BJC British Journal of Cancer

www.nature.com/bjc

ARTICLE

Epidemiology

# Lack of an association between gallstone disease and bilirubin levels with risk of colorectal cancer: a Mendelian randomisation analysis

Richard Culliford[1], Alex J. Cornish[1], Philip J. Law[1], Susan M. Farrington[2], Kimmo Palin[3], Mark A. Jenkins[4], Graham Casey[5], Michael Hoffmeister[6], Hermann Brenner[6,7,8], Jenny Chang-Claude[9,10], Iva Kirac[11], Tim Maughan[12], Stefanie Brezina[13], Andrea Gsur[13], Jeremy P. Cheadle[14], Lauri A. Aaltonen[3], Malcom G. Dunlop[2] and Richard S. Houlston[1]

**BACKGROUND:** Epidemiological studies of the relationship between gallstone disease and circulating levels of bilirubin with risk of developing colorectal cancer (CRC) have been inconsistent. To address possible confounding and reverse causation, we examine the relationship between these potential risk factors and CRC using Mendelian randomisation (MR).

**METHODS:** We used two-sample MR to examine the relationship between genetic liability to gallstone disease and circulating levels of bilirubin with CRC in 26,397 patients and 41,481 controls. We calculated the odds ratio per genetically predicted SD unit increase in log bilirubin levels ($OR_{SD}$) for CRC and tested for a non-zero causal effect of gallstones on CRC. Sensitivity analysis was applied to identify violations of estimator assumptions.

**RESULTS:** No association between either gallstone disease ($P$ value = 0.60) or circulating levels of bilirubin ($OR_{SD}$ = 1.00, 95% confidence interval (CI) = 0.96–1.03, $P$ value = 0.90) with CRC was shown.

**CONCLUSIONS:** Despite the large scale of this study, we found no evidence for a causal relationship between either circulating levels of bilirubin or gallstone disease with risk of developing CRC. While the magnitude of effect suggested by some observational studies can confidently be excluded, we cannot exclude the possibility of smaller effect sizes and non-linear relationships.

*British Journal of Cancer* (2021) 124:1169–1174; https://doi.org/10.1038/s41416-020-01211-x

## BACKGROUND

Gallstone disease has been proposed to increase the risk of colorectal cancer (CRC); however, evidence for a causal relationship from epidemiological studies is lacking.[1,2] The observational nature of studies has made them vulnerable to confounding, from measured and unmeasured risk factors and reverse causation. It is especially noteworthy that many of the risk factors for gallstone disease include factors that are well-established as risk factors for CRC (obesity, high energy intake, alcohol consumption and diabetes).[3,4] Increased production of bilirubin, the metabolic by-product of haemoglobin degradation, is also associated with gallstone disease.[5–10] This is intriguing as bilirubin has anti-oxidant and anti-inflammatory attributes, with experimental studies reporting mildly elevated levels being associated with decreased oxidative stress-related disease, including cancer.[11–13]

Mendelian randomisation (MR) is an analytic methodology in which germline genetic variants are used as proxies, or instrumental variables (IVs), for putative risk factors. We assume that these genetic variants follow Mendel's laws of inheritance, where the variants are randomly assorted at conception and are not in linkage disequilibrium (LD) with one another. These genetic variants are not influenced by reverse causation. The variants can also provide unconfounded estimates of disease risk, provided they do not directly influence the disease without complete mediation of the risk factor and are not associated with confounders of the relationship between risk factor and disease.[14–19] Confounding can, however, be introduced to MR by population stratification, assortative mating and dynastic effects.[20]

To address the shortcomings of conventional observational epidemiological studies, we investigated whether genetic liability to gallstone disease and circulating levels of bilirubin are causally

