University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Nounu, A;Richmond, RC;Stewart, ID;Surendran, P;Wareham, NJ;Butterworth, A;Weinstein, SJ;Albanes, D;Baron, JA;Hopper, JL;Figueiredo, JC;Newcomb, PA;Lindor, NM;Casey, G;Platz, EA;Marchand, LL;Ulrich, CM;Li, CI;van Dujinhoven, FJB;Gsur, A;Campbell, PT;Moreno, V;Vodicka, P;Vodickova, L;Amitay, E;Alwers, E;Chang-Claude, J;Sakoda, LC;Slattery, ML;Schoen, RE;Gunter, MJ;Castellví-Bel, S;Kim, HR;Kweon, SS;Chan, AT;Li, L;Zheng, W;Bishop, DT;Buchanan, DD;Giles, GG;Gruber, SB;Rennert, G;Stadler, ZK;Harrison, TA;Lin, Y;Keku, TO;Woods, MO;Schafmayer, C;Van Guelpen, B;Gallinger, S;Hampel, H;Berndt, SI;Pharoah, PDP;Lindblom, A;Wolk, A;Wu, AH;White, E;Peters, U;Drew, DA;Scherer, D;Bermejo, JL;Brenner, H;Hoffmeister, M;Williams, AC;Relton, CL

Title:
Salicylic acid and risk of colorectal cancer: A two-sample mendelian randomization study

Date:
2021-11-01

Citation:
Nounu, A., Richmond, R. C., Stewart, I. D., Surendran, P., Wareham, N. J., Butterworth, A., Weinstein, S. J., Albanes, D., Baron, J. A., Hopper, J. L., Figueiredo, J. C., Newcomb, P. A., Lindor, N. M., Casey, G., Platz, E. A., Marchand, L. L., Ulrich, C. M., Li, C. I., van Dujinhoven, F. J. B. ,... Relton, C. L. (2021). Salicylic acid and risk of colorectal cancer: A two-sample mendelian randomization study. Nutrients, 13 (11), https://doi.org/10.3390/nu13114164.

Persistent Link:
https://hdl.handle.net/11343/296315

License:
CC BY

nutrients

MDPI

Article

Salicylic Acid and Risk of Colorectal Cancer: A Two-Sample Mendelian Randomization Study

Aayah Nounu [1,2,*], Rebecca C. Richmond [1], Isobel D. Stewart [3], Praveen Surendran [4,5,6,7], Nicholas J. Wareham [3], Adam Butterworth [4,5,6,8,9], Stephanie J. Weinstein [10], Demetrius Albanes [10], John A. Baron [11], John L. Hopper [12,13], Jane C. Figueiredo [14,15], Polly A. Newcomb [16,17], Noralane M. Lindor [18], Graham Casey [19], Elizabeth A. Platz [20], Loïc Le Marchand [21], Cornelia M. Ulrich [22], Christopher I. Li [16], Fränzel J. B. van Duijnhoven [23], Andrea Gsur [24], Peter T. Campbell [25], Víctor Moreno [26,27,28,29], Pavel Vodicka [30,31,32], Ludmila Vodickova [30,31,32], Efrat Amitay [33], Elizabeth Alwers [33], Jenny Chang-Claude [34,35], Lori C. Sakoda [16,36], Martha L. Slattery [37], Robert E. Schoen [38], Marc J. Gunter [39], Sergi Castellví-Bel [40], Hyeong-Rok Kim [41], Sun-Seog Kweon [42,43], Andrew T. Chan [44,45,46,47,48,49], Li Li [50], Wei Zheng [51], D. Timothy Bishop [52], Daniel D. Buchanan [53,54,55], Graham G. Giles [12,56,57], Stephen B. Gruber [58], Gad Rennert [59,60,61], Zsofia K. Stadler [62], Tabitha A. Harrison [23], Yi Lin [23], Temitope O. Keku [63], Michael O. Woods [64], Clemens Schafmayer [65], Bethany Van Guelpen [66,67], Steven Gallinger [68], Heather Hampel [69], Sonja I. Berndt [10], Paul D. P. Pharoah [70], Annika Lindblom [71,72], Alicja Wolk [73,74], Anna H. Wu [75], Emily White [23,76], Ulrike Peters [23,76], David A. Drew [45,47], Dominique Scherer [77], Justo Lorenzo Bermejo [77], Hermann Brenner [34,78,79], Michael Hoffmeister [34], Ann C. Williams [2] and Caroline L. Relton [1]

1 Integrative Cancer Epidemiology Programme (ICEP), Medical Research Council (MRC) Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, Bristol BS8 2BN, UK; rebecca.richmond@bristol.ac.uk (R.C.R.); caroline.relton@bristol.ac.uk (C.L.R.)
2 School of Cellular and Molecular Medicine, University of Bristol, Bristol BS8 1TD, UK; ann.c.williams@bristol.ac.uk
3 MRC Epidemiology Unit, School of Clinical Medicine, University of Cambridge, Cambridge CB2 0SL, UK; Isobel.Stewart@mrc-epid.cam.ac.uk (I.D.S.); nick.wareham@mrc-epid.cam.ac.uk (N.J.W.)
4 British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK; praveen.x.surendran@gsk.com (P.S.); asb38@medschl.cam.ac.uk (A.B.)
5 British Heart Foundation Centre of Research Excellence, Division of Cardiovascular Medicine, University of Cambridge, Cambridge CB2 0QQ, UK
6 Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Cambridge CB10 1SA, UK
7 Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK
8 National Institute for Health Research Blood and Transplant Research Unit in Donor Health and Genomics, University of Cambridge, Cambridge CB2 1TN, UK
9 National Institute for Health Research Cambridge Biomedical Research Centre, Cambridge Biomedical Campus, University of Cambridge, Cambridge University Hospitals, Cambridge CB2 0QQ, UK
10 Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20814, USA; weinstes@mail.nih.gov (S.J.W.); albanesd@mail.nih.gov (D.A.); berndts@mail.nih.gov (S.I.B.)
11 Department of Medicine, School of Medicine, University of North Carolina, Chapel Hill, NC 27516, USA; john_baron@med.unc.edu
12 Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, VIC 3053, Australia; johnlh@unimelb.edu.au (J.L.H.); graham.giles@cancervic.org.au (G.G.G.)
13 Department of Epidemiology, Institute of Health and Environment, School of Public Health, Seoul National University, Seoul 08826, Korea
14 Department of Medicine, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA; jane.figueiredo@cshs.org
15 Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 0032, USA
16 Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109-1024, USA; pnewcomb@fredhutch.org (P.A.N.); cili@fhcrc.org (C.I.L.); lori.sakoda@kp.org (L.C.S.)
17 School of Public Health, University of Washington, Seattle, WA 98195, USA
18 Department of Health Science Research, Mayo Clinic, Scottsdale, AZ 85259, USA; nlindor@mayo.edu

check for updates

Citation: Nounu, A.; Richmond, R.C.; Stewart, I.D.; Surendran, P.; Wareham, N.J.; Butterworth, A.; Weinstein, S.J.; Albanes, D.; Baron, J.A.; Hopper, J.L.; et al. Salicylic Acid and Risk of Colorectal Cancer: A Two-Sample Mendelian Randomization Study. *Nutrients* **2021**, *13*, 4164. https://doi.org/10.3390/nu13114164

Academic Editor: Rachel Murphy

Received: 13 October 2021
Accepted: 17 November 2021
Published: 21 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

[19] Center for Public Health Genomics, Department of Public Health Sciences, University of Virginia, Charlottesville, VA 22908, USA; gc8r@virginia.edu
[20] Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA; eplatz@jhsph.edu
[21] Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI 96813, USA; loic@cc.hawaii.edu
[22] Huntsman Cancer Institute, Department of Population Health Sciences, University of Utah, Salt Lake City, UT 84112, USA; neli@hci.utah.edu
[23] Division of Human Nutrition and Health, Department of Agrotechnology and Food Sciences, Wageningen University & Research, 6700 HB Wageningen, The Netherlands; franzel.vanduijnhoven@wur.nl (F.J.B.v.D.); tharriso@fredhutch.org (T.A.H.); ylin2@fredhutch.org (Y.L.); ewhite@fredhutch.org (E.W.); upeters@fredhutch.org (U.P.)
[24] Institute of Cancer Research, Department of Medicine I, Medical University Vienna, 1090 Vienna, Austria; andrea.gsur@meduniwien.ac.at
[25] Department of Population Science, American Cancer Society, Atlanta, GA 30303, USA; Peter.Campbell@cancer.org
[26] Oncology Data Analytics Program, Catalan Institute of Oncology-IDIBELL, 08908 Barcelona, Spain; v.moreno@iconcologia.net
[27] CIBER Epidemiología y Salud Pública (CIBERESP), 28029 Madrid, Spain
[28] Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, 08007 Barcelona, Spain
[29] ONCOBEL Program, Bellvitge Biomedical Research Institute (IDIBELL), 08908 Barcelona, Spain
[30] Department of Molecular Biology of Cancer, Institute of Experimental Medicine of the Czech Academy of Sciences, 142 20 Prague, Czech Republic; pvodicka@biomed.cas.cz (P.V.); ludovod@seznam.cz (L.V.)
[31] Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Nové Město, 121 08 Prague, Czech Republic
[32] Faculty of Medicine and Biomedical Center in Pilsen, Charles University, 323 00 Pilsen, Czech Republic
[33] Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; e.amitay@dkfz-heidelberg.de (E.A.); elizabeth.alwers@dkfz-heidelberg.de (E.A.)
[34] Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; j.chang-claude@dkfz.de (J.C.-C.); h.brenner@dkfz.de (H.B.); m.hoffmeister@dkfz.de (M.H.)
[35] Department of Oncology, Haematology and BMT, University Medical Centre Hamburg-Eppendorf, University Cancer Centre Hamburg (UCCH), 20251 Hamburg, Germany
[36] Division of Research, Kaiser Permanente Northern California, Oakland, CA 94612, USA
[37] Department of Internal Medicine, University of Utah, Salt Lake City, UT 84112, USA; marty.slattery@hsc.utah.edu
[38] Department of Medicine and Epidemiology, University of Pittsburgh Medical Center, Pittsburgh, PA 15261, USA; rschoen@pitt.edu
[39] Nutrition and Metabolism Section, International Agency for Research on Cancer, World Health Organization, 69372 Lyon, France; GunterM@iarc.fr
[40] Gastroenterology Department, Hospital Clínic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD), University of Barcelona, 08036 Barcelona, Spain; sbel@clinic.cat
[41] Department of Surgery, Chonnam National University Hwasun Hospital and Medical School, Hwasun 58128, Korea; drkhr@jnu.ac.kr
[42] Department of Preventive Medicine, Chonnam National University Medical School, Gwangju 61186, Korea; ss.kweon2@gmail.com
[43] Jeonnam Regional Cancer Center, Chonnam National University Hwasun Hospital, Hwasun 58128, Korea
[44] Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA; achan@partners.org
[45] Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA; DADREW@mgh.harvard.edu
[46] Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
[47] Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[48] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA 02115, USA
[49] Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA 02115, USA
[50] Department of Family Medicine, University of Virginia, Charlottesville, VA 22903, USA; ll134q@rocketmail.com
[51] Division of Epidemiology, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt Epidemiology Center, Vanderbilt University School of Medicine, Nashville, TN 37232, USA; wei.zheng@Vanderbilt.Edu

