Body size and risk of colorectal cancer molecular defined subtypes and pathways: Mendelian randomization analyses

Nikos Papadimitriou,[a,*] Conghui Qu,[b] Tabitha A. Harrison,[b] Alaina M. Bever,[c,d] Richard M. Martin,[e,f,g] Konstantinos K. Tsilidis,[h,i] Polly A. Newcomb,[b,j] Stephen N. Thibodeau,[k] Christina C. Newton,[l] Caroline Y. Um,[l] Mireia Obón-Santacana,[m,n,o] Victor Moreno,[m,n,o,p] Hermann Brenner,[q,r,s] Marko Mandic,[q,t] Jenny Chang-Claude,[u,v] Michael Hoffmeister,[q] Andrew J. Pellatt,[w] Robert E. Schoen,[x] Sophia Harlid,[y] Shuji Ogino,[z,aa,ab,ac] Tomotaka Ugai,[z,ab] Daniel D. Buchanan,[ad,ae,af] Brigid M. Lynch,[ag,ah] Stephen B. Gruber,[ai] Yin Cao,[aj,ak,al] Li Hsu,[b,am] Jeroen R. Huyghe,[b] Yi Lin,[b] Robert S. Steinfelder,[b] Wei Sun,[b] Bethany Van Guelpen,[y,an] Syed H. Zaidi,[ao] Amanda E. Toland,[ap] Sonja I. Berndt,[aq] Wen-Yi Huang,[aq] Elom K. Aglago,[i] David A. Drew,[ar] Amy J. French,[k] Peter Georgeson,[ad,ae] Marios Giannakis,[ac,as,at] Meredith Hullar,[b] Johnathan A. Nowak,[z] Claire E. Thomas,[b] Loic Le Marchand,[au] Iona Cheng,[av] Steven Gallinger,[aw] Mark A. Jenkins,[ah] Marc J. Gunter,[a,i] Peter T. Campbell,[ax] Ulrike Peters,[b,ay] Mingyang Song,[ar,az] Amanda I. Phipps,[b,ay] and Neil Murphy[a]

[a]Nutrition and Metabolism Branch, International Agency for Research on Cancer, Lyon, France
[b]Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA, USA
[c]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[d]Harvard-MIT Division of Health Sciences and Technology, Harvard Medical School, Boston, MA, USA
[e]MRC Integrative Epidemiology Unit (IEU), Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
[f]Bristol Medical School, Department of Population Health Sciences, University of Bristol, Bristol, UK
[g]National Institute for Health Research (NIHR) Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol, Bristol, UK
[h]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece
[i]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK
[j]School of Public Health, University of Washington, Seattle, WA, USA
[k]Division of Laboratory Genetics, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA
[l]Department of Population Science, American Cancer Society, Atlanta, GA, USA
[m]Unit of Biomarkers and Suceptibility (UBS), Oncology Data Analytics Program (ODAP), Catalan Institute of Oncology (ICO), L'Hospitalet del Llobregat, Barcelona 08908, Spain
[n]ONCOBELL Program, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona 08908, Spain
[o]Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), Madrid 28029, Spain
[p]Department of Clinical Sciences, Faculty of Medicine and Health Sciences and Universitat de Barcelona Institute of Complex Systems (UBICS), University of Barcelona (UB), L'Hospitalet de Llobregat, Barcelona 08908, Spain
[q]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany
[r]Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany
[s]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany
[t]Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany
[u]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[v]University Medical Centre Hamburg-Eppendorf, University Cancer Centre Hamburg (UCCH), Hamburg, Germany
[w]Division of Cancer Medicine, MD Anderson Cancer Center, Houston, TX, USA
[x]Department of Medicine and Epidemiology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA
[y]Department of Radiation Sciences, Oncology Unit, Umeå University, Umeå, Sweden
[z]Program in MPE Molecular Pathological Epidemiology, Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[aa]Cancer Immunology Program, Dana-Farber Harvard Cancer Center, Boston, MA, USA
[ab]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[ac]Broad Institute of MIT and Harvard, Cambridge, MA, USA
[ad]Colorectal Oncogenomics Group, Department of Clinical Pathology, The University of Melbourne, Parkville, Victoria 3010, Australia
[ae]University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, Victoria 3010, Australia
[af]Genetic Medicine and Family Cancer Clinic, The Royal Melbourne Hospital, Parkville, Victoria, Australia
[ag]Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Victoria, Australia
[ah]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Victoria, Australia
[ai]Department of Medical Oncology & Therapeutics Research, City of Hope National Medical Center, Duarte, CA, USA

*Corresponding author. Branch of Nutrition and Metabolism, International Agency for Research on Cancer, 25 avenue Tony Garnier CS 90627 69366 LYON CEDEX 07, France.
E-mail address: papadimitrioun@iarc.who.int (N. Papadimitriou).

[aj]Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine, St. Louis, MO, USA
[ak]Alvin J. Siteman Cancer Center at Barnes-Jewish Hospital and Washington University School of Medicine, St. Louis, MO, USA
[al]Division of Gastroenterology, Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA
[am]Department of Biostatistics, University of Washington, Seattle, WA, USA
[an]Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden
[ao]Ontario Institute for Cancer Research, Toronto, Ontario, Canada
[ap]Departments of Cancer Biology and Genetics and Internal Medicine, Comprehensive Cancer Center, The Ohio State University, Columbus, OH, USA
[aq]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA
[ar]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
[as]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA
[at]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[au]University of Hawaii Cancer Center, Honolulu, HI, USA
[av]Department of Epidemiology and Biostatistics, University of California-San Francisco, San Francisco, CA, USA
[aw]Lunenfeld Tanenbaum Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada
[ax]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA
[ay]Department of Epidemiology, University of Washington, Seattle, WA, USA
[az]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA

eBioMedicine
2024;101: 105010
Published Online xxx
https://doi.org/10.1016/j.ebiom.2024.105010

Summary

Background Obesity has been positively associated with most molecular subtypes of colorectal cancer (CRC); however, the magnitude and the causality of these associations is uncertain.

Methods We used Mendelian randomization (MR) to examine potential causal relationships between body size traits (body mass index [BMI], waist circumference, and body fat percentage) with risks of Jass classification types and individual subtypes of CRC (microsatellite instability [MSI] status, CpG island methylator phenotype [CIMP] status, *BRAF* and *KRAS* mutations). Summary data on tumour markers were obtained from two genetic consortia (CCFR, GECCO).