[1]Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK; [2]Cancer Research UK Edinburgh Centre and Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK; [3]Medicum and Genome-Scale Biology Research Program, Research Programs Units, Department of Medical and Clinical Genetics, University of Helsinki, Helsinki, Finland; [4]Centre for Epidemiology and Biostatistics, University of Melbourne, Melbourne, VIC, Australia; [5]Centre for Public Health Genomics, University of Virginia, Virginia, VA, USA; [6]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany; [7]Division of Preventive Oncology, German Cancer Research Center, Heidelberg, Germany; [8]German Cancer Consortium, German Cancer Research Center, Heidelberg, Germany; [9]Unit of Genetic Epidemiology, German Cancer Research Center, Heidelberg, Germany; [10]Cancer Epidemiology Group, University Medical Center Hamburg-Eppendorf, University Cancer Center Hamburg, Hamburg, Germany; [11]Department of Surgical Oncology, University Hospital for Tumours, Sestre milosrdnice University Hospital Centre, Zagreb, Croatia; [12]Department of Oncology, University of Oxford, Oxford, UK; [13]Institute of Cancer Research, Department of Medicine I, Medical University of Vienna, Vienna, Austria and [14]Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff, UK
Correspondence: Richard Culliford (richard.culliford@icr.ac.uk)

Received: 14 May 2020 Revised: 9 October 2020 Accepted: 25 November 2020
Published online: 7 January 2021

associated with CRC using MR. Specifically, we used two-sample MR, in which genetic variants associated with relevant risk factors as instrumental variables were first identified from genome-wide association studies (GWAS). We then assessed the association between these instrumental variables and CRC in a large GWAS.

## METHODS

### Procedures

The genetic instruments, i.e., single-nucleotide polymorphisms (SNPs) to be used as IVs, were identified from recent meta-analyses or the largest GWAS published to date. Details on each meta-analysis are provided in Supplementary Tables 1 and 2. Briefly, summary statistics for gallstone disease were obtained from a meta-analysis of UK and Icelandic biobank data[21,22] comprising 27,174 cases and 736,838 controls;[23] effect sizes were not adjusted for covariates. Summary statistics for log bilirubin levels were obtained from a meta-analysis of three GWAS comprising 9464 individuals of Europeans ancestry;[24] SNP-exposure effect sizes were age- and sex-adjusted.

To examine the association of each genetic instrument with CRC risk, we made use of summary effect estimates and corresponding standard errors (SEs) from a recent meta-analysis of 15 GWAS by Law et al.[25,26] After imputation, this meta-analysis related over 10 million genetic variants to CRC in individuals of European ancestry. One of the 15 GWAS was an analysis of UK BioBank data. Since the published gallstone disease meta-analysis included UK BioBank data, we recomputed CRC association statistics excluding UK BioBank data to avoid bias from sample overlap. The 14 studies provided data on 26,397 patients with CRC and 41,481 controls. Aside from principal component adjustment for residual population substructure, none of the contributing GWAS were adjusted for covariates (Supplementary Table 3). Ethical approval for this study was not required as all data came from summary statistics and no individual-level data were used.

We only considered SNPs associated with each trait at $P$ value of less than $5 \times 10^{-8}$ with minor allele frequency greater than 0.01 in Europeans as potential instruments. We excluded SNPs with poor imputation quality (info score <0.8). For each SNP, the chromosome position was recovered, and effect estimates harmonised according to their effect alleles and corresponding frequencies in both datasets (failure in harmonisation, such as palindromic alleles, resulted in those SNPs being removed). Effect estimates were expressed in standard deviations (SD) of the trait per allele, along with the corresponding SE. While Ferkingstad et al.[23] used a 'multiplicative allele' model in their analysis of gallstone disease, we assumed association estimates equated to those that would have been obtained using an additive model (an assumption our MR estimator applies). To avoid co-linearity between SNPs for each trait, correlated SNPs were excluded (LD threshold, $r^2 \geq 0.01$, with LD estimates for the European population obtained from MR-Base[17]) within each trait, with SNPs with the strongest effect size retained. Post filtering (Fig. 1), 26 and 2 SNPs were used as IVs for gallstone disease and circulating bilirubin, respectively (Supplementary Table 4).

