption of no association between genetic variants and confounders, there are possible unmeasured confounders. We also performed diagnostic tests for the assumption of no pleiotropic effects. Second, we only investigated genetically elevated CRP in relation to CRC risk. Since we included studies from multiple countries, the association between genetic variants and CRP concentrations may be influenced by various environmental factors. Furthermore, there is possibility of selection bias. Compared to the average ages in the GWAS of CRP (range: 31-76) (14), our study participants were slightly older, but remained within the range of the GWAS samples. In Mendelian randomization studies, there is also possibility of survivor bias (41) where cases survived long enough and controls remained cancer-free until study recruitment in population-based case-control studies. However, most of our studies were cohorts, and we only included incident cases whose date of diagnosis were relatively close to the recruitment in case-control studies. Lastly, our results may not be generalizable to race/ethnicity groups other than European, in which the associations between genetic variants and CRP concentrations and CRC risk may be different.

In summary, we found that genetically elevated CRP concentration was not associated with increased risk of CRC among participants of European ancestry. Our findings do not support a causal role of CRP in CRC risk.

(Word count: 2 998)

Acknowledgements

The authors would like to thank all those at the GECCO Coordinating Center for helping bring together the data and people that made this project possible. The authors also acknowledge Deanna Stelling, Mark Thornquist, Greg Warnick, Carolyn Hutter, and team members at COMPASS (Comprehensive Center for the Advancement of Scientific Strategies) at the Fred Hutchinson Cancer Research Center for their work harmonizing the GECCO epidemiological data set. The authors acknowledge Dave Duggan and team members at TGEN (Translational Genomics Research Institute), the Broad Institute, and the Génome Québec Innovation Center for genotyping DNA samples of cases and controls, and for scientific input for GECCO.

The authors wish to acknowledge the contribution of Alexandre Belisle and the genotyping team of the McGill University and Génome Québec Innovation Centre, Montréal, Canada, for genotyping the Sequenom panel in the NFCCR samples.

We are also very grateful to Dr. Bruno Buecher without whom this project would not have existed. We also thank all those who agreed to participate in this study, including the patients and the healthy control persons, as well as all the physicians, technicians and students in the ASTERISK study.

Biospecimens of ColoCare studies were provided by the ColoCare Consortium, funded by the Fred Hutchinson Cancer Research Center.

The authors would like to thank the CPS-II participants and Study Management Group for their invaluable contributions to this research. The authors would also like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention National Program of Cancer Registries, and cancer registries supported by the National Cancer Institute Surveillance Epidemiology and End Results program.

We are thankful to all participants and cooperating clinicians, and Ute Handte-Daub, Utz Benscheid, Muhabbet Celik and Ursula Eilber for excellent technical assistance in the DACHS study.

We would like to acknowledge Patrice Soule and Hardeep Ranu of the Dana Farber Harvard Cancer Center High-Throughput Polymorphism Core who assisted in the genotyping for NHS, HPFS, and PHS under the supervision of Dr. Immaculata Devivo and Dr. David Hunter, Qin (Carolyn) Guo and Lixue Zhu who assisted in programming for NHS and HPFS, and Haiyan Zhang who assisted in programming for the PHS. We would like to thank the participants and staff of the Nurses' Health Study and the Health Professionals Follow-Up Study, for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. We also would like to acknowledge Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital. The authors assume full responsibility for analyses and interpretation of these data.

The authors would also like to thank Drs. Christine Berg and Philip Prorok, Division of Cancer Prevention, National Cancer Institute, the Screening Center investigators and staff of PLCO Cancer Screening Trial, Mr. Tom Riley and staff, Information Management Services, Inc., Ms. Barbara O'Brien and staff, Westat, Inc., and Drs. Bill Kopp and staff, SAIC-Frederick. Most importantly, we acknowledge the study participants for their contributions to making this study possible. The statements contained herein are solely those of the authors and do not represent or imply concurrence or endorsement by NCI.

The authors would like to thank the study participants and staff of the Hormones and Colon Cancer study.

The authors would also like to thank the investigators and staff in the Swedish Low-Risk Colorectal Cancer Study group, including David Edler, Claes Lenander, Johan Dalén, Fredrik Hjern, Nils Lundqvist, Ulrik Lindforss, Lars Påhlman, Kennet Smedh, Anders Törnqvist, Jörn Holm, Martin Janson, Magnus Andersson, Susanne Ekelund and Louise Olsson.