Findings A 1-standard deviation (SD:5.1 kg/m^2) increment in BMI levels was found to increase risks of Jass type $1_{\text{MSI-high,CIMP-high,}BRAF\text{-mutated,}KRAS\text{-wildtype}}$ (odds ratio [OR]: 2.14, 95% confidence interval [CI]: 1.46, 3.13; p-value = 9×10^{-5}) and Jass type $2_{\text{non-MSI-high,CIMP-high,}BRAF\text{-mutated,}KRAS\text{-wildtype}}$ CRC (OR: 2.20, 95% CI: 1.26, 3.86; p-value = 0.005). The magnitude of these associations was stronger compared with Jass type $4_{\text{non-MSI-high,CIMP-low/negative,}BRAF\text{-wildtype,}KRAS\text{-wildtype}}$ CRC (p-differences: 0.03 and 0.04, respectively). A 1-SD (SD:13.4 cm) increment in waist circumference increased risk of Jass type $3_{\text{non-MSI-high,CIMP-low/negative,}BRAF\text{-wildtype,}KRAS\text{-mutated}}$ (OR 1.73, 95% CI: 1.34, 2.25; p-value = 9×10^{-5}) that was stronger compared with Jass type 4 CRC (p-difference: 0.03). A higher body fat percentage (SD:8.5%) increased risk of Jass type 1 CRC (OR: 2.59, 95% CI: 1.49, 4.48; p-value = 0.001), which was greater than Jass type 4 CRC (p-difference: 0.03).

Interpretation Body size was more strongly linked to the serrated (Jass types 1 and 2) and alternate (Jass type 3) pathways of colorectal carcinogenesis in comparison to the traditional pathway (Jass type 4).

Funding Cancer Research UK, National Institute for Health Research, Medical Research Council, National Institutes of Health, National Cancer Institute, American Institute for Cancer Research, Brigham and Women's Hospital, Prevent Cancer Foundation, Victorian Cancer Agency, Swedish Research Council, Swedish Cancer Society, Region Västerbotten, Knut and Alice Wallenberg Foundation, Lion's Cancer Research Foundation, Insamlingsstiftelsen, Umeå University. Full funding details are provided in acknowledgements.

Keywords: Obesity; Colorectal cancer; Mendelian randomization; Molecular subtypes

Introduction

Colorectal cancer (CRC) is a heterogeneous disease that evolves through multiple pathways defined by genetic, epigenetic and environmental exposure events.[1] This heterogeneity is often characterized by underlying molecular markers, such as: i) microsatellite instability (MSI) resulting from alterations in the DNA mismatch repair system; and ii) CpG island methylator phenotype

Research in context

Evidence before this study

Obesity has been positively associated with most molecular subtypes of colorectal cancer; however, the magnitude and the causality of these associations is uncertain.

Added value of this study

Higher body mass index and body fat percentage were associated with elevated risks of Jass types 1 and 2 cancers (both originating from the serrated pathway) compared to Jass type 4 cancers (traditional adenoma-carcinoma pathway). For waist circumference, we found evidence of a stronger positive effect on Jass type 3 (alternate pathway) cancer compared with Jass type 4. The positive risk estimates for body mass index and Jass types 1 and 2 were stronger than those reported in observational studies.

Implications of all the available evidence

We found that larger body size had differential effects on increasing the risk of colorectal cancer subtypes defined by molecular characteristics. In comparison to the traditional pathway, body size was more strongly linked to the serrated and alternate pathways of colorectal carcinogenesis. The current results suggest also that the impact of elevated adiposity on colorectal cancer risk for the serrated pathway may have previously been underestimated.

(CIMP), which results from hypermethylation of promoter CpG island sites that inactivates several tumour suppressor or other tumour-related genes. Somatic mutations in the *BRAF* and *KRAS* oncogenes are other tumour markers that are relevant for CRC etiology and prognosis.[2,3] However, these tumour markers capture broader tumour phenotypes and can be of greater utility to clinical research, when combined to proxy specific pathways (i.e., Jass types) of colorectal carcinogenesis[4] (Table 1).

Obesity is an established causal risk factor for CRC.[5–7] However, there is relatively less evidence on the effect of obesity and other body size traits with molecular subtypes of CRC. A recent pooled observational study of 11,872 CRC cases and 11,013 controls in the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO) and the Colon Cancer Family Registry (CCFR) found consistent positive associations of body mass index (BMI) across CRC tumour subtypes defined by individual molecular markers.[8] Similarly, positive associations between BMI and CRC risk were observed for Jass type classifications 1 to 4 indicating a role for obesity for most major pathways of CRC development. In contrast, a null association was found for Jass type 5, suggesting that BMI may not be a risk factor for the development of CRC amongst individuals with familial-like/Lynch syndrome.[8]

An important limitation of the current literature on the role of obesity on molecular subtypes of CRC is that it is based on evidence from traditional observational studies.[8–12] Causal inference is therefore limited by the inherent biases of these epidemiological study designs, such as residual confounding, measurement error, and reverse causality.[13,14] In addition, these observational studies measured participants body size once in middle age, so any life course effects of adiposity in the development of CRC molecular defined subtypes and pathways are uncertain.[15] Furthermore, most prior studies only investigated associations for BMI so how central adiposity (e.g., waist circumference) or other indicators of overall adiposity (e.g., body fat percentage) are associated with molecularly defined subtypes CRC is largely unknown.

Mendelian randomization (MR) is an instrumental variable approach appraising causality using observational data. MR uses germline genetic variants as proxies (or instrumental variables) for exposures of interest to allow causal inference between a given exposure and outcome.[16] Unlike traditional observational analyses, MR analyses should be less susceptible to confounding and reverse causality due to the random assortment of alleles at meiosis and germline genetic variants being fixed at conception, and thus unaffected by the disease process.[17,18] Additionally, MR estimates may better

Jass type	Components
Jass type 1	MSI-high, CIMP-high, *BRAF*-mutated, *KRAS*-wildtype
Jass type 2	non-MSI-high, CIMP-high, *BRAF*-mutated, *KRAS*-wildtype
Jass type 3	non-MSI-high, CIMP-low/negative, *BRAF*-wildtype, *KRAS*-mutated
Jass type 4	non-MSI-high, CIMP-low/negative, *BRAF*-wildtype, *KRAS*-wildtype
Jass type 5	MSI-high, CIMP-low/negative, *BRAF*-wildtype, *KRAS*-wildtype

CIMP, CpG island methylator phenotype; MSI, microsatellite instability.

Table 1: **Definition of Jass groups.**

reflect the accumulated exposure to adiposity across the life course (given that adiposity is proxied by germline genetics), while correcting for the exposure measurement error related to a single time point body size measurement.[15,19]

We used two-sample MR to assess whether BMI, waist circumference, and body fat percentage are causally associated with risks of individual molecularly defined subtypes of CRC (MSI status, CIMP status, *BRAF* and *KRAS* mutations) and Jass classification types (Table 1). Genetic variants associated with body size traits were identified from recent genome-wide association studies (GWAS). We then examined how these genetic variants related to CRC using GWAS data from two genetic consortia.[20–23]

Methods

Data on body size traits

Summary-level data on BMI was obtained from a recent meta-analysis of three GWAS of up to 1.1 million individuals of European ancestry in UK Biobank, Million Veteran Program, and GIANT consortium.[23] This study identified 1253 genome-wide-significant SNPs (p-value $< 5 \times 10^{-8}$) using a linkage disequilibrium (LD) R^2 of <0.01. Summary data on waist circumference and body fat percentage were collected from UK Biobank GWASs (up to 462,166 participants) through the MRC-IEU consortium platform (https://gwas.mrcieu.ac.uk/) using the codes 'ukb-b-9405' and 'ukb-b-8909'. Using a similar LD cut-off point of ≤0.01, 569 and 671 independent and genome-wide-significant SNPs were retained for waist circumference and body fat percentage, respectively (Supplemental Tables S1–S3).