### Statistical analysis

The MR methodology assumes that genetic variants used as instruments for a risk factor are associated with only the risk factor

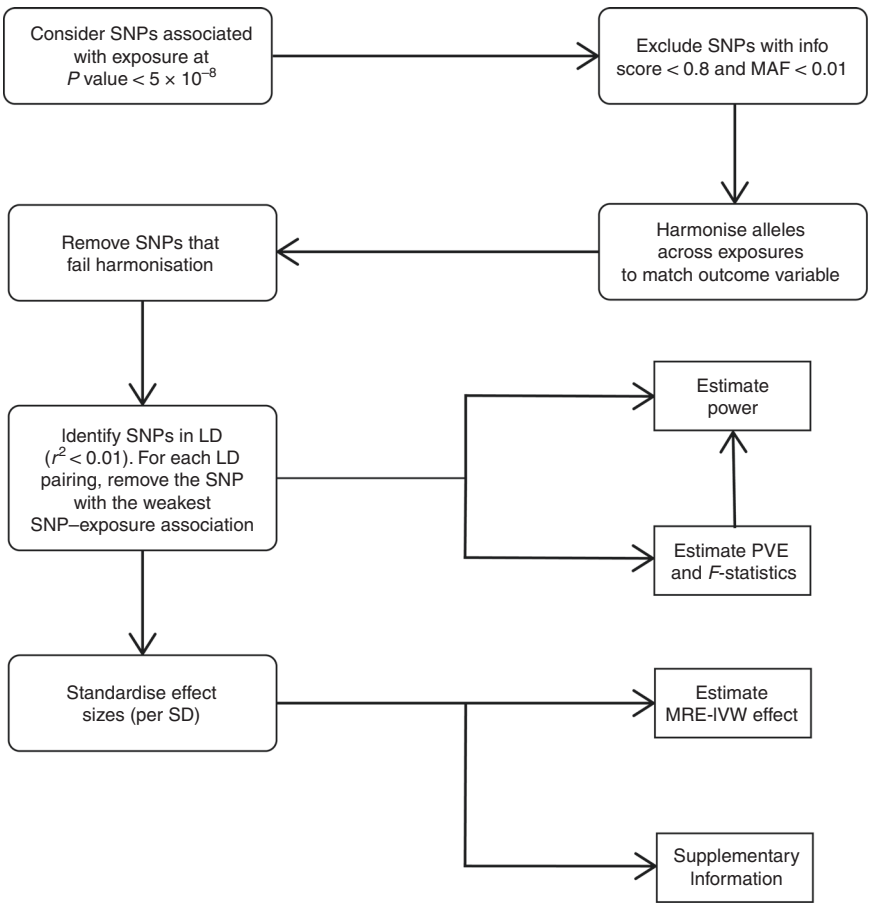


**Fig. 1 Flowchart of single-nucleotide polymorphism filtering.** EA effect allele, MAF minor allele frequency, LD linkage disequilibrium, PVE, proportion of variation explained, SD standard deviation, MRE-IVW multiplicative random effect inverse variance weighted.

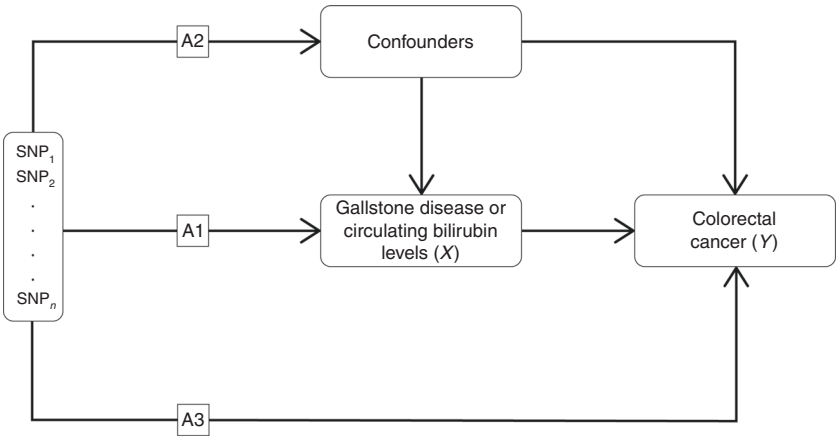


**Fig. 2 Mendelian randomisation (MR) and the key assumptions required to obtain an unbiased estimate of the causal effect.** (A1) genetic variants used as instrumental variables are only associated with the modifiable risk factor (X, gallstone disease or circulating bilirubin levels); (A2) there exists no instrument-outcome confounding including, but not limited to, conventional confounders of the exposure-outcome relationship; (A3) genetic variants only influence the risk of CRC (Y) through the risk factor (X).

and not with any confounders or another causal pathway (Fig. 2).[15–18,27,28]