[52] Leeds Institute of Cancer and Pathology, School of Medicine, University of Leeds, Leeds LS2 9JT, UK; d.t.bishop@leeds.ac.uk
[53] Colorectal Oncogenomics Group, Department of Clinical Pathology, The University of Melbourne, Parkville, VIC 3010, Australia; daniel.buchanan@unimelb.edu.au
[54] Melbourne Medical School, University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, VIC 3010, Australia
[55] Genetic Medicine and Family Cancer Clinic, The Royal Melbourne Hospital, Parkville, VIC 3000, Australia
[56] Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, VIC 3004, Australia
[57] Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC 3168, Australia
[58] Department of Preventive Medicine & USC Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA; sgruber@coh.org
[59] Department of Community Medicine and Epidemiology, Lady Davis Carmel Medical Center, Haifa 3448516, Israel; rennert@tx.technion.ac.il
[60] Ruth and Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa 3200003, Israel
[61] Clalit National Cancer Control Center, Haifa 3436212, Israel
[62] Memorial Sloan Kettering Cancer Center, Department of Medicine, New York, NY 10065, USA; stadlerz@mskcc.org
[63] Center for Gastrointestinal Biology and Disease, School of Medicine, University of North Carolina, Chapel Hill, NC 27599-7555, USA; temitope_keku@med.unc.edu
[64] Discipline of Genetics, Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL A1B 3V6, Canada; mwoods@mun.ca
[65] Department of General Surgery, University Hospital Rostock, 18057 Rostock, Germany; Clemens.Schafmayer@uksh.de
[66] Department of Radiation Sciences, Oncology Unit, Umeå University, 901 87 Umeå, Sweden; bethany.vanguelpen@umu.se
[67] Wallenberg Centre for Molecular Medicine, Department of Biomedical and Clinical Sciences, Umeå University, 901 87 Umeå, Sweden
[68] Lunenfeld Tanenbaum Research Institute, Mount Sinai Hospital, Faculty of Medicine, University of Toronto, Toronto, ON M5G 1X5, Canada; Steven.Gallinger@uhn.on.ca
[69] Division of Human Genetics, Department of Internal Medicine, The Ohio State University Comprehensive Cancer Center, Columbus, OH 43210, USA; Heather.Hampel@osumc.edu
[70] Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK; pp10001@medschl.cam.ac.uk
[71] Department of Clinical Genetics, Karolinska University Hospital, 171 64 Solna, Sweden; Annika.Lindblom@ki.se
[72] Department of Molecular Medicine and Surgery, Karolinska Institutet, 171 64 Solna, Sweden
[73] Institute of Environmental Medicine, Karolinska Institutet, 171 64 Solna, Sweden; Alicja.Wolk@ki.se
[74] Department of Surgical Sciences, Uppsala University, 751 85 Uppsala, Sweden
[75] Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA 90032, USA; annawu@usc.edu
[76] Department of Epidemiology, University of Washington School of Public Health, Seattle, WA 98195, USA
[77] Institute of Medical Biometry and Informatics, Medical Faculty, University of Heidelberg, Im Neuenheimer Feld 130.3, 69120 Heidelberg, Germany; Dominique.Scherer@uni-heidelberg.de (D.S.); lorenzo@imbi.uni-heidelberg.de (J.L.B.)
[78] Division of Preventive Oncology, German Cancer Research Center (DKFZ), National Center for Tumor Diseases (NCT), 69120 Heidelberg, Germany
[79] German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany
* Correspondence: an0435@bristol.ac.uk

Abstract: Salicylic acid (SA) has observationally been shown to decrease colorectal cancer (CRC) risk. Aspirin (acetylsalicylic acid, that rapidly deacetylates to SA) is an effective primary and secondary chemopreventive agent. Through a Mendelian randomization (MR) approach, we aimed to address whether levels of SA affected CRC risk, stratifying by aspirin use. A two-sample MR analysis was performed using GWAS summary statistics of SA (INTERVAL and EPIC-Norfolk, N = 14,149) and CRC (CCFR, CORECT, GECCO and UK Biobank, 55,168 cases and 65,160 controls). The DACHS study (4410 cases and 3441 controls) was used for replication and stratification of aspirin-use. SNPs proxying SA were selected via three methods: (1) functional SNPs that influence the activity of aspirin-metabolising enzymes; (2) pathway SNPs present in enzymes' coding regions; and (3) genome-wide significant SNPs. We found no association between functional SNPs and SA levels. The pathway and genome-wide SNPs showed no association between SA and CRC risk (OR: 1.03, 95% CI: 0.84–1.27 and

OR: 1.08, 95% CI: 0.86–1.34, respectively). Results remained unchanged upon aspirin use stratification. We found little evidence to suggest that an SD increase in genetically predicted SA protects against CRC risk in the general population and upon stratification by aspirin use.

Keywords: salicylic acid; aspirin; colorectal cancer; Mendelian randomization

1. Introduction

Colorectal cancer (CRC) is the fourth most common cancer in the UK and worldwide [1,2]. Although incidence rates among the over 50s have remained relatively stable, rates in younger age groups have increased in both the UK and US populations [3,4]. This highlights a need to find better and complementary prevention strategies to reduce risk of cancer.

Salicylic acid (SA) is a dietary metabolite that can be found in various fruits, vegetables, herbs, and spices [5–7]. Results from a meta-analysis of 19 cohort studies found that combined intake of fruits and vegetables reduced the risk of colorectal cancer (summary relative risk (RR): 0.90, 95% CI: 0.83–0.98) [8]. Whilst dietary fibre obtained from fruits and vegetables indicates a possible mechanism for decreased risk [9], it has also been suggested that increased levels of SA obtained through their consumption may play a role [7]. In addition, salicylates can be obtained through pharmacological intervention in the form of aspirin (acetylsalicylic acid), a well-known analgesic used to treat fever, inflammation, and acute pain [10], which is rapidly deacetylated to form SA [11,12] (Figure 1), the active form of the aspirin metabolic pathway [13,14]. Whilst SA can be obtained from the diet, the concentrations achieved (male and female median intake from diet 4.4 mg/day and 3.2 mg/day, respectively [6]) are much lower than through aspirin ingestion (aspirin doses ranging between 75 mg to ≥325 mg given daily/alternate days) [15]. Therefore, it is unclear whether concentrations achieved from the diet alone are sufficient to protect against cancer or whether larger doses obtained through pharmacological intervention are required.

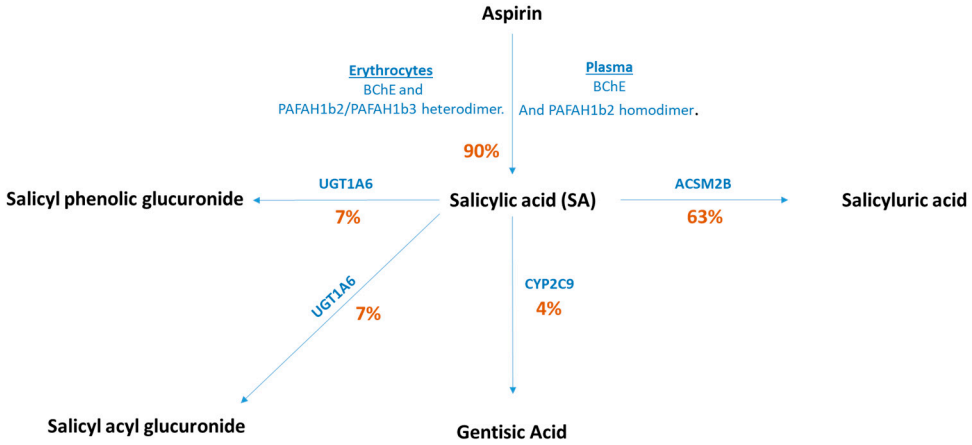


Figure 1. Aspirin metabolism pathway. Roughly 10% of aspirin remains unchanged and is excreted in the urine as aspirin. Aspirin is broken down into various metabolites, the most active of them being salicylic acid [14,16]. Various enzymes are involved in the metabolism pathway. The percentages indicate how much of the drug is being metabolised in that pathway. Abbreviations: BChE, butyrylcholinesterase; PAFAH1b2, platelet-activating factor acetylhydrolase 2; PAFAH1b3, platelet-activating factor acetylhydrolase 3; UGT1A6, UDP-glucuronosyltransferase 1–6; ACSM2B, Acyl-CoA Synthetase Medium-Chain Family Member 2B and CYP450, cytochrome P450.

As of yet, no primary prevention trials have been carried out to assess the effect of SA intervention on CRC risk, but the evidence of aspirin as a chemopreventive agent is

clear [17]. A long-term follow up of a randomised controlled trial (RCT) in the Women's Health Study (WHS) showed that alternate day aspirin intake reduced the risk of CRC after a median of 17.5 years follow up (HR: 0.80, 95% confidence intervals (CI): 0.67–0.97) [18] and a meta-analysis of observational studies showed that aspirin is protective against CRC (relative risk (RR): 0.79, 95% CI: 0.74–0.85) [19]. Further evidence comes from RCTs for primary and secondary prevention of vascular events. These showed that aspirin reduces the risk of CRC incidence and mortality (HR: 0.76, 95% CI: 0.60–0.96 and odds ratio (OR): 0.79, 95% CI: 0.68–0.92, respectively) [20,21]. Considering aspirin is rapidly deacetylated to form SA in under 30 min [22], and that evidence in the form of in vivo and in vitro experiments have previously shown SA to be an antiproliferative and antitumour agent [23–25], it may be that metabolism of aspirin leading to increased circulating SA levels may partially explain aspirin's chemopreventive mode of action.

Although many observational studies have shown an inverse association between aspirin use and CRC risk, few have directly assessed the association between SA itself and CRC. In order to identify the true effect of SA on CRC risk, conducting an RCT would be the ideal study design. However, RCTs for cancer primary prevention are lengthy and costly, therefore it would be helpful to test this association using statistical methods such as Mendelian Randomization (MR). MR uses genetic variants (mostly single nucleotide polymorphisms (SNPs)) related to modifiable factors (such as metabolite levels) to investigate the causal role of these factors on risk of disease [26–28]. Through this method, MR has been likened to RCTs in that genetic variants are randomly allocated at conception the same way that an intervention is randomly allocated at the start of a trial [29,30]. This lends many advantages such as overcoming the issues of confounding and reverse causation, which are commonly encountered in observational epidemiology [29]. MR has previously been useful in predicting trial outcomes such as the case of selenium and prostate cancer in The Selenium and Vitamin E Cancer Prevention Trial (SELECT) [31]. Results from an MR study mimicked the findings of this RCT and may have been useful to inform whether to conduct a trial that cost $114 million and that was weakly associated with increasing high-grade prostate cancer risk [32].

For this reason, we applied an MR approach using genetic "instruments" or proxies for SA to assess the causal effect of this metabolite on the risk of CRC. Since aspirin is rapidly deacetylated to SA [22], and therefore a plausible proxy of increased SA levels, we also stratified our analysis between aspirin users and non-users to test the hypothesis of whether diet-derived levels of SA alone would affect risk of CRC or whether higher concentrations achieved through pharmacological intervention in the form of aspirin was required to identify an effect. Based on previous observational evidence, we hypothesise that a genetically predicted increase in SA levels would reduce the risk of CRC, with a stronger effect observed in aspirin users.