The authors also thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf

Funding

GECCO (Genetics and Epidemiology of Colorectal Cancer Consortium) is supported by the National Cancer Institute (NCI), National Institutes of Health (NIH), U.S. Department of Health and Human Services (U01 CA137088; R01 CA059045; U01 CA164930). ASTERISK was funded by a Hospital Clinical Research Program (PHRC-BRD09/C) from the University Hospital Center of Nantes (CHU de Nantes) and supported by the Regional Council of Pays de la Loire, the Groupement des Entreprises Françaises dans la Lutte Contre le Cancer (GEFLUC), the Association Anne de Bretagne Génétique and the Ligue Régionale Contre le Cancer (LRCC). COLO2&3 (Hawai'i Colorectal Cancer Studies 2 & 3) is supported by the NIH (R01 CA60987). DACHS (Darmkrebs: Chancen der Verhutüng durch Screening) was supported by the German Research Council (BR 1704/6-1, BR 1704/6-3, BR 1704/6-4, CH 117/1-1, HO 5117/2-1, HE 5998/2-1, KL 2354/3-1, RO 2270/8-1 and BR 1704/17-1), the Interdisciplinary Research Program of the National Center for Tumor Diseases (NCT), Germany, and the German Federal Ministry of Education and Research (01KH0404, 01ER0814, 01ER0815, 01ER1505A and 01ER1505B). DALS (Diet, Activity and Lifestyle Survey) is supported by the NIH (R01 CA48998 to P.A. Slattery). HPFS (Health Professionals Follow-up Study) is supported by the NIH (P01 CA055075, UM1 CA167552, R01 CA137178, R01 CA151993, R35 CA197735, K07 CA190673, and P50 CA127003), NHS (Nurses' Health Study) by the NIH (R01 CA137178, P01 CA087969, UM1 CA186107, R01 CA151993, R35 CA197735, K07 CA190673, and P50 CA127003) and PHS (Physician's Health Study) by the NIH (R01 CA042182). MEC (Multiethnic Cohort Study) is supported by the NIH (R37 CA54281, P01 CA033619, and R01 CA63464). OFCCR (the Ontario Registry for Studies of Familial Colorectal Cancer) was supported by NIH (U01

CA074783; see CCFR section above), and additional funding toward genetic analyses of OFCCR was supported by a GL2 grant from the Ontario Research Fund, Canadian Institutes of Health Research and a Cancer Risk Evaluation (CaRE) Program grant from the Canadian Cancer Society Research Institute. PLCO (Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial) is supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics and supported by contracts from the Division of Cancer Prevention, National Cancer Institute, NIH, DHHS. PMH-CCFR (Postmenopausal Hormone Study-Colon Cancer Family Registry) is supported by the NIH (R01 CA076366 to P.A. Newcomb). VITAL (VITamins And Lifestyle) is supported by the NIH (K05 CA154337 to E.White). WHI (Women's Health Initiative) is supported by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. X.Wang and E. White were also supported by the NCI (R25 CA094880).

CORECT (the Colorectal Transdisciplinary Study) is supported by the NCI under RFA # CA-09-002 as part of the GAME-ON consortium (US NIH, U19 CA148107) with additional support from the NCI grants (R01 CA81488 and P30 CA14089), the National Human Genome Research Institute at the US NIH (T32 HG000040) and the National Institute of Environmental Health Sciences at the US NIH (T32 ES013678). The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in CORECT, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government or CORECT. CCFR (the Colon Cancer Family Registry) is supported by grant UM1 CA167551 from the US NCI and through cooperative agreements with members of the CCFR and principal investigators of the Australasian Colorectal Cancer Family Registry (US NIH, U01 CA074778 and U01/U24 CA097735), University of South California Consortium Colorectal Cancer Family Registry for Colon Cancer Studies (US NIH, U01/U24 CA074799), the Mayo Clinic Cooperative Family Registry for Colon Cancer Studies (US NIH, U01/U24 CA074800), Ontario Registry for Studies of Familial Colorectal Cancer (US NIH, U01/U24 CA074783), Seattle Colorectal Cancer Family Registry (US NIH, U01/U24 CA074794), and the University of Hawaii Colorectal Cancer Family Registry (US NIH, U01/U24 CA074806). The Colon CFR Illumina GWAS was supported by NCI/NIH grant U01 CA122839 and R01 CA143237 to G Casey. CPSII (the Cancer Prevention Study-II Nutrition Cohort) is funded by the American Cancer Society. MECC was supported by the NIH, U.S. Department of Health and Human Services (R01 CA81488 to SB Gruber and G Rennert). The MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria, GALEON: FIS Intrasalud (PI13/01136). The MCCS was further supported by Australian NHMRC grants 509348, 209057, 251553 and 504711, and by infrastructure provided by Cancer Council Victoria. The NFCCR (Newfoundland Colorectal Cancer Registry) was supported by an Interdisciplinary Health Research Team award from the Canadian Institutes of Health Research (CRT 43821); the NIH, U.S. Department of Health and Human Serivces (U01 CA74783); and National Cancer Institute of Canada grants (18223 and 18226). The Kentucky study was supported by the US NCI R01 CA136726, and the Clinical Investigator Award from Damon Runyon Cancer Research Foundation (CI-8). The Spain study was supported by Instituto de Salud Carlos III, co-funded by FEDER funds –a way to build Europe–