To examine for a causal relationship between genetic liability to circulating levels of bilirubin and gallstone disease with CRC, we used a multiplicative random effect inverse variance weighted estimator (MRE-IVW)[28,29] as our primary measurement. We adopted this method as it is robust in the presence of pleiotropic effects, provided any heterogeneity is balanced or normally distributed (centred at zero) and that the instrument strength independent of direct effect (InSIDE) and no measurement error (NoME) assumption is met.[19,28] The results for the relationship between the levels of circulating bilirubin and CRC are reported as odds ratios per genetically predicted standard deviation unit increase ($OR_{SD}$) and 95% confidence intervals (CIs). Two-sample MR analysis of the association between a binary risk factor (i.e., gallstone disease) and a binary outcome (i.e., CRC) is problematic as estimates can be positively biased should there exist a true causal association, although bias becomes less of an issue as the true causal effect tends to unity.[30,31] In view of this when testing the association between gallstone disease and CRC, we primarily consider whether there exists a significant non-zero effect, and only report ORs for completeness.[32,33] We assessed evidence of heterogeneity by Cochran's $Q$,[28,34,35] $I^2$ statistics[35,36] and by examining funnel and forest plots of Wald ratio estimates for SNPs. To test the robustness of findings, we also derived weighted median[37] and weighted mode[38] MR estimators. Evidence of directional pleiotropy was assessed using MR-Egger regression.[39,40] We also formally assessed the effect of SNP heterogeneity on the MRE-IVW estimate, by implementing MR-PRESSO outlier and distortion tests.[41]

A Bonferroni-corrected $P$ value of 0.025 (i.e., 0.05/2 putative risk factors) was considered significant, with a $0.025 < P$ value $< 0.05$ being considered suggestive of a causal association. The power of our MR analysis to demonstrate causal effects was estimated considering the proportion of variance (PVE) in the risk factor explained by the genetic instruments,[42,43] stipulating a $P$ value of 0.05. We assumed that the PVE explained by all variants combined was approximately equal to the sum of the individual PVEs as per Shim et al..[44] Since calculation of power for binary exposure and binary outcome in a two-sample setting is problematic, we did not estimate study power for gallstone disease. Bias from weak instruments, or violation of the strong variant–risk-factor association assumption, affecting bilirubin effect estimates, were analysed using the estimated total PVE, the mean $F$ statistic and the Staiger–Stock rule.[45] Analyses were conducted using R v3.4.0[46] and the TwoSampleMR R-package.[17]

**Table 1.** Two-sample Mendelian randomisation analysis of the relationship between circulating levels of bilirubin and gallstone disease with risk of colorectal cancer.

| | $OR_{SD}$ | CI (95%) | $P$ value |
|---|---|---|---|
| *Bilirubin* | | | |
| MRE-IVW | 1.00 | (0.96, 1.03) | 0.90 |
| *Gallstone disease* | *OR** | | |
| MRE-IVW | 1.23 | (0.57, 2.66) | 0.60 |
| WME | 0.71 | (0.41, 1.24) | 0.23 |
| WMO | 0.68 | (0.39, 1.19) | 0.19 |
| MR-Egger | 0.52 | (0.12, 2.27) | 0.4 |

MRE-IVW multiplicative random effects inverse variance weighted, WME weighted median, WMO weighted mode, CI (95%) 95% confidence interval, $OR_{SD}$ odds ratio per genetically predicted standard deviation unit increase in bilirubin, $OR_*$ odds ratio per genetically predicted standard deviation (of the log odds of gallstone disease) log unit increase of the risk of gallstone disease, caveated by issues discussed in the main text.

## RESULTS

Analysis on the basis of MRE-IVW provided no evidence to support a causal association between gallstone disease and CRC ($P = 0.60$). MR-Egger showed no evidence to suggest directional pleiotropy ($P = 0.20$). Following on from MRE-IVW, we examined the relationship between gallstone disease and CRC under weighted median, weighted mode and MR-Egger regression methodologies, all providing little evidence to support a causal association ($P$ values of 0.23, 0.19 and 0.4, respectively, Table 1, Supplementary Tables 5–6, Figs. 3–4, Supplementary Fig. 1). There was evidence of heterogeneity between the SNPs used as IVs for gallstone disease (Cochran's $Q$ $P = 6.57 \times 10^{-23}$, $I^2 = 85\%$), and the outlier test from MR-PRESSO identified six SNPs being responsible for this heterogeneity (Supplementary Table 4). Effect re-estimation excluding the outlying SNPs showed no evidence to support a causal association ($P = 0.37$). There was also no evidence to indicate that outliers were causing a large distortion in the causal association estimate ($P = 0.77$, Supplementary Tables 5–6, Fig. 4 and Supplementary Fig. 2).