Data on colorectal cancer

Summary data were drawn from a meta-analysis of the Colon Cancer Family Registry (CCFR), and the Genetics and Epidemiology of Colorectal Cancer (GECCO) consortia within participants of European descent.[24,25] The current study includes 10,472 controls and 8178 CRC cases of European ancestry (99.7% of total sample) with available information on the four molecular markers from 10 studies within the two consortia (Table 2, Supplemental Tables S4–S22). Polytomous regressions were performed for all Jass types (Table 1) and individual tumour markers adjusting for age at diagnosis or recruitment, sex, GWAS set, and 3 principal components to adjust for underlying population structures. Logistic regression was performed for case-only analysis to compare mutated and wildtype cases with the same covariates.

Tumour marker data–molecular subtyping and pathways

The process of data collection and harmonization of the tumour marker data has been previously described.[24,25]

Molecular subtype	Men	Women	Total
MS			
MSI high	505	660	1165
non-MSI-high	3520	2985	6505
CIMP			
CIMP-high	383	617	1000
CIMP-low/negative	2951	2311	5262
KRAS			
Mutation	1147	1018	2165
Wildtype	2367	2107	4474
BRAF			
Mutation	317	555	872
Wildtype	3526	2900	6426
Jass groups			
Type 1	84	258	342
Type 2	51	97	148
Type 3	739	682	1421
Type 4	1446	1023	2469
Type 5	104	83	187
Controls	5180	5292	10,472

CIMP, CpG island methylator phenotype; MSI, microsatellite instability. Type 1 (MSI-high, CIMP-high, *BRAF*-mutated, *KRAS*-wildtype). Type 2 (non-MSI-high, CIMP-high, *BRAF*-mutated, *KRAS*-wildtype). Type 3 (non-MSI-high, CIMP-low/negative, *BRAF*-wildtype, *KRAS*-mutated). Type 4 (non-MSI-high, CIMP-low/negative, *BRAF*-wildtype, *KRAS*-wildtype). Type 5 (MSI-high, CIMP-low/negative, *BRAF*-wildtype, *KRAS*-wildtype).

Table 2: **Sample size by molecular subtype status and sex.**

In summary, MSI testing was primarily conducted using polymerase chain reaction (PCR) following accepted guidelines (CCFR, CPS-II, MCCS, NHS)[26] with >4 interpretable markers typically required to classify tumours (Supplemental Materials and Supplemental Table S23). DACHS used a mononucleotide panel of 3 markers that has high concordance with the Bethesda Consensus Panel for the detection of MSI-high status.[27] Tumours were classified as MSI-high if at least 30% of the markers showed instability. Other studies used immunohistochemistry for MSH2, MLH1, MSH6, and PMS2 and loss of any of those proteins was classified as "mismatch repair deficiency", which very highly correlates with MSI-high (NSHDS, EPIC-Sweden and subsets of CCFR and MCCS).[28,29]

CIMP status was determined using methylation analyses (Supplemental Materials and Supplemental Table S24). Briefly, MethyLight was used in the CCFR, CPS-II, HPFS, MCCS, NSHDS, EPIC Sweden and NHS to determine CIMP status. CPS-II, HPFS, NSHDS, EPIC Sweden, and NHS used an 8-gene panel; CCFR and MCCS used a 5-gene panel. DACHS determined CIMP status using a different 5-gene panel.[30] For the current analysis two CIMP categories were created: CIMP-high and CIMP-low/negative.

BRAF and *KRAS* mutations were assessed using PCR, sequencing, and immunohistochemistry (Supplemental Materials). Most studies evaluated *BRAF*

via c.1799T > A (p.V600E) mutations in exon 15 and *KRAS* via mutations in codons 12 and 13.

We defined 5 combined colorectal tumour subtypes consistent with the Jass classifications (Table 1).[4,31]

Statistical power

Statistical power (*a priori*) was calculated using an online tool at http://cnsgenomics.com/shiny/mRnd/.[32] The selected SNPs explained approximately 8.8%, 4.7%, and 4.0% of the variability for BMI, waist circumference, and body fat percentage, respectively. Under the scenario of a type 1 error of 5%, the expected odds ratio (OR) per 1 standard deviation (SD) needed to have adequate statistical power of 80% ranged from 1.22 for BMI for the Jass type 4 CRC to 2.40 for body fat percentage and Jass type 2 CRC. Supplemental Table S25 details the minimum ORs needed for 80% power for the different exposures and Jass types and individual subtypes of CRC.

Statistical analysis

Given the large number of SNPs included in the current study and the likelihood some of them being pleiotropic, the random-effects IVW method was used to adjust for heterogeneity due to horizontal pleiotropy. All results correspond to an OR per 1-SD increment in BMI (SD: 5.1 kg/m^2), waist circumference (SD: 13.4 cm), and body fat percentage (SD: 8.5%). We investigated the effect of body size on CRC Jass types (primary analysis) and individual molecular subtypes (secondary analysis). Heterogeneity of the causal estimates in the pathway analysis was measured by conducting an additional MR analysis using data derived from the cancer cases only using the Jass type 4 (traditional) pathway as a reference. Similarly, we evaluated the heterogeneity for each molecular marker using non-MSI-high, CIMP-low/negative, *BRAF*-wildtype, or *KRAS*-wildtype status as the reference categories for each marker. We used false discovery rate (FDR) corrected p-values to assess statistical significance for the cancer case-only analyses across CRC Jass type pathways and individual markers.[33]

We also conducted multivariable a MR analysis adjusting the genetic instruments of the three body size phenotypes for cigarette smoking and alcohol consumption, two major risk factors associated with CRC.[5,34] The data for lifetime smoking were obtained from a recent GWAS and MR study on causal effects of lifetime smoking on risk for depression and schizophrenia.[35] Data on alcohol consumption (drinks per week) were drawn from a GWAS of 2.6 million individuals[36] (Supplemental Tables S26–S28).