grants PI14-613 and PI09-1286. SEARCH was supported by the Cancer Research UK (C490/A16561). The Sweden-Wolk was supported by grants from the Swedish Research Council/Infrastructure grant, the Swedish Cancer Foundation and Karolinska Institute's Distinguished Professor Award to Alicja Wolk. The ATBC Study is supported by the Intramural Research Program of the U.S. NCI and by U.S. Public Health Service contract HHSN261201500005C from the NCI, Department of Health and Human Services. The ColoCare-heidelberg and the ColoCare-Seattle studies were funded by the US NIH (grants 2P30CA015704-40, R01 CA189184 and U01 CA152756), the Matthias Lackas-Foundation, the German Consortium for Translational Cancer Research, and the EU TRANSCAN initiative. ESTER_VERDI was supported by grants from the Baden-Württemberg Ministry of Science. The work at MSKCC (Memorial Sloan Kettering Cancer Center in New York) was supported by the Robert and Kate Niehaus Center for Inherited Cancer Genomics and the Romeo Milio Foundation. D. Buchanan is also supported by University of Melbourne Research at Melbourne Accelerator Program (R@MAP) and NHMRC R.D. Wright Career Development Fellowship. Dr. M. Song was supported by the American Cancer Society (Grant number MRSG-17-220-01 – NEC to M.S.); by the 2017 AACR-AstraZeneca Fellowship in Immuno-oncology Research (Grant Number 17-40-12-SONG to M.S.); by the U.S. National Institutes of Health (NIH) grants [K99 CA215314 to M.S.].