The mean $F$ statistic for the IVs used for genetically determined circulating level of bilirubin was 1038, corresponding to a PVE of 18%. Hence, the risk of weak instrument bias and violation of the

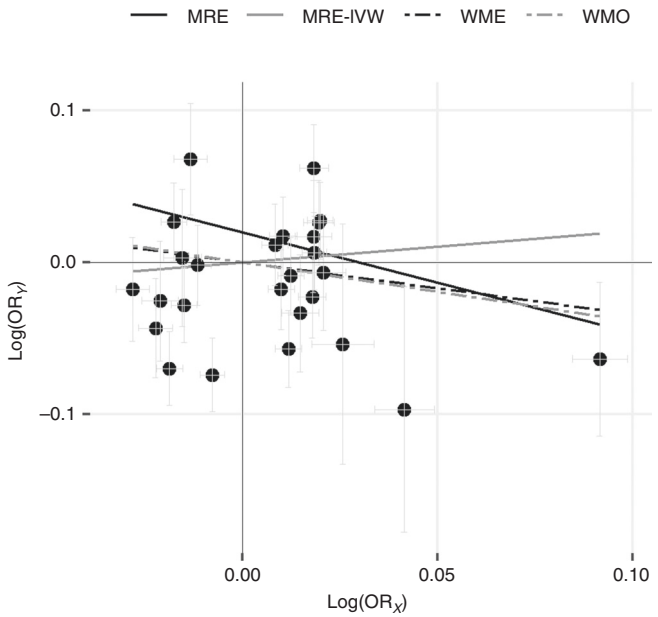


**Fig. 3 Single-nucleotide polymorphism (SNP) exposure association estimates for gallstone disease risk against the SNP-outcome association estimates for colorectal cancer risk.** Causal effect given by each Mendelian randomisation (MR) estimators, caveated by issues discussed in the main text. MRE-IVW multiplicative random effects inverse variance weighted, MRE MR-Egger, WME weighted median, WMO weighted mode, log(OR$_X$) log odds ratio (OR) per genetically predicted standard deviation increase in the exposure for each allele, log(OR$_Y$) log odds ratio in the outcome for each additional allele.

NOME assumption was low. Specifically, we had 80% power to detect an OR$_{SD}$ of 1.05 and over 90% power to detect an OR$_{SD}$ of 1.10 for CRC (Supplementary Tables 4 and 5). Nevertheless, there was no evidence for a causal association between levels of circulating bilirubin and CRC risk using MRE-IVW methodology (OR$_{SD}$ = 1.00, 95% CI = 0.96–1.03, $P$ = 0.90, Table 1, Fig. 4). As only two IVs were used, other estimators such as MR-Egger could not be applied.

## DISCUSSION

There has been interest in exploring a possible relationship between gallstone disease and risk of CRC since an association could be reflective of pathways linking bilirubin/bile acid metabolism, sex hormones and cholesterol metabolism. Previous epidemiological evidence for an association between gallstone disease and CRC has been inconclusive. Chen et al. reported that individuals with cholelithiasis had a higher risk of developing CRC (adjusted hazards ratio (AHR) = 1.36, 95% CI = 1.12–1.66). The risk increased for patients with cholelithiasis who went on to have a cholecystectomy (AHR = 1.56, 95% CI = 1.12–2.17).[1] In comparison, Ward et al. concluded that the incidence of gallstone disease did not increase the risk, hazard ratios (HRs) were stratified by sex, of CRC (HR$_{Men}$ = 0.81, 95% CI = 0.63–1.04; HR$_{Women}$ = 1.14, 95% CI = 0.99–1.31).[2]