2. Materials and Methods

2.1. Genetic Variants for Salicylic Acid

We applied a two-sample MR study design to test for the association of SA levels (sample 1) with risk of CRC (sample 2). GWAS and meta-analysis of salicylate levels were performed using 5841 participants from the EPIC-Norfolk study [33] and 8455 from the INTERVAL study [34]. The percentage of samples with missing salicylate measurements was low (0.43% and 1.44% in EPIC-Norfolk and INTERVAL respectively), providing a total sample size of 14,149. Salicylate was measured independently in each study as one of many metabolites measured using the Metabolon DiscoveryHD4® platform (Metabolon, Inc., Durham, NC, USA), from non-fasted plasma samples (predominantly non-fasted samples in EPIC-Norfolk) collected at baseline. Measures that were median normalised for run day were natural log transformed, winsorised to 5 standard deviations from the mean, before being regressed against age, sex, and study-specific variables (measurement consignment in EPIC-Norfolk and measurement consignment, INTERVAL centre, plate number, appointment month, lag time between blood donation appointment and sample

processing, and the first 5 ancestry principal components in INTERVAL) using linear regression. Residuals from this regression were standardised (mean 0, standard deviation 1) and used for further analysis. Genotyping was performed in both studies using the Affymetrix Axiom UK Biobank genotyping array. In INTERVAL, genotype imputation was performed using the combined UK10K+1000 Genomes Phase 3 reference panel. In EPIC-Norfolk, imputation was performed using the Haplotype Reference Consortium reference panel, with additional variants imputed using the UK10K+1000 Genomes Phase 3 reference panel. Genome-wide association analyses were performed using BOLT-LMM (version 2.2) [35], and variants with a MAF < 0.01% and INFO < 0.3 were excluded. Associations from the two studies were pooled using inverse variance weighted fixed effect meta-analysis implemented in METAL [36], applying a minor allele count threshold in each study of >10.

The causal effect of SA on risk of CRC was assessed using 3 sets of genetic variants (SNPs) related to SA: (1) functional SNPs that influence aspirin and SA metabolising enzymes' activity (derived from Figure 1)—termed "functional SNPs"; (2) pathway SNPs, those that are present in the coding regions of genes that are involved in aspirin and SA metabolism (based on the NCBI Build 37/UCSC hg19 from https://grch37.ensembl.org/index.html (accessed on 14 December 2016), Supplementary Table S1) termed "pathway SNPs"; (3) genome-wide significant SNPs associated with levels of circulating aspirin metabolites—termed "genome-wide SNPs". Pathways SNPs were defined as having a Bonferroni threshold of association (p value $0.05/2701 = 1.85 \times 10^{-5}$), a MAF $\geq$ 0.01%, as well as a consistent direction of effect in both EPIC-Norfolk and INTERVAL. Genome-wide signals were defined as having an association p value $< 5 \times 10^{-8}$ in the meta-analysis, a MAF $\geq$ 0.01%, consistent direction of effect across the two studies, and association p value < 0.01 in both studies

To account for genetic correlation, linkage disequilibrium (LD) clumping at an $R^2 < 0.001$ and 10,000 kb window was performed to retain the SNP most strongly associated with the metabolite for downstream analysis. Since an $R^2 < 0.001$ is considered highly stringent, we also used an $R^2 < 0.8$ to incorporate more variants while accounting for residual correlation in the model (see Statistical Analysis). An F-statistic for each SNP–exposure association was calculated to reflect the strength of the genetic instrument and indicate any possibility of weak instrument bias, usually inferred when F < 10 [37]. Power calculations were conducted using the mRnd online calculator to identify the OR in both directions that could be detected with the sample size available [38].

2.2. Genetic Variants for CRC Incidence

SNP–outcome associations were obtained from the Colon Cancer Family Registry (CCFR), Colorectal Cancer Transdisciplinary Study (CORECT), Genetics and Epidemiology of Colorectal Cancer (GECCO) consortia, and the UK Biobank (55,168 cases and 65,160 controls), hereafter collectively termed as GECCO [39–41]. Genetic data from a population-based case-control study from southwestern Germany (Darmkrebs: Chancen der Verhütung durch Screening (DACHS)) was used to assess replication of the findings, and to run an MR analysis stratified by aspirin intake, since this study recorded aspirin use (defined as twice per week for at least a year) [42–44]. This study comprised 4410 cases, of which 810 (18.37% of cases) were aspirin users and 3340 (75.74%) were non-users, and 260 cases (5.90%) were excluded as they had reported use of other non-aspirin NSAIDs. This study also contained 3441 controls, of which 779 (22.64%) had recorded aspirin use and 2320 (67.42%) were recorded as non-users, and 342 controls (9.94%) were excluded as they had reported use of other non-aspirin NSAIDs.

To assess the causal effect of SA on CRC risk, we tested for association in GECCO, and also stratified the analysis between aspirin users and non-users in DACHS to investigate whether increased SA levels via pharmacological intervention is required to see an effect. We obtained summary association statistics from GECCO and also conducted logistic regression analyses adjusting for age and sex in the DACHS study for all the participants. We

then stratified the participants of the DACHS study to aspirin users and non-users before repeating the logistic regression analyses again. Genetic instruments that had a MAF ≤ 0.01 in both GECCO and DACHS (all participants) were excluded from further analyses.

2.3. Statistical Analyses

Analyses were carried out in R version 3.2.3 using the "Two-Sample MR" package [45]. This package allows the formatting, harmonisation, and analysis of summary data from genetic association studies in a semi-automated manner. The Two-Sample MR package automatically assigns effect alleles so that SNP associations with the exposure are positive i.e., so the effect allele is "metabolite-increasing". The SNPs identified as associated with SA can then be extracted from the outcome datasets. Allele harmonization ensures that the effect (metabolite-increasing) allele in the exposure dataset is also treated as the effect allele in the outcome dataset. When only one SNP was associated with the metabolite, Wald ratios (SNP–outcome estimate ÷ SNP–exposure estimate) were calculated to assess the change in log OR per SD increase in the metabolite. When more than one SNP was available, a weighted mean weighted by the inverse variance of the Wald ratio estimates (inverse-variance weighted (IVW) method) was used to assess the causal effect of increased metabolite levels on risk of CRC incidence [46]. To assess the quality of our instruments, we calculated the variance in SA levels explained by the SNPs and the F statistic. The variance explained for each SNP was calculated using the formula: $\frac{2b^2p\ (1-p)}{var}$, where p is the minor allele frequency, b is the SNP effect on the exposure (beta) and var is the variance of the exposure. The F statistic was calculated using the formula: $\frac{r^2(n-1-k)}{((1-r^2)k)}$ where r is the sum of the variance explained by the set of SNPs, n is the sample size of the exposure GWAS and k is the number of SNPs used to proxy the exposure. In the presence of weak instruments, we conducted an MR robust adjusted profile score (MR RAPS), which is a method that provides robust inference when many weak instruments are present [47].

Furthermore, the presence of one invalid instrument, e.g., one that is associated with exposures other than the exposure of interest (horizontal pleiotropy), may bias the results from the IVW method [48]. For this reason, alternative methods that produce an unbiased estimator even when some of the genetic instruments are invalid were used as a sensitivity analysis when more than 2 SNPs were used as exposure instruments (weighted mode, weighted median, and MR Egger) [45,49–51]. The MR Egger test is not constrained to pass through an effect size of 0, unlike the IVW method, allowing the assessment of the presence of directional pleiotropy through the y intercept [48,51]. We also measured the Q statistic to measure the presence of pleiotropy between our instruments. If all the SNPs are valid instruments, then the individual MR estimates for each SNP will only vary by chance. A larger amount of heterogeneity would indicate that one or more of the SNPs are pleiotropic [52].

Due to the presence of a small number of independent SNPs associated with the metabolite, we also conducted a weighted generalised linear regression (WGLR), whereby SNPs in LD ($R^2 < 0.8$) could be used with the incorporation of their correlation as weights in the regression analysis [53]. This was performed using the "LDlinkR" and "Mendelian-Randomization" packages in R (version 3.5.1). The use of multiple SNPs explains more of the variance in the metabolite levels and therefore improves power to detect an effect [53].

We also assessed the possibility of reverse causation through the use of the MR Steiger test, found in the "TwoSampleMR" package, which does so by comparing the variance explained by the SNPs for the exposure and the outcome [45].

3. Results

3.1. Functional SNPs and CRC Risk

To interrogate the effect of SA on CRC risk, we used three methods to select our exposure instruments (Figure 2). In our first approach, we identified four functional SNPs that have been shown to affect enzyme efficiency in the aspirin metabolic pathway (Figure 1). For BChE (rs6445035), the presence of an A allele increase has been associated

with a decrease in aspirin hydrolysis by around 1.2 nmol/mL/min [54]. The UGT1A6 variants, rs2070959 and rs1105879, predict a higher metabolic activity of the enzyme than the wild type [55,56]. Furthermore, a variant in CYP2C9 (rs1799853) encodes an enzyme with reduced activity [57].

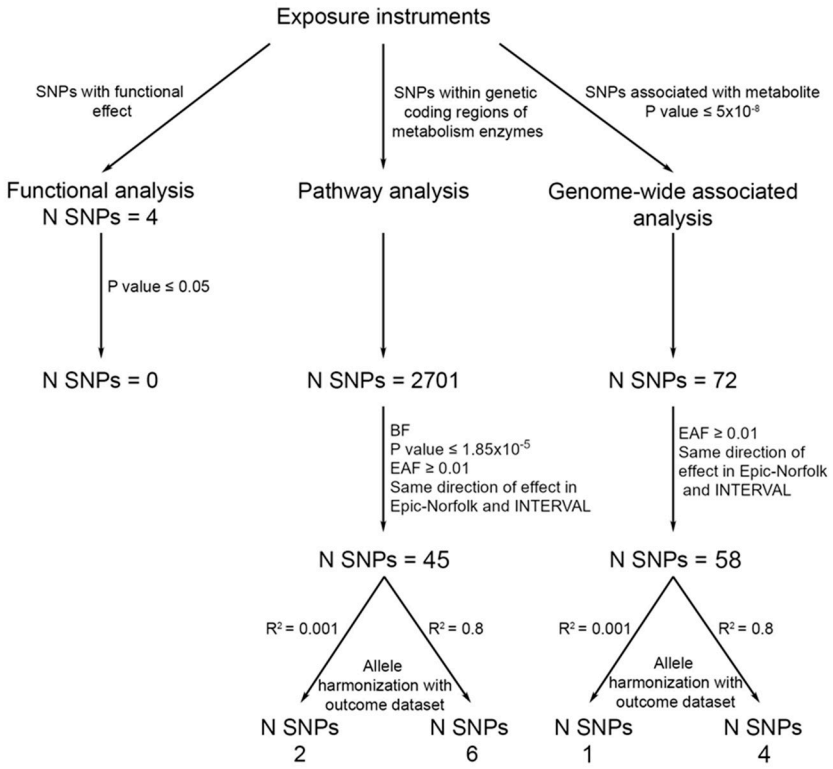


Figure 2. Instrument selection for functional, pathway, and genome-wide SNPs. Abbreviations: SA, salicylic acid; EAF, effect allele frequency; BF, Bonferroni.

These SNPs were tested for association with SA in the INTERVAL and (EPIC)-Norfolk study, however none of the SNPs reached nominal significance with the metabolite (Figure 3A) (Supplementary Table S2). For this reason, these SNPs were therefore not taken forward in an MR analysis.

3.2. Pathway SNPs and CRC Risk

We investigated genetic variants within the coding regions of the enzymes involved in aspirin and SA metabolism (Figure 1). These were BChE, PAFAH1b2, PAFAH1b3, UGT1A6, ACSM2B and CYP2C9.

We obtained summary statistics for 2701 SNPs within the genetic coding regions of the enzymes for SA. We applied a Bonferroni threshold of association (P value 0.05/2701 = 1.85×10^{-5}) for SNPS and restricted to SNPs with a consistent direction of effects in both studies and a minor allele frequency of $\geq$0.01 in the exposure and outcome studies. This identified 45 SNPs that could be used to instrument SA. These SNPs were then clumped at an $R^2 < 0.001$ and 0.8, providing two and six SNPs, respectively, to instrument SA levels (Figure 2). These explained 0.03% and 0.09% of the variance in SA levels, and had an F statistic of 1.74 and 2.16, respectively (Table 1).