MR assumption testing and sensitivity analyses

For the causal estimates to be valid, there are three main assumptions that must hold: 1) the genetic instrument is strongly associated with the body size traits; 2) the genetic instrument is not associated with any potential confounder of the exposure (body size)–outcome (cancer) association; and 3) the genetic instrument does not affect outcome (cancer) independently of exposure (body size) (i.e., exclusion of horizontal pleiotropy) (Supplemental Fig. S1). The strength of each genetic instrument can be evaluated through the F-statistic using the following formula: $F = R^2(N-2)/(1-R^2)$, where R^2 is the proportion of the variability explained by each instrument for each adiposity trait and N the sample size of the GWAS for the SNP-adiposity trait association.[37] The R^2 was calculated using the following formula: $2 \times EAF \times (1-EAF) \times beta^2$, where EAF is the effect allele frequency, beta is the estimated genetic effect on the exposure (body size trait) and N is the sample size of the GWAS for the SNP-exposure association. Several sensitivity analyses were done to identify and correct for the presence of pleiotropy in the main results. Cochran's Q was computed to quantify heterogeneity across the individual causal effects, with a p-value ≤ 0.05 indicating the presence of pleiotropy.[38,39] MR-Egger regression was applied where deviations from zero for the intercept term denote presence of horizontal pleiotropic effects across the genetic variants and the slope provides valid estimates when the pleiotropic effects of the genetic variants are independent from the genetic associations with the exposure.[40,41] Causal estimates were also computed using the weighted-median approach which can give valid MR estimates under the presence of horizontal pleiotropy when up to 50% of the included instruments are invalid.[42] Signals that appeared to be outliers were removed using the Radial method, a simulation based approach and then the analyses were rerun after excluding any outlying variants.[43] Finally, to satisfy the third MR assumption, we removed cancer-related SNPs ($p < 1 \times 10^{-5}$) from the analyses.

All the analyses were conducted using the ieugwasr, Mendelian Randomization, Two Sample MR, and Radial MR packages and the R programming language.[20,43,44] Reporting guidelines for MR studies were followed.[45,46]

Ethics

All analyses were conducted using summary-level data generated by previous studies. All participants provided written informed consent, and each study was approved by the relevant research ethics committee or institutional review board.

Role of funders

The funders had no role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication.

Results

Jass classification CRC types

A 1-SD increment in BMI levels increased the risks of Jass type 1 (OR per 1-SD: 2.14, 95% confidence interval

[CI]: 1.46, 3.13; p-value = 9×10^{-5}) and Jass type 2 CRC (OR per 1-SD:2.20, 95% CI: 1.26, 3.86; p-value = 0.005) while a positive effect estimate was observed for Jass type 4 (OR per 1-SD:1.17, 95% CI: 1.00, 1.36; p-value = 0.05). After FDR correction, the estimates for Jass types 1 and 2 were stronger compared to Jass type 4 CRC (Fig. 1; p-differences: 0.03 and 0.04, respectively).

Positive effects of varying strengths were observed between waist circumference and all CRC Jass types reaching the threshold of statistical significance for type 1 (OR per 1-SD: 1.77, 95% CI: 1.06, 2.97; p-value = 0.03) and type 3 CRC (OR per 1-SD: 1.73, 95% CI: 1.34, 2.25; p-value = 3×10^{-5}) only. In the case only analysis, compared to the type 4 reference group, statistically significant heterogeneity was found for type 3 CRC only (Fig. 1; p-difference: 0.03). A 1-SD increment in body fat percentage level increased risk for Jass type 1 CRC and which was larger than the effect for type 4 CRC (OR per 1-SD:2.59, 95% CI: 1.49, 4.48; p-difference: 0.03).

Individual molecular CRC markers

Higher BMI, waist circumference, and body fat percentage levels were similarly positively related with CRC tumour subtypes defined by individual molecular markers (Fig. 2, Supplemental Table S29). The individual marker results seemed to be more consistent across the three exposures with the effects being stronger for MSI-high, CIMP-high and *BRAF*-mutated tumours than the counterparts. In the case only analysis, no statistically significant heterogeneity was found for the relationship between body size traits and each individual molecular CRC marker (Supplemental Table S29).

Multivariable MR

The multivariable MR results in which we adjusted for alcohol consumption and number of cigarettes are presented in Supplemental Table S30. A similar pattern of results was found for analyses of BMI, waist circumference, and body fat percentage with Jass classification CRC types and individual markers after adjustment for alcohol consumption and number of cigarettes.

Sensitivity analyses

Based on the F-statistics, the genetic instruments were deemed strong (F-statistic all ≥16) (Supplemental Tables S1–S3). Little evidence of directional pleiotropy was observed based on the MR-Egger's test (MR-Egger intercept p-values > 0.05) and the effect estimates from the lower powered MR Egger regression model and the weighted-median approach were generally consistent in direction and magnitude to the IVW models except of the analysis of BMI and Jass type 4 CRC where none of the two methods found similar to the IVW estimates (Supplemental Table S29). The radial method identified only a few outlying SNPs, however, their exclusion had little impact on the original effect estimates (Supplemental Table S29).

Discussion

This MR study investigated the effects of body size traits with risk of CRC molecular defined pathways and subtypes. We found that higher BMI and body fat percentage elevated risks of Jass types 1 and 2, suggesting that overall body adiposity is a stronger risk factor for CRC risk originating from the serrated pathway than for Jass type 4 CRC (traditional adenoma-carcinoma pathway). For waist circumference, we found evidence of a stronger positive effect on Jass type 3 (alternate pathway) CRC compared with Jass type 4 CRC. Interestingly, these positive MR risk estimates for Jass types

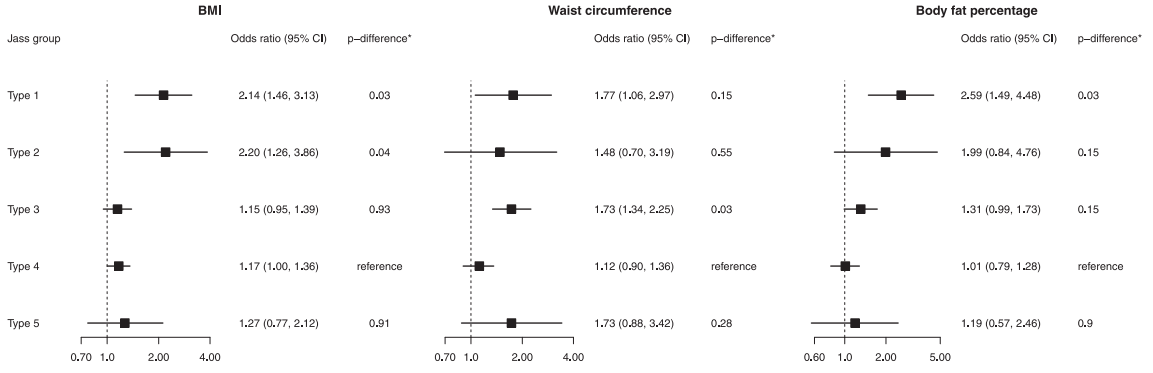

***Fig. 1:* Association between BMI, waist circumference, and body fat percentage and Jass classified types of colorectal cancer.** All estimates correspond to a 1 SD change in the levels of the exposures. Type 1 (MSI-high, CIMP-high, *BRAF*-mutated, *KRAS*-wildtype); Type 2 (non-MSI-high, CIMP-high, *BRAF*-mutated, *KRAS*-wildtype); Type 3 (non-MSI-high, CIMP-low/negative, *BRAF*-wildtype, *KRAS*-mutated); Type 4 (non-MSI-high, CIMP-low/negative, *BRAF*-wildtype, *KRAS*-wildtype); Type 5 (MSI-high, CIMP-low/negative, *BRAF*-wildtype, *KRAS*-wildtype). The odds ratios correspond to the associations of BMI, waist circumference, and fat percentage with each Jass colorectal cancer type compared to the controls. The p-difference* corresponds to the false discovery rate (FDR) adjusted p values for heterogeneity comparing Jass types 1,2,3, and 5 with Jass colorectal cancer type 4.