References

1. Itzkowitz SH, Yio X. Inflammation and cancer IV. Colorectal cancer in inflammatory bowel disease: the role of inflammation. American journal of physiology Gastrointestinal and liver physiology. 2004 Jul;287(1):G7-17. PubMed PMID: 15194558.
2. Tsilidis KK, Branchini C, Guallar E, Helzlsouer KJ, Erlinger TP, Platz EA. C-reactive protein and colorectal cancer risk: a systematic review of prospective studies. International journal of cancer Journal international du cancer. 2008 Sep 1;123(5):1133-40. PubMed PMID: 18528865.
3. Zhou B, Shu B, Yang J, Liu J, Xi T, Xing Y. C-reactive protein, interleukin-6 and the risk of colorectal cancer: a meta-analysis. Cancer causes & control : CCC. 2014 Oct;25(10):1397-405. PubMed PMID: 25053407.
4. Salvioli S, Capri M, Valensin S, Tieri P, Monti D, Ottaviani E, et al. Inflamm-aging, cytokines and aging: state of the art, new hypotheses on the role of mitochondria and new perspectives from systems biology. Current pharmaceutical design. 2006;12(24):3161-71. PubMed PMID: 16918441.
5. de Ferranti S, Rifai N. C-reactive protein and cardiovascular disease: a review of risk prediction and interventions. Clinica chimica acta; international journal of clinical chemistry. 2002 Mar;317(1-2):1-15. PubMed PMID: 11814453.
6. Doll R. Uncovering the effects of smoking: historical perspective. Statistical methods in medical research. 1998 Jun;7(2):87-117. PubMed PMID: 9654637.
7. Wannamethee SG, Lowe GD, Shaper AG, Rumley A, Lennon L, Whincup PH. Associations between cigarette smoking, pipe/cigar smoking, and smoking cessation, and haemostatic and inflammatory markers for cardiovascular disease. European heart journal. 2005 Sep;26(17):1765-73. PubMed PMID: 15817606.
8. Colbert LH, Visser M, Simonsick EM, Tracy RP, Newman AB, Kritchevsky SB, et al. Physical activity, exercise, and inflammatory markers in older adults: findings from the Health, Aging and Body Composition Study. Journal of the American Geriatrics Society. 2004 Jul;52(7):1098-104. PubMed PMID: 15209647.
9. Prasad K. C-reactive protein (CRP)-lowering agents. Cardiovascular drug reviews. 2006 Spring;24(1):33-50. PubMed PMID: 16939632.
10. Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. Nature. 2008 Jul 24;454(7203):436-44. PubMed PMID: 18650914.
11. O'Hanlon DM, Lynch J, Cormican M, Given HF. The acute phase response in breast carcinoma. Anticancer Res. 2002 Mar-Apr;22(2B):1289-93. PubMed PMID: 12168939.
12. Song M, Wu K, Ogino S, Fuchs CS, Giovannucci EL, Chan AT. A prospective study of plasma inflammatory markers and risk of colorectal cancer in men. British journal of cancer. 2013 May 14;108(9):1891-8. PubMed PMID: 23591192. Pubmed Central PMCID: 3658520.
13. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Statistics in medicine. 2008 Apr 15;27(8):1133-63. PubMed PMID: 17886233.
14. Dehghan A, Dupuis J, Barbalic M, Bis JC, Eiriksdottir G, Lu C, et al. Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. Circulation. 2011 Feb 22;123(7):731-8. PubMed PMID: 21300955. Pubmed Central PMCID: 3147232.
15. Nimptsch K, Aleksandrova K, Boeing H, Janke J, Lee YA, Jenab M, et al. Association of CRP genetic variants with blood concentrations of C-reactive protein and colorectal cancer risk. International journal of cancer Journal international du cancer. 2015 Mar 1;136(5):1181-92. PubMed PMID: 25043606.
16. Slattery ML, Curtin K, Poole EM, Duggan DJ, Samowitz WS, Peters U, et al. Genetic variation in C-reactive protein in relation to colon and rectal cancer risk and survival. International journal of cancer Journal international du cancer. 2011 Jun 1;128(11):2726-34. PubMed PMID: 20949557. Pubmed Central PMCID: 3229275.
17. Siemes C, Visser LE, Coebergh JW, Splinter TA, Witteman JC, Uitterlinden AG, et al. C-reactive protein levels, variation in the C-reactive protein gene, and cancer risk: the Rotterdam Study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2006 Nov 20;24(33):5216-22. PubMed PMID: 17114654.
18. Heikkila K, Silander K, Salomaa V, Jousilahti P, Koskinen S, Pukkala E, et al. C-reactive protein-associated genetic variants and cancer risk: findings from FINRISK 1992, FINRISK 1997 and Health 2000 studies. European journal of cancer. 2011 Feb;47(3):404-12. PubMed PMID: 20727736.