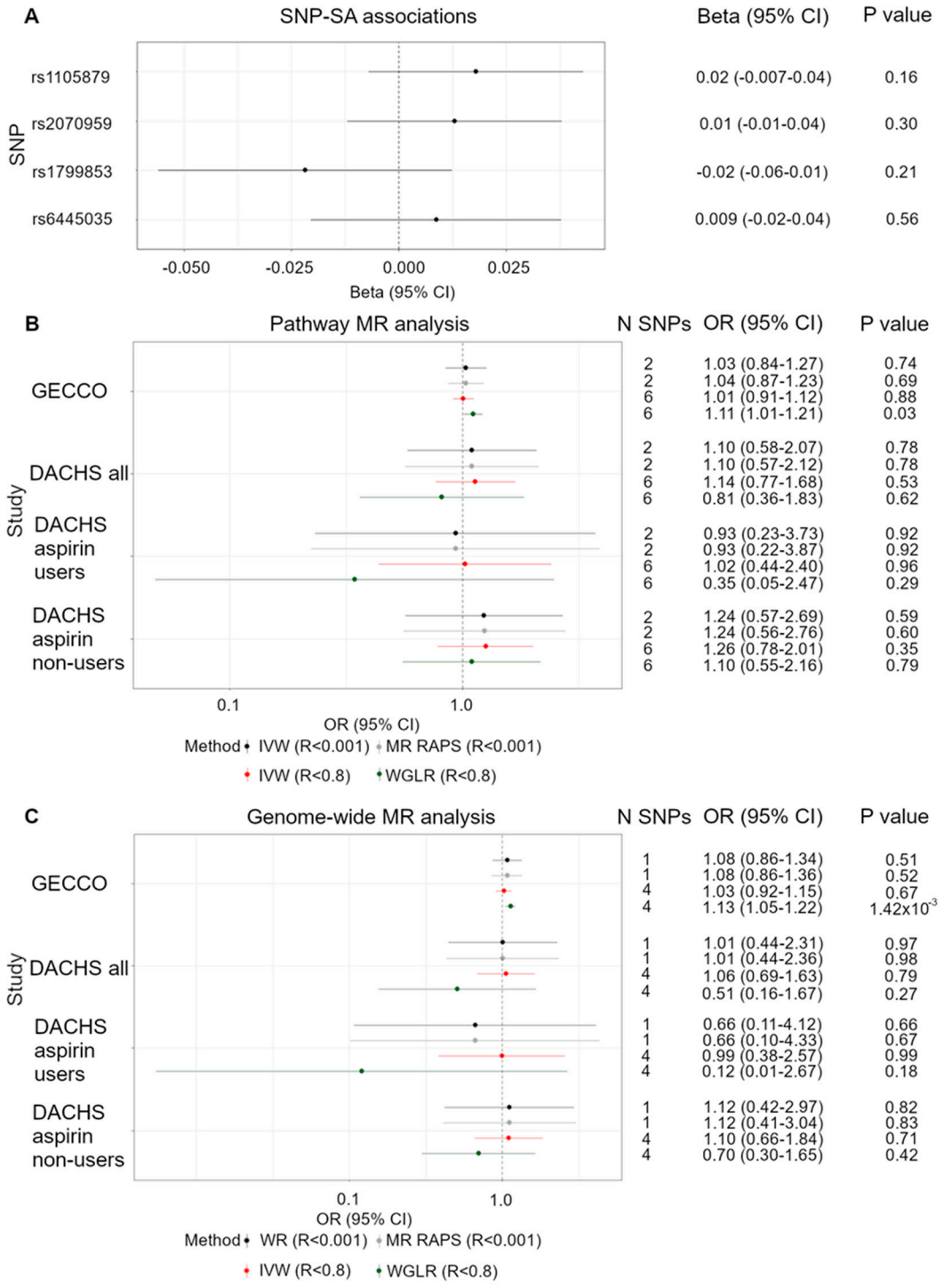


Figure 3. Functional SNP metabolite associations and two-sample pathway MR analysis. (**A**) Forest plot of single SNP associations with salicylic acid for the functional SNPs. (**B**) Forest plot of one SD increase in SA and its effect on CRC risk, instrumented by pathway SNPs and applying three methods: IVW after applying an LD threshold of $R^2 < 0.001$ (black), MR RAPS after applying an LD threshold of $R^2 < 0.001$ (grey), IVW after applying an LD threshold of $R^2 < 0.8$ (red) and a WGLR after applying an LD threshold of $R^2 < 0.8$ (green). (**C**) Forest plot of one SD increase in SA and its effect on CRC risk, instrumented by genome-wide SNPs and applying three methods: WR after applying an LD threshold of $R^2 < 0.001$ (black), MR RAPS after applying an LD threshold of $R^2 < 0.001$ (grey), IVW after applying an LD threshold of $R^2 < 0.8$ (red) and a WGLR after applying an LD threshold of $R^2 < 0.8$ (green). Abbreviations: OR, odds ratio; IVW, inverse variance weighted; WGLR, weighted generalised linear regression; WR, Wald ratio; LD, linkage disequilibrium.

Table 1. Exposure instruments used in the MR analysis.

SNP Set	Study	Outcome Sample Size	Percentage Cases (%)	N SNPs	LD R^2	Variance Explained R^2 (%)	F Statistics	OR Detected at 80% Power	
								Decreased Risk	Increased Risk
Pathway SNPs	GECCO	120,328	45.85 55,168/120,328)	2	0.001	0.025	1.74	0.90	1.11
	DACHS	7851	56.17	2				0.68	1.51
	DACHS aspirin users	1589	(4410/7851)	2				0.43	2.38
	DACHS aspirin non-users	5660	50.98	2				0.64	1.64
	GECCO	120,328	(810/1589)	6	0.8	0.092	2.16	0.95	1.06
	DACHS	7851	59.01	6				0.81	1.24
	DACHS aspirin users	1589	(3340/5660)	6				0.63	1.58
	DACHS aspirin non-users	5660	45.85 (55,168/120,328)	6				0.78	1.30
Genome-wide SNPs	GECCO	120,328	45.85 55,168/120,328)	2	0.001	0.053	7.44	0.93	1.07
	DACHS	7851	56.17	2				0.76	1.32
	DACHS aspirin users	1589	(4410/7851)	2				0.55	1.83
	DACHS aspirin non-users	5660	50.98	2				0.73	1.42
	GECCO	120,328	(810/1589)	6	0.8	0.090	3.18	0.95	1.06
	DACHS	7851	59.01	6				0.81	1.24
	DACHS aspirin users	1589	(3340/5660)	6				0.63	1.59
	DACHS aspirin non-users	5660	45.85 (55,168/120,328)	6				0.78	1.30

Abbreviations: SA, salicylic acid; LD, linkage disequilibrium; NA, not applicable; OR, odds ratio.

After LD clumping at an $R^2 < 0.001$, 2 SNPs were taken forward in an IVW analysis. We found little evidence of an association between an SD increase in SA and CRC risk (GECCO OR: 1.03, 95% CI: 0.84–1.27 and DACHS OR: 1.10, 95% CI: 0.58–2.07) (Figure 3B). Since aspirin is rapidly deacetylated to form SA [22] and therefore a plausible proxy for increased SA levels, we stratified our analysis between aspirin users and non-users in the DACHS study. Our power calculations show that after stratification, we had 80% power to detect an effect of an SD increase in SA on CRC risk with an OR of $\leq$0.43 and $\geq$2.38 in the reciprocal direction for aspirin users (n = 1589). For non-users (n = 5660), we had 80% power to detect an OR of $\leq$0.64 and $\geq$1.64 in the reciprocal direction (Table 1). However, our MR analysis showed little evidence of an association between SA and CRC risk (OR: 0.93, 95% CI: 0.23–3.73 and OR: 1.24, 95% CI: 0.57–2.69, respectively) (Figure 3B).

The variance explained by these two instruments and their F statistics indicate the possibility of weak instrument bias. For this reason, we conducted MR RAPS, a method that provides robust inference even in the presence of weak instruments [47]. Through this method, no association was found between an SD increase in SA and CRC risk (GECCO OR: 1.04, 95% CI: 0.87–1.23 and DACHS OR:1.10, 95% CI: 0.57–2.12). Results remain unchanged, even after stratification between aspirin users and non-users (OR: 0.93, 95% CI: 0.22–3.87 and OR: 1.24, 95% CI: 0.56–2.76).

Since this LD threshold is known to be very stringent, we used a more relaxed threshold ($R^2 < 0.8$) to increase the number of SNPs available to instrument the metabolite and therefore explain more of the variance in SA levels. This provided six SNPs associated with SA shown in Supplementary Table S3, of which SNP associations with the outcome are also provided in Supplementary Table S4. These SNPs showed no association between SA and CRC risk (GECCO OR: 1.01, 95% CI: 0.91–1.12 and DACHS OR:1.14, 95% CI: 0.77–1.68). Stratification between aspirin use and non-use found no association between the metabolite and CRC risk in aspirin users or non-users (OR: 1.02, 95% CI: 0.44–2.40 and OR: 1.26, 95% CI: 0.78–2.01, respectively).

Using the alternative MR methods (weighted mode, weighted median, and MR Egger), no other association between SA and CRC in both GECCO and DACHS was observed, regardless of stratification (Supplementary Table S5).

Since all the SNPs were found to be on chromosome 16 (Supplementary Table S3), a WGLR method was carried out to account for the SNP correlations and include them as weights into the regression. Through this method, there was no association between SA and CRC risk in DACHS (OR: 0.81, 95% CI: 0.36–1.83) but a positive association in the GECCO sample (OR: 1.11, 95% CI: 1.01–1.21). No association was observed between SA and CRC risk in aspirin users or non-users (OR: 0.35, 95% CI: 0.05–2.47 and OR: 1.10, 95% CI: 0.55–2.16, respectively) (Figure 3B). As a sensitivity analysis, the heterogeneity of the results was appraised through a Q statistic, but no evidence of pleiotropy was observed-i.e., no evidence that the instruments may also be associated with another phenotype (Supplementary Table S6).

To identify whether the causal pathway was in the direction from SA to CRC and not the reverse, we performed the MR Steiger method using the functional SNPs [58]. This method suggested that the causal direction was indeed from SA to CRC, because the SNPs explained more variation in SA levels than CRC risk (Supplementary Table S7).

3.3. Genome-Wide Significant SNPs and CRC Risk

Initially, 72 SNPs were associated with SA at genome-wide significance. We restricted our analysis to SNPs with a MAF threshold of $\geq$0.01 in the exposure and outcome studies, and only included those with a consistent direction of effect in both studies. This resulted in 58 SNPs that were available to instrument SA. After removing SNPs in LD at an $R^2 < 0.001$ and $R^2 < 0.8$, one SNP and four SNPs were available to instrument SA, respectively (Figure 2). These explained 0.05% and 0.09% of the variance in SA levels and had an F statistic of 7.44 and 3.18, respectively (Table 1).