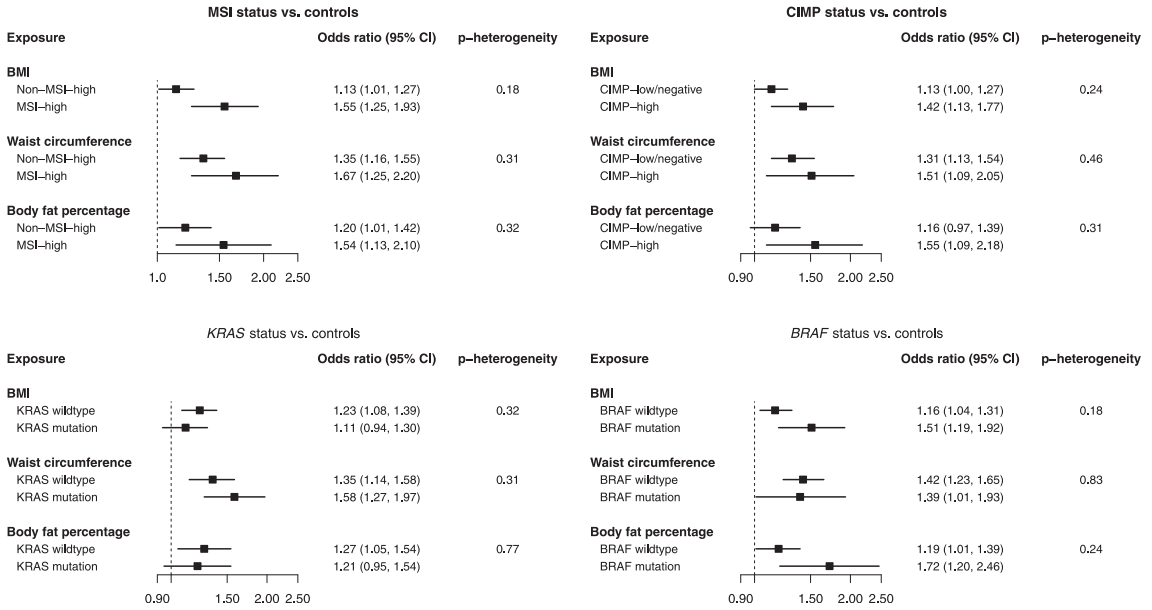

***Fig. 2:* Associations of BMI, waist circumference, and body fat percentage with cancer risk based on the individual molecular markers.** All estimates correspond to a 1 SD change in the levels of the exposures.

1–2 were stronger than those previously reported from prior observational studies, suggesting that the impact of elevated adiposity on CRC risk may have previously been underestimated.

We found that higher levels of body size traits were associated with a higher CRC risk originating from the serrated pathway (Jass types 1 and 2). Consistent with these results, a US cohort analysis reported that BMI was more strongly associated with serrated polyps (OR BMI ≥ 35 kg/m^2 vs < 25 kg/m^2: 1.34, 95% CI: 1.23–1.46) than conventional adenomas (OR BMI ≥ 35 kg/m^2 vs <25 kg/m^2: 1.07, 95% CI: 0.98, 1.17) (p-heterogeneity < 0.001).[47] Our recent pooled observational study also found positive associations between BMI and both Jass type 1 CRC (OR per 5 kg/m^2: 1.24, 95% CI: 1.12, 1.36) and Jass type 2 CRC (OR per 5 kg/m^2: 1.33, 95% CI: 1.17, 1.52) with the effect estimates being stronger than those for Jass type 4 (OR per 5 kg/m^2: 1.18, 95% CI: 1.13, 1.24).[8] However, our MR estimates were stronger than those observed in our previous observational study, which might be a consequence of the MR estimates better reflecting accumulated exposure to adiposity across the life course, the avoidance of biases like reverse causation and residual confounding in an MR study, and also correcting for exposure measurement error related to a body size measurement single time point.[15,19]

Mechanisms underlying the stronger positive effects of adiposity on serrated pathway CRC are unclear. Obesity may promote colorectal tumourigenesis through chronic inflammation which may play a greater role in neoplastic progression for serrated polyps than for conventional adenomas as it was found that the expression levels of inflammatory proteins, COX-2, IL-4 and TNF-α were higher in serrated polyps.[48] Additionally, there is growing evidence implicating the gut microbiota in colorectal carcinogenesis and *Fusobacterium nucleatum* has been consistently associated with greater CRC risk.[49,50] *Fusobacterium nucleatum* levels have been found to be higher among people with obesity and are more strongly associated with serrated pathway than conventional pathway CRC.[51–54] Additional studies are needed to gain a better understanding of the biological pathways underlying the strong positive relationship found between adiposity and the serrated CRC pathway.

Positive effect estimates were observed for all three body size traits and Jass type 3 CRC (alternate pathway) with waist circumference showing the strongest effect among the exposure variables. Our recent pooled observational analysis also found positive associations between BMI and Jass type 3 CRC (OR per 5 kg/m^2:1.15; 95% CI: 1.09, 1.22).[8] *KRAS* mutation is the characteristic that separates this cancer type from Jass type 4. In our *KRAS* mutation subtype analysis we did not find any heterogeneity in the effects of body size traits according to mutation status. However, it has been found that silencing of the *MGMT* gene is associated with *KRAS*-mutated and CIMP-low status CRC and therefore *MGMT* methylation may be another characteristic of this pathway which requires further investigation in relation to obesity.[4]

For the Jass type 4 defined CRC (traditional adenoma-carcinoma pathway), we observed positive

effect estimates for BMI and waist circumference (ORs of 1.17 and 1.12 per SD, respectively) that did not pass the conventional threshold of statistical significance. It is important to note that our study was not sufficiently powered to detect ORs of such magnitude for Jass type 4 CRC. However, our MR effect estimate for BMI was of similar magnitude to the Jass type 4 CRC result from our recent pooled observational study (OR per 5 kg/m2:1.18; 95% CI: 1.13, 1.24).[8] Together these results provide support for a positive association between BMI and CRC development through the traditional adenoma-carcinoma pathway.