19. Allin KH, Nordestgaard BG, Zacho J, Tybjaerg-Hansen A, Bojesen SE. C-reactive protein and the risk of cancer: a mendelian randomization study. Journal of the National Cancer Institute. 2010 Feb 3;102(3):202-6. PubMed PMID: 20056955.
20. Prizment AE, Folsom AR, Dreyfus J, Anderson KE, Visvanathan K, Joshu CE, et al. Plasma C-reactive protein, genetic risk score, and risk of common cancers in the Atherosclerosis Risk in Communities study. Cancer causes & control : CCC. 2013 Dec;24(12):2077-87. PubMed PMID: 24036889. Pubmed Central PMCID: 3836434.
21. Kocarnik JM, Pendergrass SA, Carty CL, Pankow JS, Schumacher FR, Cheng I, et al. Multiancestral analysis of inflammation-related genetic variants and C-reactive protein in the population architecture using genomics and epidemiology study. Circulation Cardiovascular genetics. 2014 Apr;7(2):178-88. PubMed PMID: 24622110. Pubmed Central PMCID: 4104750.
22. Hutter CM, Chang-Claude J, Slattery ML, Pflugeisen BM, Lin Y, Duggan D, et al. Characterization of gene-environment interactions for colorectal cancer susceptibility loci. Cancer Res. 2012 Apr 15;72(8):2036-44. PubMed PMID: 22367214. Pubmed Central PMCID: PMC3374720.
23. Newcomb PA, Baron J, Cotterchio M, Gallinger S, Grove J, Haile R, et al. Colon Cancer Family Registry: an international resource for studies of the genetic epidemiology of colon cancer. Cancer Epidemiol Biomarkers Prev. 2007 Nov;16(11):2331-43. PubMed PMID: 17982118.
24. Peters U, Jiao S, Schumacher FR, Hutter CM, Aragaki AK, Baron JA, et al. Identification of Genetic Susceptibility Loci for Colorectal Tumors in a Genome-Wide Meta-analysis. Gastroenterology. 2013 Apr;144(4):799-807 e24. PubMed PMID: 23266556. Pubmed Central PMCID: 3636812.
25. Peters U, Hutter CM, Hsu L, Schumacher FR, Conti DV, Carlson CS, et al. Meta-analysis of new genome-wide association studies of colorectal cancer risk. Human genetics. 2012 Feb;131(2):217-34. PubMed PMID: 21761138. Pubmed Central PMCID: 3257356.
26. Zanke BW, Greenwood CM, Rangrej J, Kustra R, Tenesa A, Farrington SM, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nature genetics. 2007 Aug;39(8):989-94. PubMed PMID: 17618283.
27. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. Epidemiology. 2014 May;25(3):427-35. PubMed PMID: 24681576. Pubmed Central PMCID: 3981897.
28. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genetic epidemiology. 2013 Nov;37(7):658-65. PubMed PMID: 24114802.
29. Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in Mendelian randomization studies. International journal of epidemiology. 2011 Jun;40(3):755-64. PubMed PMID: 21414999.
30. Baum CF, Schaffer ME, Stillman S. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. Stata J. 2007;7(4):465-506. PubMed PMID: WOS:000254886000002. English.
31. Stock JH, Wright JH, Yogo M. A survey of weak instruments and weak identification in generalized method of moments. J Bus Econ Stat. 2002 Oct;20(4):518-29. PubMed PMID: WOS:000182731800010. English.
32. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. International journal of epidemiology. 2015 Apr;44(2):512-25. PubMed PMID: 26050253. Pubmed Central PMCID: 4469799.
33. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. International journal of epidemiology. 2014 Jun;43(3):922-9. PubMed PMID: 24608958. Pubmed Central PMCID: 4052137.
34. Coussens LM, Werb Z. Inflammation and cancer. Nature. 2002 Dec 19-26;420(6917):860-7. PubMed PMID: 12490959. Pubmed Central PMCID: PMC2803035.
35. Philip M, Rowley DA, Schreiber H. Inflammation as a tumor promoter in cancer induction. Semin Cancer Biol. 2004 Dec;14(6):433-9. PubMed PMID: 15489136.
36. Chen K, Li F, Li J, Cai H, Strom S, Bisello A, et al. Induction of leptin resistance through direct interaction of C-reactive protein with leptin. Nat Med. 2006 Apr;12(4):425-32. PubMed PMID: 16582918.
37. Xu XT, Xu Q, Tong JL, Zhu MM, Huang ML, Ran ZH, et al. Meta-analysis: circulating adiponectin levels and risk of colorectal cancer and adenoma. J Dig Dis. 2011 Aug;12(4):234-44. PubMed PMID: 21791018.
38. Joshu CE, Tsilidis KK, Peskoe SB, Giardiello FM, Dluzniewski PJ, Nelson WG, et al. The association between circulating high-sensitivity C-reactive protein concentration and pathologic measures of colonic inflammation. Cancer Causes Control. 2014 Apr;25(4):409-18. PubMed PMID: 24435936. Pubmed Central PMCID: PMC3951758.

39. Collaboration CRPCHDG, Wensley F, Gao P, Burgess S, Kaptoge S, Di Angelantonio E, et al. Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. BMJ. 2011 Feb 15;342:d548. PubMed PMID: 21325005. Pubmed Central PMCID: PMC3039696.
40. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. N Engl J Med. 2005 Apr 21;352(16):1685-95. PubMed PMID: 15843671.
41. Vansteelandt S, Dukes O, Martinussen T. Survivor bias in Mendelian randomization analysis. Biostatistics. 2017 Sep 27. PubMed PMID: 29028924.