Using the one independent SNP associated with SA at genome-wide significance, WR results showed no association between the genetically predicted metabolite levels and cancer risk (GECCO OR: 1.08, 95% CI: 0.86–1.34 and DACHS OR: 1.01, 95% CI: 0.44–2.31). Our power calculations show that after stratification between aspirin users and non-users in the DACHS study, we had 80% power to detect an effect of an SD increase in SA on CRC risk with an OR of $\leq$0.55 and $\geq$1.83 in the reciprocal direction for aspirin users (n = 1589). For non-users (n = 5660), we had 80% power to detect an OR of $\leq$0.73 and $\geq$1.42 in the reciprocal direction (Table 1), however, we found no association between SA levels and CRC in aspirin users (OR: 0.66, 95% CI: 0.11–4.12) and non-users (OR: 1.12, 95% CI: 0.42–2.97) (Figure 3C).

Due to the possibility of weak instrument bias, we also conducted an MR RAPS approach, but results remained unchanged (GECCO OR: 1.08, 95% CI: 0.86–1.36, DACHS OR: 1.01, 95% CI: 0.44–2.36, DACHS aspirin users OR: 0.66, 95% CI: 0.10–4.33 and DACHS aspirin non-users OR: 1.12, 95% CI: 0.41–3.04).

To explain more of the variance, we used a less stringent LD threshold of $R^2 < 0.8$, and therefore four SNPs to instrument SA (Supplementary Table S8, associations with CRC are also found in Supplementary Table S9). IVW results also showed no association between the metabolites and CRC risk (GECCO OR: 1.03, 95% CI: 0.92–1.15, and DACHS OR: 1.06, 95% CI: 0.69–1.63) and no association was found upon stratification by aspirin use (users OR: 0.99, 95% CI: 0.38–2.57, non-users OR: 1.10, 95% CI: 0.66–1.84).

Using the alternative MR methods (weighted mode, weighted median, and MR Egger), no association between SA and CRC in both GECCO and DACHS was seen, regardless of stratification (Supplementary Table S10).

Since these four SNPs were all found on chromosome 16 (Supplementary Table S8), a WGLR method was applied to account for their correlation. We found a positive association between SA and CRC risk in the GECCO sample (OR: 1.13, 95% CI: 1.05–1.22) but no association in the DACHS sample (OR: 0.51, 95% CI: 0.16–1.67), DACHS aspirin users (OR: 0.12, 95% CI: 0.01–2.67) and DACHS aspirin non-users (OR: 0.70, 95% CI: 0.30–1.65)

(Figure 3C). As a sensitivity analysis, the heterogeneity of the results was assessed through a Q statistic, but no evidence of heterogeneity was seen (Supplementary Table S11).

We repeated the MR Steiger method using the genome-wide significant SNPs, and results again suggested that the causal direction was indeed from SA to CRC, as the SNPs explained more variation in SA levels than CRC risk (Supplementary Table S6).

4. Discussion

In this study, we aimed to assess whether increasing levels of SA affected risk of CRC, using an MR approach, and whether higher levels of SA proxied by pharmacological intervention in the form of aspirin use was required to identify an effect. Our analysis focused on aspirin, since 90% of the drug is rapidly deacetylated to form SA [16], which is the active metabolite of the drug [13,14], and therefore increases SA levels more than would be achieved through the diet. Three different approaches were applied to identify genetic variants (instrument variables) which could serve as proxies for SA and understand the causal nature of their role in determining CRC risk. The three approaches involved selecting (i) functional, (ii) pathway and (iii) genome-wide SNPs each associated with SA. The functional genetic variants were selected through the established role of the genes in aspirin metabolism from various sources of evidence. With regards to the pathway and genome-wide significant SNPs, all were found on chromosome 16, either within or near the coding region for the enzyme ACSM2B, which is the enzyme involved in breaking down SA into its metabolite salicyluric acid, thereby providing a plausible biological link between these SNPs and levels of SA.

We found no association between the functional SNPs and levels of SA, therefore did not take them forward to instrument SA levels. Using pathway and genome-wide SNPs, we identified two and one independent SNPs ($R^2 < 0.001$) to proxy for SA levels, respectively, and found no association between increasing metabolite levels and CRC risk using an IVW and MR RAPS approach, regardless of aspirin stratification. Furthermore, due to the small number of instruments, we applied a less stringent LD threshold ($R^2 < 0.8$) and identified six pathway SNPs and four genome-wide SNPs to proxy for an SD increase in SA levels. Using these SNPs, we found consistent null results using the IVW method and alternative MR methods (weighted median, weighted mode, and MR Egger). However, after accounting for SNP correlation using a WGLR method, we found that an SD increase in SA increased the risk of CRC in GECCO (OR: 1.11, 95% CI: 1.01–1.21, *p*-value: 0.03 and OR: 1.13, 95% CI: 1.05–1.22, *p*-value: 1.42×10^{-3}, respectively). We acknowledge that when the LD clumping threshold was relaxed to 0.8, there may have been some overlap with SNPs used in both the functional and the genome-wide analysis. We also acknowledge that the sensitivity analyses used were limited in detecting heterogeneity due to the low number of SNPs. The Cochran Q statistic also requires a large number of SNPs, otherwise there is little power to detect heterogeneity [59]. Overall, we found little evidence to suggest that SA affects risk of CRC, regardless of stratification.

It is thought that one reason why fruit and vegetable consumption may prevent CRC [8] is due to the presence of SA [60], although no formal RCTs have been carried out to confirm this. In vitro studies have also shown that salicylic acid inhibits the growth of colorectal cancer cells [61]. SA is the primary metabolite of aspirin, of which both observational and RCT evidence have shown aspirin as a chemopreventive agent [18–21]. Our MR results show little evidence of an association between the metabolite and CRC risk, regardless of aspirin use. However, we discuss some of the possible reasons why below.

Whilst we found no association between functional SNPs known to affect aspirin metabolism enzymes' activity and levels of SA, this may be due to a more complex relationship between genotype and metabolite levels, rather than the assumed linear additive model. For example, with regards to the functional SNPs, Nagar et al. (2004) identified that whilst individuals with homozygous mutant alleles of UGT1A6 had the highest metabolic activity, those that were heterozygous for alleles in three SNPs (including rs1105879 and rs2070959) were actually less active than homozygous wildtype enzymes [56], indicating

a non-linear association between the alleles and the metabolites, which is a common assumption made in regression analyses [62]. This non-linear association between alleles and enzyme activity needs to also be addressed between alleles and metabolite levels to derive instrumental variables.

To our knowledge, our GWAS for SA is the largest performed for this metabolite (n = 14,149), with others having much smaller sample sizes and not being publicly available. By using a much larger sample size, we were able to identify genome-wide significant associations that would have otherwise been missed in smaller studies. However, the variance explained and the strength of the instruments still indicated weak instruments despite strong associations with the metabolite. In order to improve the results and conclusions observed in this study, ideally we would need to identify the SNP associations with SA levels stratified between aspirin users and non-users, similar to what was carried out in our CRC outcome sample. However, to our knowledge, metabolite, genotype, and phenotype data (of aspirin use) are not currently large enough to run this analysis. If a stronger association exists between the SNPs and SA levels in aspirin users, this would provide more strength of the appropriateness of the genetic instruments used to proxy for SA levels.

We also acknowledge another limitation in this study is that the measurement of metabolites was through an untargeted metabolomics approach, and so the variables generated are assessed in units of measurement called "ion counts" which are calculated from the area under the curve of the corresponding peak in the mass spectrum. This means that metabolite measurements are quantitative values of relative changes as opposed to the absolute quantification of metabolite concentrations that can be achieved through targeted metabolomics [63]. For this reason, it is important to focus on the direction of effect and strength of association (p-values) in this study, as opposed to the magnitude of effect. This may have also impacted on the calculation of variance explained and the F statistics, which mostly indicate that the instrumental variables used in the MR were weak as they explain little of the variance, and the F-statistic is below the conventionally applied indicative threshold of 10 [64]. However, without carrying out a more targeted metabolomic approach and quantifying the exact effect of these SNPs on the metabolite levels, it is difficult to draw firm conclusions about the strength of the instruments used for MR.

Furthermore, larger sample sizes of recorded aspirin use are required as currently, our study may have been underpowered to detect an effect, hence explaining the null results using the IVW approach. Our power calculations show that in aspirin users, we had sufficient power to detect an effect of SA on CRC risk with an OR of $\leq$0.43 and $\geq$2.38, whereas observationally, the effect of aspirin on CRC risk is RR: 0.79 (95% CI: 0.74–0.85) [19], and in a long-term observational follow-up of a trail, the hazard ratio was 0.80 (95% CI: 0.67–0.97) [18]. Therefore, it would be useful to repeat this analysis in a larger sample with comprehensive data on aspirin use.

5. Conclusions

Overall, the analyses presented have shown that dietary levels of SA, as well as increased levels proxied by aspirin use, may be insufficient at reducing risk of CRC, although based on the variance explained in SA levels by our SNPs and the F statistic, we acknowledge that the analysis needs to be repeated again with stronger instruments that proxy the metabolite levels.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/nu13114164/s1, Table S1: Enzyme genomic regions based on NCBI Build 37/UCSC hg19, Table S2: Associations of the 4 functional SNPs with salicylic acid, Table S3: Pathway SNP associations with salicylic acid, Table S4: Pathway SNP associations with colorectal cancer, Table S5: Pathway SNP associations with CRC using the other MR methods, Table S6: Results of the Q statistic heterogeneity test for pathway SNPs, Table S7: MR Steiger test results, Table S8: Genome-wide SNP associations with salicylic acid, Table S9: Genome-wide SNP associations with colorectal cancer, Table S10:

Genome-wide SNP associations with CRC using the other MR methods, and Table S11: Results of the Q statistic heterogeneity test for genome-wide SNPs.