These epidemiological studies, however, differ substantially in design, sample size and their ability to adjust for key covariates. The large cohort studies with high numbers of gallstone cases having very limited covariate data and being heavily reliant on self-reported information. For example, in the EPIC cohort study, data on gallstones were missing from 17% of participants and cholecystectomy status was unknown.[2] Moreover, the analysis was based solely on baseline data. Studies with more extensive data on potential confounders have typically been small, being based on fewer gallstone cases.[47]

An important strength of our analysis is that by utilising the random allocation of genetic variants, we were able to overcome potential confounding, for example, from other interrelated traits. Furthermore, reverse causation and selection bias may have biased estimates from previously published observational studies. Exploiting data from large genetic consortia for circulating bilirubin levels, gallstone disease and CRC risk has enabled us to more precisely test study hypotheses than if we had been reliant on individual-level data from a small study. The use of summary test statistics in two-sample MR analyses does, however, require consideration of sample overlap, the winner's curse and genotype uncertainty. We have sought to avoid sample overlap between the association studies of the exposure traits and outcome trait by excluding CRC data based on analysis of UK BioBank.

In our MR, we have used genetic variants as proxies for circulating log levels of bilirubin and gallstone disease. Our findings suggest that neither trait is a major influence on the development of CRC. Our study has several advantages over those conducted previously. Firstly, we made use of summary GWAS data from a large published study of CRC. Secondly, provided the underlying assumptions of the MR estimator we use are met, the common sources of probable bias in previous conventional observational epidemiological studies, including reverse causation and residual confounding, will have been avoided, albeit with limitations and interpretation issues discussed below.

Firstly, as is the case with many MR analyses, despite our study being based on a large case–control data, we were not powered to detect small effects. For gallstone disease, we were unable to assess IV strength and power to detect causal effects as we are not aware of suitable methods for doing so when considering a binary exposure in a two-sample MR investigation. Moreover, we were not able to stratify by subgroups of interest, including sex and cancer site. This may be especially relevant, as some epidemiological studies have only reported significant associations for colon cancer in women. This lack of stratification or adjustment may also have led to bias in the causal estimates[18,48] for bilirubin, as SNP–CRC association estimates were not adjusted for by age and sex. For bilirubin, we were not able to model non-linear associations as this would have required us to have access to individual data.

Analysing a binary exposure and a binary outcome in a two-sample MR setting is problematic with Wald-type ratio estimates potentially being positively biased should there exist a true causal association between a binary exposure and a binary outcome.[30,31] Simulations performed by Didelez et al. demonstrated that if standard IV assumptions were met, one-sample IV estimators do not misspecify a no-association effect as being false. Disney-Hogg et al.[31] arrived at the same conclusion using two-sample IVW estimators. Both simulation studies did however report inflated causal effect magnitudes when causal associations were simulated. Based on such considerations, we believe that our gallstone-CRC causal effect estimates, and the corresponding $P$ values, will only have been subjected to minimal bias. Nevertheless, in analysing the relationship between gallstone disease and CRC, we focused on testing whether there exists a causal association and not principally on the magnitude of any effect.[32,33]

In conclusion, our findings shed light on an area for which the evidence to date has been mixed. Specifically, they provide evidence against the levels of circulating bilirubin or gallstone disease as major risk factors for CRC development.

## AUTHOR CONTRIBUTIONS

R.C. performed statistical analyses. A.J.C. and R.S.H. managed the project. R.C., A.J.C. and P.J.L. acquired and analysed the data. S.M.F., K.P., M.A.J., G.C., M.H., H.B., J.C.C., I.K.,