We observed consistently positive but imprecise and statistically non-significant effect estimates for body size traits and risk of CRC Jass type 5 tumours (considered familial-like/Lynch syndrome), with the largest effect estimate observed for waist circumference. In our prior pooled observational study we found no association between BMI and CRC Jass type 5 tumours.[8] Overall, the MR and observational evidence provide limited support for a positive association between overall adiposity and CRC Jass type 5 tumours. However, the non-significant positive effect estimate we observed for waist circumference requires further investigation as it may suggest that central adiposity is a risk factor for CRC Jass type 5 tumours.

Our study has several strengths. Compared to prior observational evidence, our MR analyses are less prone to bias from reverse causation and residual confounding.[17,18] Large-scale summary genetic data from UK Biobank, the Million Veteran Program, and GIANT consortium was used to collect the genome-wide significant SNPs for the three exposures of interest resulting in strong genetic instruments. Independent datasets for the exposures and outcome phenotypes were used, thus avoiding potential bias due to sample overlap.[55] Finally, unlike prior observational studies that focused on BMI only, we were able to examine relationship also for body fat percentage and waist circumference.

A limitation of our study was the low CRC case numbers for some of the Jass type groups thus limiting our power to identify potential relationships. For the same reason we did not conduct any sex-specific analyses; however, in our prior observational study we found limited evidence of heterogeneity by sex for the associations between BMI and CRC molecular subtypes.[8] Nevertheless, larger studies are needed to verify the current results, include additional rarer molecular subtypes, and investigate potential heterogeneity by sex. Given the large number of instruments included in the current study, we cannot exclude potential pleiotropic effects, meaning that the instruments might affect cancer risk through other pathways outside their effects on body size; however, our results were consistent according to the comprehensive sensitivity analyses we conducted. No inference could be made regarding the effects of cigarettes smoking and alcohol consumption in the multivariable MR analysis. The complete summary GWAS data on BMI was not publicly available and we were not able to conduct a unified multivariable MR analysis merging the genetic instruments of each of our exposures with the confounding variables. As previously mentioned, the measurement of tumour markers differed slightly across studies which may have introduced some heterogeneity in the tumour marker classifications. Stratified analyses by contributing study or by CRC subtype classification methods could potentially help us quantify this heterogeneity if present. However, the use of GWAS summary statistics, precluded us conducting this type of analysis which would also be prone to loss of statistical power due to the smaller sample size. Additionally, a high correlation between the different tumour marker measurement methods has been reported in the literature of colorectal cancer research.[25,27–29] Furthermore, another recent study using data from the same consortia as the current investigation found a high level of consistency between the classification of MSI, BRAF, and KRAS mutation status for participants with existing tumour marker and newly centrally generated tumour sequencing data.[25] More specifically, the tumor classifications from the two approaches were highly concordant with 98.6% concordance for the 1534 individuals with information on MSI status, 91.4% concordance for the 1696 individuals with KRAS mutation data, and 93.1% concordance for the 1738 individuals with BRAF mutation data from both sources.[25] Finally, the results cannot be generalised to diverse populations due to the lack of ancestral diversity in the genetic data used for our analyses.

In summary, using MR we found that larger body size had differential effects on increasing the risk of CRC subtypes defined by molecular characteristics. In comparison to the traditional pathway (Jass type 4), body size was more strongly linked to the serrated (Jass types 1 and 2) and alternate (Jass type 3) pathways of colorectal carcinogenesis. The positive risk estimates for BMI and Jass types 1 and 2 were stronger than those reported in observational studies, suggesting that the impact of elevated adiposity on CRC risk for the serrated CRC pathway may have previously been underestimated.

Contributors

Conceptualization: NP, NM, PTC, MS, UP.
Data curation: UP, TAH, CQ, AIP.
Data verification: UP, TAH, CQ, AIP.
Methodology: NP, NM, PTC, MS, UP.
Investigation: All authors.
Writing—original draft: NP.
Writing—review & editing: All authors.
All authors read and approved the final version of the manuscript.

Data sharing statement

All data are available in the Supplemental data files.

Declaration of interests

AB has received a Scholar-in-Training Award from AACR for 2023. VM has received grants from Instituto de Salud Carlos III and FC AECC.

BML has received support from Uppsala University as a part of a presentation in the Svedberg Seminar Series, and she was the president of the Australasian Epidemiological Association (2020–2023). BVG has received support grant from World Cancer Research Fund for a separate project within the same research field. AET has received an editorial board fee from NIH PDQ Cancer Genetics. MG has received research funding from Janssen and Servier, consulting fees from Nerviano Medical Sciences, Chromacode, and AstraZeneca and support to attend a meeting from Dava Oncology. JN has received research funding from Natera Inc and consulting fees from Leica Biosciences. MAJ has received an NIH funding. The remaining authors declare no competing interests.

Acknowledgements

This study was supported by Cancer Research UK (C18281/A29019) and Cancer Research UK (PPRCPJT\100005; Dr Tsilidis). RMM is a National Institute for Health Research Senior Investigator (NIHR202411). RMM is supported by a Cancer Research UK 25 (C18281/A29019) programme grant (the Integrative Cancer Epidemiology Programme). RMM is also supported by the NIHR Bristol Biomedical Research Centre which is funded by the NIHR (BRC-1215-20011) and is a partnership between University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. RMM is affiliated with the Medical Research Council Integrative Epidemiology Unit at the University of Bristol which is supported by the Medical Research Council (MC_UU_00011/1, MC_UU_00011/3, MC_UU_00011/6, and MC_UU_00011/4) and the University of Bristol. AB has received support from the National Cancer Institute (F30CA265012). SO has received a grant from National Institutes of Health. TU has received support from National Institutes of Health/National Cancer Institute (R50CA274122), American Institute for Cancer Research Investigator-Initiated Research Grant, Brigham and Women's Hospital Faculty Career Development Award, and Prevent Cancer Foundation Grant. BML has received support from the Victorian Cancer Agency (MCRF18005). BVG has received support from Swedish Research Council, Swedish Cancer Society, Region Västerbotten, Knut and Alice Wallenberg Foundation, Lion's Cancer Research Foundation and, Insamlingsstiftelsen, Umeå University. AET has received an R01 National Institutes of Health/National Cancer Institute funding. SB has received support from National Cancer Institute, NIH. DAD has received grants from the National Institutes of Health.

Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO): National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services (U01 CA137088, R01 CA059045, U01 CA164930, R21 CA191312, R01 CA244588, R01 CA201407). Genotyping/Sequencing services were provided by the Center for Inherited Disease Research (CIDR) contract number HHSN2682017000006I and HHSN2682012000008I. This research was funded in part through the NIH/NCI Cancer Center Support Grant P30 CA015704. Scientific Computing Infrastructure at Fred Hutch funded by ORIP grant S10OD028685.