Author Contributions: Conceptualization, A.N., A.C.W. and C.L.R.; methodology. A.N., R.C.R., D.S., J.L.B., A.C.W. and C.L.R.; formal analysis, A.N.; investigation, A.N., R.C.R., A.C.W. and C.L.R.; data curation, A.N., I.D.S., P.S., N.J.W., A.B., S.J.W., D.A., J.A.B., J.L.H., J.C.F., P.A.N., N.M.L., G.C., E.A.P., L.L.M., C.M.U., C.I.L., F.J.B.v.D., A.G., P.T.C., V.M., P.V., L.V., E.A. (Efrat Amitay), E.A. (Elizabeth Alwers), J.C.-C., L.C.S., M.L.S., R.E.S., M.J.G., S.C.-B., H.-R.K., S.-S.K., A.T.C., L.L., W.Z., D.T.B., D.D.B., G.G.G., S.B.G., G.R., Z.K.S., T.A.H., Y.L., T.O.K., M.O.W., C.S., B.V.G., S.G., H.H., S.I.B., P.D.P.P., A.L., A.W., A.H.W., E.W., U.P., D.A.D., H.B. and M.H.; writing—original draft preparation, A.N.; writing—review and editing, A.N., R.C.R., I.D.S., S.J.W., J.A.B., J.C.F., E.A.P., F.J.B.v.D., P.T.C., P.V., L.V., E.A., J.C.-C., L.C.S., S.C.-B., A.T.C., B.V.G., E.W., U.P., D.A.D., H.B., M.H., A.C.W. and C.L.R.; supervision, R.C.R., A.C.W. and C.L.R.; funding acquisition A.C.W. and C.L.R. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by a PhD studentship from the Medical Research Council (AN), a Cancer Research UK Programme Grant (C19/A11975, ACW), an MRC Research Grant (MR/R017247/1, ACW) and by the John James Bristol Foundation. Further funding was provided by The UK Medical Research Council Integrative Epidemiology Unit (MC_UU_12013_2, CLR) and Cancer Research UK (C18281/A19169, CLR). RCR is a de Pass Vice Chancellor Research Fellow at the University of Bristol. PS is supported by a Rutherford Fund Fellowship from the Medical Research Council grant MR/S003746/1. DS was supported by the German Federal Ministry of Education and Research (01KT1510). Consortia funding: Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO): National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services (U01 CA164930, U01 CA137088, R01 CA059045, R21 CA191312, R01201407). Genotyping/Sequencing services were provided by the Center for Inherited Disease Research (CIDR) contract number HHSN268201700006I and HHSN268201200008I. 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. ASTERISK: 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). The ATBC Study is supported by the Intramural Research Program of the U.S. National Cancer Institute, National Institutes of Health. CLUE II funding was from the National Cancer Institute (U01 CA86308, Early Detection Research Network; P30 CA006973), National Institute on Aging (U01 AG18033), and the American Institute for Cancer Research. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government. Maryland Cancer Registry (MCR) cancer data was provided by the Maryland Cancer Registry, Center for Cancer Prevention and Control, Maryland Department of Health, with funding from the State of Maryland and the Maryland Cigarette Restitution Fund. The collection and availability of cancer registry data is also supported by the Cooperative Agreement NU58DP006333, funded by the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors, and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Department of Health and Human Services. ColoCare: This work was supported by the National Institutes of Health (grant numbers R01 CA189184 (Li/Ulrich), U01 CA206110 (Ulrich/Li/Siegel/Figueireido/Colditz, 2P30CA015704-40 (Gilliland), R01 CA207371 (Ulrich/Li)), the Matthias Lackas Foundation, the German Consortium for Translational Cancer Research, and the EU TRANSCAN initiative. The Colon Cancer Family Registry (CCFR, www.coloncfr.org, accessed on 16 November 2021) is supported in part by funding from the National Cancer Institute (NCI), National Institutes of Health (NIH) (award U01 CA167551). The CCFR Set-1 (Illumina 1M/1M-Duo) and Set-2 (Illumina Omni1-Quad) scans were supported by NIH awards U01 CA122839 and R01 CA143247 (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 HHSN268201200008I. The SCCFR Illumina HumanCytoSNP array was supported through NCI award R01 CA076366 (to PAN). Additional funding for the OFCCR/ARCTIC was through award GL201-043 from the Ontario Research Fund (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 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. COLON: The COLON study is sponsored by Wereld Kanker Onderzoek Fonds, including funds from grant 2014/1179 as part of the World Cancer Research Fund International Regular Grant Programme, by Alpe d'Huzes and the Dutch Cancer Society (UM 2012–5653, UW 2013-5927, UW2015-7946), and by TRANSCAN (JTC2012-MetaboCCC, JTC2013-FOCUS). The Nqplus study is sponsored by a ZonMW investment grant (98-10030); by PREVIEW, the project PREVention of diabetes through lifestyle intervention and population studies in Europe and around the World (PREVIEW) project which received funding from the European Union Seventh Framework Programme (FP7/2007–2013) under grant no. 312057; by funds from TI Food and Nutrition (cardiovascular health theme), a public–private partnership on precompetitive research in food and nutrition; and by FOODBALL, the Food Biomarker Alliance, a project from JPI Healthy Diet for a Healthy Life. Colorectal Cancer Transdisciplinary (CORECT) Study: The CORECT Study was supported by the National Cancer Institute, National Institutes of Health (NCI/NIH), U.S. Department of Health and Human Services (grant numbers U19 CA148107, R01 CA81488, P30 CA014089, R01 CA197350; P01 CA196569; R01 CA201407) and National Institutes of Environmental Health Sciences, National Institutes of Health (grant number T32 ES013678). CORSA: "Österreichische Nationalbank Jubiläumsfondsprojekt" (12511) and Austrian Research Funding Agency (FFG) grant 829675. CPS-II: The American Cancer Society funds the creation, maintenance, and updating of the Cancer Prevention Study-II (CPS-II) cohort. This study was conducted with Institutional Review Board approval. CRCGEN: Colorectal Cancer Genetics & Genomics, Spanish 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), Agency for Management of University and Research Grants (AGAUR) of the Catalan Government (grant 2017SGR723), and Junta de Castilla y León (grant LE22A10-2). Sample collection of this work was supported by the Xarxa de Bancs de Tumors de Catalunya sponsored by Pla Director d'Oncología de Catalunya (XBTC), Plataforma Biobancos PT13/0010/0013 and ICOBIOBANC, sponsored by the Catalan Institute of Oncology. Czech Republic CCS: This work was supported by the Czech Science Foundation (20-03997S) and by the Grant Agency of the Ministry of Health of the Czech Republic (grants NV18/03/00199 and NU21-07-00247). 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 01GL1712). The sponsors had no role in the study design and the collection, analysis, and interpretation of data. DALS: National Institutes of Health (R01 CA48998 to M. L. Slattery). EDRN: This work is funded and supported by the NCI, EDRN Grant (U01 CA 84968-06). EPIC: The coordination of EPIC is financially supported by the European Commission (DGSANCO) and the International Agency for Research on Cancer. The national cohorts are supported by Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF), Deutsche Krebshilfe, Deutsches Krebsforschungszentrum and Federal Ministry of Education and Research (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRCItaly and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); ERC-2009-AdG 232997 and Nordforsk, Nordic Centre of Excellence programme on Food, Nutrition and Health (Norway); Health Research Fund (FIS), PI13/00061 to Granada, PI13/01162 to EPIC-Murcia, Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, ISCIII RETIC (RD06/0020) (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPICOxford) (United Kingdom). The EPIC-Norfolk study (https://doi.org/10.22025/2019.10.105.00004, accessed on 16 November 2021) has received funding from the Medical Research Council (MR/N003284/1, MC-UU_12015/1 and MC_UU_00006/1) and

Cancer Research UK (C864/A14136). The genetics work in the EPIC-Norfolk study was funded by the Medical Research Council (MC_PC_13048). Metabolite measurements in the EPIC-Norfolk study were supported by the MRC Cambridge Initiative in Metabolic Science (MR/L00002/1) and the Innovative Medicines Initiative Joint Undertaking under EMIF grant agreement no. 115372. EPICOLON: This work was supported by grants from Fondo de Investigación Sanitaria/FEDER (PI08/0024, PI08/1276, PS09/02368, PI11/00219, PI11/00681, PI14/00173, PI14/00230, PI17/00509, 17/00878, PI20/00113, PI20/00226, Acción Transversal de Cáncer), Xunta de Galicia (PGIDIT07PXIB9101209PR), Ministerio de Economia y Competitividad (SAF07-64873, SAF 2010-19273, SAF2014-54453R), Fundación Científica de la Asociación Española contra el Cáncer (GCB13131592CAST), Beca Grupo de Trabajo "Oncología" AEG (Asociación Española de Gastroenterología), Fundación Privada Olga Torres, FP7 CHIBCHA Consortium, Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR, Generalitat de Catalunya, 2014SGR135, 2014SGR255, 2017SGR21, 2017SGR653), Catalan Tumour Bank Network (Pla Director d'Oncologia, Generalitat de Catalunya), PERIS (SLT002/16/00398, Generalitat de Catalunya), CERCA Programme (Generalitat de Catalunya) and COST Actions BM1206 and CA17118. CIBERehd is funded by the Instituto de Salud Carlos III. ESTHER/VERDI. This work was supported by grants from the Baden-Württemberg Ministry of Science, Research and Arts and the German Cancer Aid. Harvard cohorts (HPFS, NHS, PHS): HPFS is supported by the National Institutes of Health (P01 CA055075, UM1 CA167552, U01 CA167552, R01 CA137178, R01 CA151993, R35 CA197735, K07 CA190673, and P50 CA127003), NHS by the National Institutes of Health (R01 CA137178, P01 CA087969, UM1 CA186107, R01 CA151993, R35 CA197735, K07CA190673, and P50 CA127003) and PHS by the National Institutes of Health (R01 CA042182). Hawaii Adenoma Study: NCI grants R01 CA72520. HCES-CRC: the Hwasun Cancer Epidemiology Study–Colon and Rectum Cancer (HCES-CRC; grants from Chonnam National University Hwasun Hospital, HCRI15011-1). Interval: Participants in the INTERVAL randomised controlled trial were recruited with the active collaboration of NHS Blood and Transplant England (www.nhsbt.nhs.uk, accessed on 16 November 2021), which has supported field work and other elements of the trial. DNA extraction and genotyping was co-funded by the National Institute for Health Research (NIHR), the NIHR BioResource (http://bioresource.nihr.ac.uk, accessed on 16 November 2021) and the NIHR [Cambridge Biomedical Research Centre at the Cambridge University Hospitals NHS Foundation Trust] [*]. Metabolon Metabolomics assays were funded by the NIHR BioResource and the NIHR [Cambridge Biomedical Research Centre at the Cambridge University Hospitals NHS Foundation Trust] [*]. The academic coordinating centre for INTERVAL was supported by core funding from: NIHR Blood and Transplant Research Unit in Donor Health and Genomics (NIHR BTRU-2014-10024), UK Medical Research Council (MR/L003120/1), British Heart Foundation (SP/09/002; RG/13/13/30194; RG/18/13/33946) and the NIHR (Cambridge Biomedical Research Centre at the Cambridge University Hospitals NHS Foundation Trust). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. This work was supported by Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation and Wellcome. Kentucky: This work was supported by the following grant support: Clinical Investigator Award from Damon Runyon Cancer Research Foundation (CI-8); NCI R01CA136726. LCCS: The Leeds Colorectal Cancer Study was funded by the Food Standards Agency and Cancer Research UK Programme Award (C588/A19167). Melbourne Collaborative Cohort Study (MCCS) cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further augmented by Australian National Health and Medical Research Council grants 209057, 396414 and 1074383 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry and the Australian Institute of Health and Welfare, including the National Death Index and the Australian Cancer Database. Multiethnic Cohort (MEC) Study: National Institutes of Health (R37 CA54281, P01 CA033619, R01 CA063464 and U01 CA164973). MECC: This work was supported by the National Institutes of Health, U.S. Department of Health and Human Services (R01 CA81488 to SBG and GR). MSKCC: The work at Sloan Kettering in New York was supported by the Robert and Kate Niehaus Center for Inherited Cancer Genomics and the Romeo Milio Foundation. Moffitt: This work was supported by funding from the National Institutes of Health (grant numbers R01 CA189184, P30 CA076292), Florida Department of Health Bankhead-Coley Grant 09BN-13, and the University of South Florida Oehler Foundation. Moffitt contributions were supported in part by the

Total Cancer Care Initiative, Collaborative Data Services Core, and Tissue Core at the H. Lee Moffitt Cancer Center & Research Institute, a National Cancer Institute-designated Comprehensive Cancer Center (grant number P30 CA076292). NCCCS I & II: We acknowledge funding support for this project from the National Institutes of Health, R01 CA66635 and P30 DK034987. NFCCR: This work was supported by an Interdisciplinary Health Research Team award from the Canadian Institutes of Health Research (CRT 43821); the National Institutes of Health, U.S. Department of Health and Human Services (U01 CA74783); and National Cancer Institute of Canada grants (18223 and 18226). 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. Funding was provided to Michael O. Woods by the Canadian Cancer Society Research Institute. NSHDS: Swedish Research Council; Swedish Cancer Society; Cutting-Edge Research Grant and other grants from Region Västerbotten; Knut and Alice Wallenberg Foundation; Lion's Cancer Research Foundation at Umeå University; the Cancer Research Foundation in Northern Sweden; and the Faculty of Medicine, Umeå University, Umeå, Sweden. OFCCR: The Ontario Familial Colorectal Cancer Registry was supported in part by the National Cancer Institute (NCI) of the National Institutes of Health (NIH) under award U01 CA167551 and award U01/U24 CA074783 (to SG). Additional funding for the OFCCR and ARCTIC testing and genetic analysis was through and a Canadian Cancer Society CaRE (Cancer Risk Evaluation) program grant and Ontario Research Fund award GL201-043 (to BWZ), through the Canadian Institutes of Health Research award 112746 (to TJH), and through generous support from the Ontario Ministry of Research and Innovation. OSUMC: OCCPI funding was provided by Pelotonia and HNPCC funding was provided by the NCI (CA16058 and CA67941). PLCO: 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. Funding was provided by National Institutes of Health (NIH), Genes, Environment and Health Initiative (GEI) Z01 CP 010200, NIH U01 HG004446, and NIH GEI U01 HG 004438. SCCFR: The Seattle Colon Cancer Family Registry was supported in part by the National Cancer Institute (NCI) of the National Institutes of Health (NIH) under awards U01 CA167551, U01 CA074794 (to JDP), and awards U24 CA074794 and R01 CA076366 (to PAN). SEARCH: The University of Cambridge has received salary support in respect of PDPP from the NHS in the East of England through the Clinical Academic Reserve. Cancer Research UK (C490/A16561); the UK National Institute for Health Research Biomedical Research Centres at the University of Cambridge. SELECT: Research reported in this publication was supported in part by the National Cancer Institute of the National Institutes of Health under Award Numbers U10 CA37429 (CD Blanke), and UM1 CA182883 (CM Tangen/IM Thompson). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. SMS: This work was supported by the National Cancer Institute (grant P01 CA074184 to J.D.P. and P.A.N., grants R01 CA097325, R03 CA153323, and K05 CA152715 to P.A.N., and the National Center for Advancing Translational Sciences at the National Institutes of Health (grant KL2 TR000421 to A.N.B.-H.). The Swedish Low-Risk Colorectal Cancer Study: The study was supported by grants from the Swedish research council; K2015-55X-22674-01-4, K2008-55X-20157-03-3, K2006-72X-20157-01-2 and the Stockholm County Council (ALF project). Swedish Mammography Cohort and Cohort of Swedish Men: This work is supported by the Swedish Research Council /Infrastructure grant, the Swedish Cancer Foundation, and the Karolinska Institute's Distinguished Professor Award to Alicja Wolk. UK Biobank: This research has been conducted using the UK Biobank Resource under Application Number 8614. VITAL: National Institutes of Health (K05 CA154337). WHI: The WHI program is funded 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.

Institutional Review Board Statement: All participants for INTERVAL, CCFR, CORECT, GECCO and UK Biobank provided written informed consent, and each study was approved by the relevant research ethics committee or institutional review board. The EPIC-Norfolk study was approved by the Norwich Local Ethics Committee (previously known as Norwich District Ethics Committee) (REC Ref: 98CN01); all participants gave their informed written consent before entering the study.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Summary data is available upon request by contacting the respective studies.

Acknowledgments: ASTERISK: We are very grateful to 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 controls, as well as all the physicians, technicians and students. 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). CLUE II: We thank the participants of Clue II and appreciate the continued efforts of the staff at the Johns Hopkins George W. Comstock Center for Public Health Research and Prevention in the conduct of the Clue II Cohort Study. COLON and NQplus: the authors would like to thank the COLON and NQplus investigators at Wageningen University & Research and the involved clinicians in the participating hospitals. CORSA: We kindly thank all those who contributed to the screening project Burgenland against CRC. Furthermore, we are grateful to Doris Mejri and Monika Hunjadi for laboratory assistance. CPS-II: The authors 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. Czech Republic CCS: We are thankful to all clinicians in major hospitals in the Czech Republic, without whom the study would not be practicable. We are also sincerely grateful to all patients participating in this study. DACHS: We thank all participants and cooperating clinicians, and Ute Handte-Daub, Utz Benscheid, Muhabbet Celik and Ursula Eilber for excellent technical assistance. EDRN: We acknowledge all the following contributors to the development of the resource: University of Pittsburgh School of Medicine, Department of Gastroenterology, Hepatology and Nutrition: Lynda Dzubinski; University of Pittsburgh School of Medicine, Department of Pathology: Michelle Bisceglia; and University of Pittsburgh School of Medicine, Department of 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. The EPIC-Norfolk study: we are grateful to all the participants who have been part of the project and to the many members of the study teams at the University of Cambridge who have enabled this research. EPICOLON: We are sincerely grateful to all patients participating in this study who were recruited as part of the EPICOLON project. We acknowledge the Spanish National DNA Bank, Biobank of Hospital Clínic–IDIBAPS and Biobanco Vasco for the availability of the samples. The work was carried out (in part) at the Esther Koplowitz Centre, Barcelona. Harvard cohorts (HPFS, NHS, PHS): 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 would like to thank the participants and staff of the HPFS, NHS and PHS 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. The authors assume full responsibility for analyses and interpretation of these data. Interval: A complete list of the investigators and contributors to the INTERVAL trial is provided in reference [33]. The academic coordinating centre would like to thank blood donor centre staff and blood donors for participating in the INTERVAL trial. Kentucky: We would like to acknowledge the staff at the Kentucky Cancer Registry. LCCS: We acknowledge the contributions of Jennifer Barrett, Robin Waxman, Gillian Smith and Emma Northwood in conducting this study. NCCCS I & II: We would like to thank the study participants, and the NC Colorectal Cancer Study staff. NSHDS investigators thank the Biobank Research Unit at Umeå University, the Västerbotten Intervention Programme, the Northern Sweden MONICA study and Region Västerbotten for providing data and samples and acknowledge the contribution from Biobank Sweden, supported by the Swedish Research Council (VR 2017-00650). PLCO: The authors thank the PLCO Cancer Screening Trial screening centre investigators and the staff from Information Management Services Inc and Westat Inc. Most importantly, we thank the study participants for their contributions that made this study possible. SCCFR: The authors would like to thank the study participants and staff of the Hormones and Colon Cancer and Seattle Cancer Family Registry studies (CORE Studies). SEARCH: We thank the SEARCH team. SELECT: We thank the research and clinical staff at the sites that participated on SELECT study, without whom the trial would not have been successful. We are also grateful to the 35,533 dedicated men who participated

in SELECT. UK Biobank: We would like to thank the participants and researchers of the UK Biobank for their participation and acquisition of data. WHI: The authors 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 (accessed on 16 November 2021).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Bowel Cancer Statistics | Cancer Research, UK. Available online: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer#heading-Zero (accessed on 13 April 2018).
2. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [CrossRef]
3. Cancer Research, UK. Bowel Cancer Incidence Trends over Time. Available online: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#heading-Two (accessed on 7 August 2018).
4. Siegel, R.L.; Fedewa, S.A.; Anderson, W.F.; Miller, K.D.; Ma, J.; Rosenberg, P.; Jemal, A. Colorectal Cancer Incidence Patterns in the United States, 1974–2013. *J. Natl. Cancer Inst.* **2017**, *109*. [CrossRef] [PubMed]
5. Malakar, S.; Gibson, P.R.; Barrett, J.S.; Muir, J.G. Naturally occurring dietary salicylates: A closer look at common Australian foods. *J. Food Compos. Anal.* **2017**, *57*, 31–39. [CrossRef]
6. Wood, A.; Baxter, G.; Thies, F.; Kyle, J.; Duthie, G. A systematic review of salicylates in foods: Estimated daily intake of a Scottish population. *Mol. Nutr. Food Res.* **2011**, *55*, S7–S14. [CrossRef] [PubMed]
7. Spadafranca, A.; Bertoli, S.; Fiorillo, G.; Testolin, G.; Battezzati, A. Circulating salicylic acid is related to fruit and vegetable consumption in healthy subjects. *Br. J. Nutr.* **2007**, *98*, 802–806. [CrossRef] [PubMed]
8. Aune, D.; Lau, R.; Chan, D.S.; Vieira, R.; Greenwood, D.C.; Kampman, E.; Norat, T. Nonlinear Reduction in Risk for Colorectal Cancer by Fruit and Vegetable Intake Based on Meta-analysis of Prospective Studies. *Gastroenterology* **2011**, *141*, 106–118. [CrossRef]
9. Aune, D.; Chan, D.S.M.; Lau, R.; Vieira, R.; Greenwood, D.C.; Kampman, E.; Norat, T. Dietary fibre, whole grains, and risk of colorectal cancer: Systematic review and dose-response meta-analysis of prospective studies. *BMJ* **2011**, *343*, d6617. [CrossRef] [PubMed]
10. Tatham, M.H.; Cole, C.; Scullion, P.; Wilkie, R.; Westwood, N.J.; Stark, L.A.; Hay, R.T. A Proteomic Approach to Analyze the Aspirin-mediated Lysine Acetylome. *Mol. Cell. Proteom.* **2017**, *16*, 310–326. [CrossRef] [PubMed]
11. Wu, K.K. Aspirin and salicylate: An old remedy with a new twist. *Circulation* **2000**, *102*, 2022–2023. [CrossRef]
12. Paterson, J.R.; Srivastava, R.; Baxter, G.J.; Graham, A.B.; Lawrence, J.R. Salicylic acid content of spices and its implications. *J. Agric. Food Chem.* **2006**, *54*, 2891–2896. [CrossRef]
13. Marimuthu, S.; Chivukula, R.S.V.; Alfonso, L.F.; Moridani, M.; Hagen, F.K.; Bhat, G.J. Aspirin acetylates multiple cellular proteins in HCT-116 colon cancer cells: Identification of novel targets. *Int. J. Oncol.* **2011**, *39*, 1273–1283. [PubMed]
14. Rang, H.; Dale, M.M. *Rang & Dale's Pharmacology*, 7th ed.; Elsevier Churchill Livingstone: Edinburgh, Scotland, 2012; 113p.
15. Rothwell, P.M.; Cook, N.R.; Gaziano, J.M.; Price, J.F.; Belch, J.J.; Roncaglioni, M.C.; Morimoto, T.; Mehta, Z. Effects of aspirin on risks of vascular events and cancer according to bodyweight and dose: Analysis of individual patient data from randomised trials. *Lancet* **2018**, *392*, 387–399. [CrossRef]
16. Agundez, J.; Martinez, C.; Perez-Sala, D.; Carballo, M.; Torres, M.; Garcia-Martin, E. Pharmacogenomics in Aspirin Intolerance. *Curr. Drug Metab.* **2009**, *10*, 998–1008. [CrossRef] [PubMed]
17. Drew, D.A.; Chan, A.T. Aspirin in the Prevention of Colorectal Neoplasia. *Annu. Rev. Med.* **2021**, *72*, 415–430. [CrossRef]
18. Cook, N.R.; Lee, I.-M.; Zhang, S.M.; Moorthy, M.V.; Buring, J.E. Alternate-Day, Low-Dose Aspirin and Cancer Risk: Long-Term Observational Follow-up of a Randomized Trial. *Ann. Intern. Med.* **2013**, *159*, 77–85. [CrossRef] [PubMed]
19. Qiao, Y.; Yang, T.; Gan, Y.; Li, W.; Wang, C.; Gong, Y.; Lu, Z. Associations between aspirin use and the risk of cancers: A meta-analysis of observational studies. *BMC Cancer* **2018**, *18*, 1–57. [CrossRef] [PubMed]
20. Rothwell, P.M.; Wilson, M.; Elwin, C.-E.; Norrving, B.; Algra, A.; Warlow, C.P.; Meade, T.W. Long-term effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials. *Lancet* **2010**, *376*, 1741–1750. [CrossRef]
21. Rothwell, P.M.; Fowkes, F.G.R.; Belch, J.J.; Ogawa, H.; Warlow, C.P.; Meade, T.W. Effect of daily aspirin on long-term risk of death due to cancer: Analysis of individual patient data from randomised trials. *Lancet* **2011**, *377*, 31–41. [CrossRef]
22. Needs, C.J.; Brooks, P.P.M. Clinical Pharmacokinetics of the Salicylates. *Clin. Pharmacokinet.* **1985**, *10*, 164–177. [CrossRef] [PubMed]
23. Law, B.K.; Waltner-Law, M.E.; Entingh, A.J.; Chytil, A.; Aakre, M.E.; Nørgaard, P.; Moses, H.L. Salicylate-induced Growth Arrest Is Associated with Inhibition of p70s6k and Down-regulation of c-Myc, Cyclin D1, Cyclin A, and Proliferating Cell Nuclear Antigen. *J. Biol. Chem.* **2000**, *275*, 38261–38267. [CrossRef]