The Colon Cancer Family Registry (CCFR, www.coloncfr.org) is supported in part by funding from the National Cancer Institute (NCI), National Institutes of Health (NIH) (award U01 CA167551). Support for case ascertainment was provided in part from the Surveillance, Epidemiology, and End Results (SEER) Program and the following U.S. state cancer registries: AZ, CO, MN, NC, NH; and by the Victoria Cancer Registry (Australia) and Ontario Cancer Registry (Canada). The CCFR Set-1 (Illumina 1M/1M-Duo) and Set-2 (Illumina Omni1-Quad) scans were supported by NIH awards U01 CA122839 and R01 CA143237 (to GC). The CCFR Set-3 (Affymetrix Axiom CORECT Set array) was supported by NIH award U19 CA148107 and R01 CA81488 (to SBG). The CCFR Set-4 (Illumina OncoArray 600K SNP array) was supported by NIH award U19 CA148107 (to SBG) and by the Center for Inherited Disease Research (CIDR), which is funded by the NIH to the Johns Hopkins University, contract number HHSN2682012000008I. Additional funding for the OFCCR/ARCTIC was through award GL201-043 from the Ontario Research Foundation (to BWZ), award 112746 from the Canadian Institutes of Health Research (to TJH), through a Cancer Risk Evaluation (CaRE) Program grant from the Canadian Cancer Society (to SG), and through generous support from the Ontario Ministry of Research and Innovation. The SFCCR Illumina HumanCytoSNP array was supported in part through NCI/NIH awards U01/U24 CA074794 and R01 CA076366 (to PAN). The content of this manuscript does not necessarily reflect the views or policies of the NCI, NIH or any of the collaborating centers in the Colon Cancer Family Registry (CCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government, any cancer registry, or the CCFR.

CPS-II: The American Cancer Society funds the creation, maintenance, and updating of the Cancer Prevention Study-II (CPS-II) cohort. The study protocol was approved by the institutional review boards of Emory University, and those of participating registries as required.

DACHS: This work 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, 01ER1505B, and 01KD2104A).

DALS: National Institutes of Health (R01 CA048998 to M. L. Slattery).

EDRN: This work is funded and supported by the NCI, EDRN Grant (U01-CA152753).

EPIC: 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). EPIC Sweden is supported by Swedish Cancer Society, Swedish Research Council and Region Skåne and Region Västerbotten (Sweden).

Harvard cohorts: HPFS is supported by the National Institutes of Health (P01 CA055075, UM1 CA167552, U01 CA167552, R01 CA137178, R01 CA151993, and R35 CA197735), NHS by the National Institutes of Health (P01 CA087969, UM1 CA186107, R01 CA137178, R01 CA151993, and R35 CA197735).

MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 509348, 209057, 251553 and 504711 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry (VCR) and the Australian Institute of Health and Welfare (AIHW), including the National Death Index and the Australian Cancer Database.

NSHDS: The research was supported by Biobank Sweden through funding from the Swedish Research Council (VR 2017-00650, VR 2017-01737), the Swedish Cancer Society (CAN 2017/581), Region Västerbotten (VLL-841671, VLL-833291), Knut and Alice Wallenberg Foundation (VLL-765961), and the Lion's Cancer Research Foundation (several grants) and Insamlingsstiftelsen, both at Umeå University.

CCFR: The Colon CFR graciously thanks the generous contributions of their study participants, dedication of study staff, and the financial support from the U.S. National Cancer Institute, without which this important registry would not exist. The authors would like to thank the study participants and staff of the Seattle Colon Cancer Family Registry and the Hormones and Colon Cancer study (CORE Studies).

CPS-II: The authors express sincere appreciation to all Cancer Prevention Study-II participants, and to each member of the study and biospecimen management group. The authors would like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention's National Program of Cancer Registries and cancer registries supported by the National Cancer Institute's Surveillance Epidemiology and End Results Program. The authors assume full responsibility for all analyses and interpretation of results. The views expressed here are those of the authors and do not necessarily represent the American Cancer Society or the American Cancer Society—Cancer Action Network.

DACHS: We thank all participants and cooperating clinicians, and everyone who provided excellent technical assistance.

EDRN: We acknowledge all contributors to the development of the resource at University of Pittsburgh School of Medicine, Department of Gastroenterology, Department of Pathology, Hepatology and Nutrition and Biomedical Informatics.

EPIC: Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

Harvard cohorts: The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required. We acknowledge Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital as home of the NHS. The authors would like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention's National Program of Cancer Registries (NPCR) and/or the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. Central registries may also be supported by state agencies, universities, and cancer centers. Participating central cancer registries include the following: Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Indiana, Iowa, Kentucky, Louisiana, Massachusetts, Maine, Maryland, Michigan, Mississippi, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, Seattle SEER Registry, South Carolina, Tennessee, Texas, Utah, Virginia, West Virginia, Wyoming. The authors assume full responsibility for analyses and interpretation of these data.

NSHDS investigators thank the Västerbotten Intervention Programme, the Northern Sweden MONICA study, the Biobank Research Unit at Umeå University and Biobanken Norr at Region Västerbotten for providing data and samples and acknowledge the contribution from Biobank Sweden, supported by the Swedish Research Council.

Disclaimer: Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2024.105010.