24. Borthwick, G.M.; Johnson, A.S.; Partington, M.; Burn, J.; Wilson, R.; Arthur, H.M. Therapeutic levels of aspirin and salicylate directly inhibit a model of angiogenesis through a Cox- independent mechanism. *FASEB J.* **2006**, *20*, 2009–2016. [CrossRef]
25. Pathi, S.; Jutooru, I.; Chadalapaka, G.; Nair, V.; Lee, S.O.; Safe, S. Aspirin Inhibits Colon Cancer Cell and Tumor Growth and Downregulates Specificity Protein (Sp) Transcription Factors. *PLoS ONE* **2012**, *7*, e48208. [CrossRef] [PubMed]
26. Davey Smith, G.; Ebrahim, S. "Mendelian randomization": Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **2003**, *32*, 1–22. [CrossRef]
27. Davey Smith, G.; Hemani, G. Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.* **2014**, *23*, R89–R98. [CrossRef] [PubMed]
28. Lawlor, D.A.; Harbord, R.M.; Sterne, J.A.C.; Timpson, N.; Smith, G.D. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat. Med.* **2008**, *27*, 1133–1163. [CrossRef] [PubMed]
29. Davey Smith, G.; Ebrahim, S. What can Mendelian randomisation tell us about modifiable behavioural and environmental exposures? *Br. Med. J.* **2005**, *330*, 1076–1079. [CrossRef]
30. Nitsch, D.; Molokhia, M.; Smeeth, L.; DeStavola, B.L.; Whittaker, J.; Leon, D. Limits to Causal Inference based on Mendelian Randomization: A Comparison with Randomized Controlled Trials. *Am. J. Epidemiol.* **2006**, *163*, 397–403. [CrossRef]
31. Klein, E.A.; Thompson, I.; Tangen, C.M.; Lucia, M.S.; Goodman, P.; Minasian, L.M.; Ford, L.G.; Parnes, H.L.; Gaziano, J.M.; Karp, D.D.; et al. Vitamin E and the risk of prostate cancer: Updated results of the Selenium and Vitamin E Cancer Prevention Trial (SELECT). *J. Clin. Oncol.* **2012**, *30*, 7. [CrossRef]
32. Yarmolinsky, J.; Bonilla, C.; Haycock, P.C.; Langdon, R.J.Q.; Lotta, L.A.; Langenberg, C.; Relton, C.L.; Lewis, S.J.; Evans, D.M.; Smith, G.D.; et al. Circulating Selenium and Prostate Cancer Risk: A Mendelian Randomization Analysis. *J. Natl. Cancer Inst.* **2018**, *110*, 1035–1038. [CrossRef]
33. Day, N.; Oakes, S.; Luben, R.; Khaw, K.T.; Bingham, S.; Welch, A.; Wareham, N. EPIC-Norfolk: Study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br. J. Cancer* **1999**, *80* (Suppl. 1), 95–103.
34. Di Angelantonio, E.; Thompson, S.G.; Kaptoge, S.; Moore, C.; Walker, M.; Armitage, J.; Ouwehand, W.H.; Roberts, D.J.; Danesh, J.; Donovan, J.; et al. Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): A randomised trial of 45,000 donors. *Lancet* **2017**, *390*, 2360–2371. [CrossRef]
35. Loh, P.-R.; Tucker, G.J.; Bulik-Sullivan, B.K.; Vilhjalmsson, B.; Finucane, H.K.; Salem, R.; Chasman, D.I.; Ridker, P.M.; Neale, B.M.; Berger, B.; et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **2015**, *47*, 284–290. [CrossRef]
36. Willer, C.; Li, Y.; Abecasis, G.R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **2010**, *26*, 2190–2191. [CrossRef] [PubMed]
37. Staiger, D.; Stock, J.H. Instrumental Variables Regression with Weak Instruments Author. *Econometrica* **1997**, *65*, 557–586. [CrossRef]
38. mRnd: Power Calculations for Mendelian Randomization. Available online: http://cnsgenomics.com/shiny/mRnd/ (accessed on 14 February 2019).
39. Huyghe, J.R.; Bien, S.A.; Harrison, T.A.; Kang, H.M.; Chen, S.; Schmit, S.L.; Conti, D.V.; Qu, C.; Jeon, J.; Edlund, C.K.; et al. Discovery of common and rare genetic risk variants for colorectal cancer. *Nat. Genet.* **2019**, *51*, 76–87. [CrossRef] [PubMed]
40. Schumacher, F.; Schmit, S.L.; Jiao, S.; Edlund, C.K.; Wang, H.; Zhang, B.; Hsu, L.; Huang, S.-C.; Fischer, C.P.; Harju, J.F.; et al. Genome-wide association study of colorectal cancer identifies six new susceptibility loci. *Nat. Commun.* **2015**, *6*, 1–7. [CrossRef]
41. Bycroft, C.; Freeman, C.; Petkova, D.; Band, G.; Elliott, L.T.; Sharp, K.; Motyer, A.; Vukcevic, D.; Delaneau, O.; O'Connell, J.; et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **2018**, *562*, 203–209. [CrossRef] [PubMed]
42. Brenner, H.; Chang-Claude, J.; Seiler, C.M.; Rickert, A.; Hoffmeister, M. Protection from Colorectal Cancer After Colonoscopy: A Population-Based, Case-Control Study. *Ann. Intern. Med.* **2011**, *154*, 22–30. [CrossRef]
43. Brenner, H.; Chang–Claude, J.; Jansen, L.; Knebel, P.; Stock, C.; Hoffmeister, M. Reduced Risk of Colorectal Cancer Up to 10 Years After Screening, Surveillance, or Diagnostic Colonoscopy. *Gastroenterology* **2014**, *146*, 709–717. [CrossRef]
44. Amitay, E.L.; Carr, P.; Jansen, L.; Walter, V.; Roth, W.; Herpel, E.; Kloor, M.; Bläker, H.; Chang-Claude, J.; Brenner, H.; et al. Association of Aspirin and Nonsteroidal Anti-Inflammatory Drugs with Colorectal Cancer Risk by Molecular Subtypes. *J. Natl. Cancer Inst.* **2018**, *111*, 475–483. [CrossRef]
45. Hemani, G.; Zheng, J.; Elsworth, B.; Wade, K.H.; Haberland, V.; Baird, D.; Laurin, C.; Burgess, S.; Bowden, J.; Langdon, R.; et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **2018**, *7*, e34408. [CrossRef] [PubMed]
46. Burgess, S.; Butterworth, A.; Thompson, S.G. Mendelian Randomization Analysis with Multiple Genetic Variants Using Summarized Data. *Genet. Epidemiol.* **2013**, *37*, 658–665. [CrossRef] [PubMed]
47. Zhao, Q.; Wang, J.; Hemani, G.; Bowden, J.; Small, D.S. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *Ann. Stat.* **2020**, *48*, 1742–1769. [CrossRef]
48. Slob, E.A.; Burgess, S. A comparison of robust Mendelian randomization methods using summary data. *Genet Epidemiol.* **2020**. [CrossRef] [PubMed]
49. Hartwig, F.P.; Smith, G.D.; Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.* **2017**, *46*, 1985–1998. [CrossRef]

50. Bowden, J.; Smith, G.D.; Haycock, P.C.; Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* **2016**, *40*, 304–314. [CrossRef]
51. Bowden, J.; Smith, G.D.; Burgess, S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **2015**, *44*, 512–525. [CrossRef]
52. Greco, M.F.; Del Minelli, C.; Sheehan, N.A.; Thompson, J.R. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med.* **2015**, *34*, 2926–2940. [CrossRef]
53. Burgess, S.; Dudbridge, F.; Thompson, S.G. Combining information on multiple instrumental variables in Mendelian randomization: Comparison of allele score and summarized data methods. *Stat. Med.* **2015**, *35*, 1880–1906. [CrossRef]
54. Zhou, G.; Marathe, G.; Hartiala, J.; Hazen, S.L.; Allayee, H.; Tang, W.H.W.; McIntyre, T.M. Aspirin Hydrolysis in Plasma Is a Variable Function of Butyrylcholinesterase and Platelet-activating Factor Acetylhydrolase 1b2 (PAFAH1b2). *J. Biol. Chem.* **2013**, *288*, 11940–11948. [CrossRef]
55. Ciotti, M.; Marrone, A.; Potter, C.; Owens, I.S. Genetic polymorphism in the human UGT1A6 (planar phenol) UDP-glucuronosyltransferase: Pharmacological implications. *Pharmacogenetics* **1997**, *7*, 485–495. [CrossRef] [PubMed]
56. Nagar, S.; Zalatoris, J.J.; Blanchard, R.L. Human UGT1A6 pharmacogenetics: Identification of a novel SNP, characterization of allele frequencies and functional analysis of recombinant allozymes in human liver tissue and in cultured cells. *Pharmacogenetics* **2004**, *14*, 487–499. [CrossRef] [PubMed]
57. Crespi, C.L.; Miller, V.P. The R144C change in the CYP2C9*2 allele alters interaction of the cytochrome P450 with NADPH:cytochrome P450 oxidoreductase. *Pharmacogenetics* **1997**, *7*, 203–210. [CrossRef] [PubMed]
58. Hemani, G.; Tilling, K.; Davey Smith, G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* **2017**, *13*, e1007081.
59. von Hippel, P.T. The heterogeneity statistic I2 can be biased in small meta-analyses. *BMC Med. Res. Methodol.* **2015**, *15*, 35. [CrossRef]
60. Paterson, J.R.; Lawrence, J.R. Salicylic acid: A link between aspirin, diet and the prevention of colorectal cancer. *QJM Int. J. Med.* **2001**, *94*, 445–448. [CrossRef]
61. Dachineni, R.; Ai, G.; Kumar, D.R.; Sadhu, S.S.; Tummala, H.; Bhat, G.J. Cyclin A2 and CDK2 as Novel Targets of Aspirin and Salicylic acid: A Potential Role in Cancer Prevention. *Mol. Cancer Res.* **2016**, *14*, 241–252. [CrossRef]
62. Sheehan, N.A.; Didelez, V. Epidemiology, genetic epidemiology and Mendelian randomisation: More need than ever to attend to detail. *Qual. Life Res.* **2019**, *139*, 121–136. [CrossRef]
63. Kapoore, R.V.; Vaidyanathan, S. Towards quantitative mass spectrometry-based metabolomics in microbial and mammalian systems. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2016**, *374*, 20150363. [CrossRef]
64. Pierce, B.L.; Ahsan, H.; VanderWeele, T.J. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int. J. Epidemiol.* **2010**, *40*, 740–752. [CrossRef]