References

1 Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology*. 2007;50(1):113–130.
2 Nazemalhosseini Mojarad E, Kuppen PJ, Aghdaei HA, Zali MR. The CpG island methylator phenotype (CIMP) in colorectal cancer. *Gastroenterol Hepatol Bed Bench*. 2013;6(3):120–128.
3 Samadder NJ, Vierkant RA, Tillmans LS, et al. Associations between colorectal cancer molecular markers and pathways with clinicopathologic features in older women. *Gastroenterology*. 2013;145(2):348–356.e1-2.
4 Kocarnik JM, Shiovitz S, Phipps AI. Molecular phenotypes of colorectal cancer and potential clinical applications. *Gastroenterol Rep (Oxf)*. 2015;3(4):269–276.
5 World Cancer Research Fund/American Institute for Cancer Research. *Continuous Update Project Expert Report. Diet, nutrition, physical activity and colorectal cancer*. 2018.
6 Bull CJ, Bell JA, Murphy N, et al. Adiposity, metabolites, and colorectal cancer risk: mendelian randomization study. *BMC Med*. 2020;18(1):396.
7 Lauby-Secretan B, Scoccianti C, Loomis D, et al. Body fatness and cancer–viewpoint of the IARC working group. *N Engl J Med*. 2016;375(8):794–798.
8 Murphy N, Newton CC, Song M, et al. Body mass index and molecular subtypes of colorectal cancer. *J Natl Cancer Inst*. 2022;115(2):165–173.
9 Carr PR, Alwers E, Bienert S, et al. Lifestyle factors and risk of sporadic colorectal cancer by microsatellite instability status: a systematic review and meta-analyses. *Ann Oncol*. 2018;29(4):825–834.
10 Weisenberger DJ, Levine AJ, Long TI, et al. Association of the colorectal CpG island methylator phenotype with molecular features, risk factors, and family history. *Cancer Epidemiol Biomarkers Prev*. 2015;24(3):512–519.
11 Brandstedt J, Wangefjord S, Nodin B, et al. Associations of anthropometric factors with KRAS and BRAF mutation status of primary colorectal cancer in men and women: a cohort study. *PLoS One*. 2014;9(2):e98964.
12 Carr PR, Amitay EL, Jansen L, et al. Association of BMI and major molecular pathological markers of colorectal cancer in men and women. *Am J Clin Nutr*. 2020;111(3):562–569.
13 Lawlor DA, Davey Smith G, Kundu D, Bruckdorfer KR, Ebrahim S. Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? *Lancet*. 2004;363(9422):1724–1727.
14 Davey Smith G, Ebrahim S. Epidemiology–is it time to call it a day? *Int J Epidemiol*. 2001;30(1):1–11.
15 Mariosa D, Carreras-Torres R, Martin RM, Johansson M, Brennan P. Commentary: what can Mendelian randomization tell us about causes of cancer? *Int J Epidemiol*. 2019;48(3):816–821.
16 Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32(1):1–22.
17 Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23(R1):R89–R98.
18 Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27(8):1133–1163.
19 Li X, Jansen L, Chang-Claude J, Hoffmeister M, Brenner H. Risk of colorectal cancer associated with lifetime excess weight. *JAMA Oncol*. 2022;8(5):730–737.
20 Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7:e34408.
21 Elsworth B, Lyon M, Alexander T, et al. The MRC IEU OpenGWAS data infrastructure. *bioRxiv*. 2020. https://doi.org/10.1101/2020.08.10.244293v1.
22 Huyghe JR, Bien SA, Harrison TA, et al. Discovery of common and rare genetic risk variants for colorectal cancer. *Nat Genet*. 2019;51(1):76–87.
23 Huang J, Huffman JE, Huang Y, et al. Genomics and phenomics of body mass index reveals a complex disease network. *Nat Commun*. 2022;13(1):7973.
24 Labadie JD, Harrison TA, Banbury B, et al. Postmenopausal hormone therapy and colorectal cancer risk by molecularly defined subtypes and tumor location. *JNCI Cancer Spectr*. 2020;4(5):pkaa042.
25 Hidaka A, Harrison TA, Cao Y, et al. Intake of dietary fruit, vegetables, and fiber and risk of colorectal cancer according to molecular subtypes: a pooled analysis of 9 studies. *Cancer Res*. 2020;80(20):4578–4590.
26 Boland CR, Thibodeau SN, Hamilton SR, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res*. 1998;58(22):5248–5257.
27 Phipps AI, Alwers E, Harrison T, et al. Association between molecular subtypes of colorectal tumors and patient survival, based on pooled analysis of 7 international studies. *Gastroenterology*. 2020;158(8):2158–2168.e4.
28 Lindor NM, Burgart LJ, Leontovich O, et al. Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. *J Clin Oncol*. 2002;20(4):1043–1048.
29 Shia J. Immunohistochemistry versus microsatellite instability testing for screening colorectal cancer patients at risk for hereditary nonpolyposis colorectal cancer syndrome. Part I. The utility of immunohistochemistry. *J Mol Diagn*. 2008;10(4):293–300.

30 Warth A, Kloor M, Schirmacher P, Blaker H. Genetics and epigenetics of small bowel adenocarcinoma: the interactions of CIN, MSI, and CIMP. *Mod Pathol*. 2011;24(4):564–570.
31 Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. *Gastroenterology*. 2010;138(6):2088–2100.
32 Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol*. 2013;42(5):1497–1501.
33 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B*. 1995;57(1):289–300.
34 Botteri E, Borroni E, Sloan EK, et al. Smoking and colorectal cancer risk, overall and by molecular subtypes: a meta-analysis. *Am J Gastroenterol*. 2020;115(12):1940–1949.
35 Wootton RE, Richmond RC, Stuijfzand BG, et al. Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a Mendelian randomisation study. *Psychol Med*. 2020;50(14):2435–2443.
36 Saunders GRB, Wang X, Chen F, et al. Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature*. 2022;612(7941):720–724.
37 Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol*. 2011;40(3):755–764.
38 Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–560.
39 Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med*. 2017;36(11):1783–1802.
40 Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44(2):512–525.
41 Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. 2017;32(5):377–389.
42 Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. 2016;40(4):304–314.
43 Bowden J, Spiller W, Del Greco MF, et al. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. *Int J Epidemiol*. 2018;47(4):1264–1278.
44 Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int J Epidemiol*. 2017;46(6):1734–1739.
45 Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. *BMJ*. 2021;375:n2233.
46 Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomization: the STROBE-MR statement. *JAMA*. 2021;326(16):1614–1621.
47 He X, Wu K, Ogino S, Giovannucci EL, Chan AT, Song M. Association between risk factors for colorectal cancer and risk of serrated polyps and conventional adenomas. *Gastroenterology*. 2018;155(2):355–373.e18.
48 Szylberg L, Janiczek M, Popiel A, Marszalek A. Expression of COX-2, IL-1beta, TNF-alpha and IL-4 in epithelium of serrated adenoma, adenoma and hyperplastic polyp. *Arch Med Sci*. 2016;12(1):172–178.
49 Chen J, Pitmon E, Wang K. Microbiome, inflammation and colorectal cancer. *Semin Immunol*. 2017;32:43–53.
50 Inamura K, Hamada T, Bullman S, Ugai T, Yachida S, Ogino S. Cancer as microenvironmental, systemic and environmental diseases: opportunity for transdisciplinary microbiomics science. *Gut*. 2022. https://doi.org/10.1136/gutjnl-2022-327209.
51 Maciel SS, Feres M, Goncalves TE, et al. Does obesity influence the subgingival microbiota composition in periodontal health and disease? *J Clin Periodontol*. 2016;43(12):1003–1012.
52 Yu J, Chen Y, Fu X, et al. Invasive Fusobacterium nucleatum may play a role in the carcinogenesis of proximal colon cancer through the serrated neoplasia pathway. *Int J Cancer*. 2016;139(6):1318–1326.
53 Ito M, Kanno S, Nosho K, et al. Association of Fusobacterium nucleatum with clinical and molecular features in colorectal serrated pathway. *Int J Cancer*. 2015;137(6):1258–1268.
54 Mima K, Nishihara R, Qian ZR, et al. Fusobacterium nucleatum in colorectal carcinoma tissue and patient prognosis. *Gut*. 2016;65(12):1973–1980.
55 Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol*. 2016;40(7):597–608.