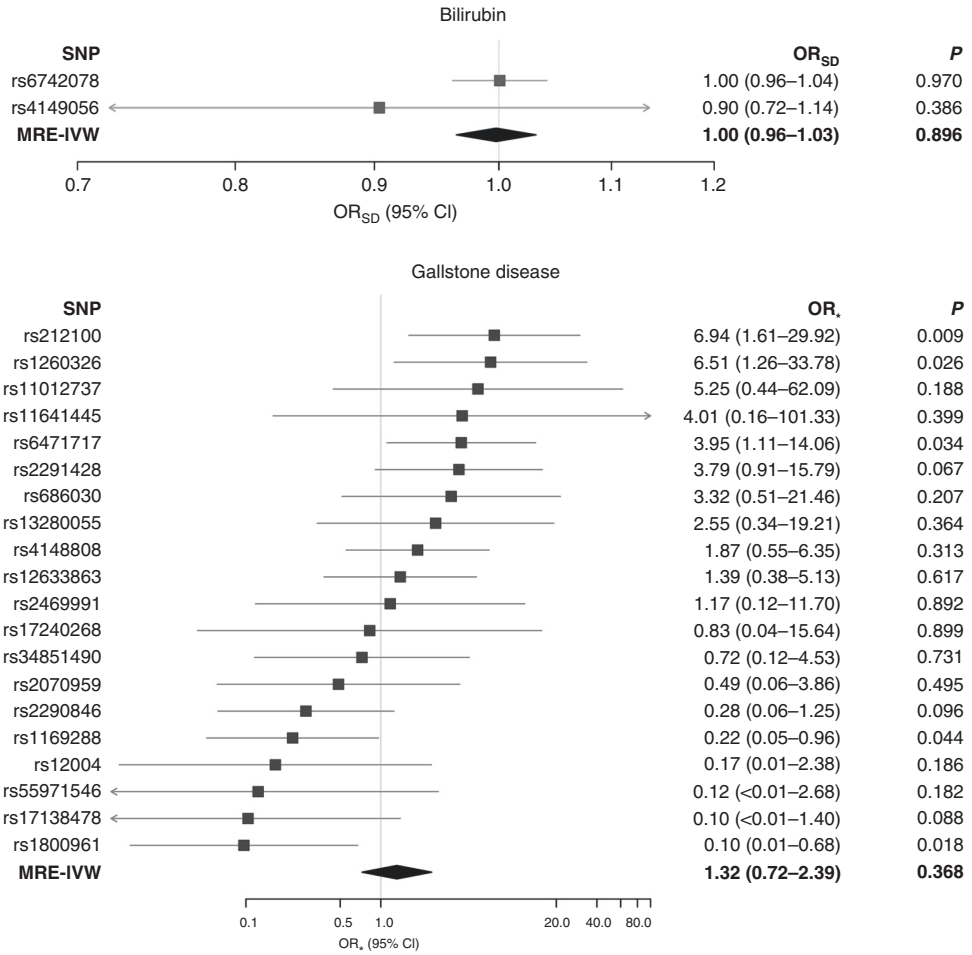


**Fig. 4 Forest plots of instrumental variable Wald ratios and causal effect estimates of the relationship between circulating bilirubin levels and gallstone disease with colorectal cancer.** Causal effects estimated using the multiplicative random effects inverse variance weighted (MRE-IVW) method. $OR_{SD}$ odds ratio per genetically predicted standard deviation unit increase in bilirubin, $OR_*$ odds ratio per genetically predicted standard deviation (of the log odds of gallstone disease) log unit increase of the risk of gallstone disease, caveated by issues discussed within the main text; SNP single-nucleotide polymorphism, CI confidence interval.

T.M., S.B., A.G., J.P.C., L.A.A. and M.G.D. were responsible for colorectal cancer data collection. R.C., A.J.C. and R.S.H. drafted the paper. The final paper has been reviewed and approved by all authors.

## ADDITIONAL INFORMATION

**Ethical approval and consent to participate** Ethical approval was not sought for this study because all data came from the summary statistics of published GWAS, and no individual-level data were used.

**Consent to publish** Not applicable.

**Data availability** Genetic instruments can be obtained through MR-Base or from published work.[23,24] Details and availability of CRC genotyping data that support the findings of this study have been previously published.[26]

**Competing interests** One of the co-authors, Richard S. Houlston, was a senior subject editor at the British Journal of Cancer during the peer-review process.

**Funding information** This work was supported by grants from Cancer Research UK (C1298/A25514, C348/A12076, C6199/A16459, C348/A18927), EU COST Action (BM1206) and the US National Cancer Institute (UM1 CA167551), by a Project Leader grant within the Medical Research Council Human Genetics Unit Centre (Edinburgh, UK, U127527198) and through cooperative agreements with the Australasian Colorectal Cancer Family Registry (U01 CA074778 and U01/U24 CA097735), the Ontario Familial Colorectal Cancer Registry (U01/U24 CA074783) and the Seattle Colorectal Cancer Family Registry (U01/U24 CA074794). Support was also provided by the DJ Fielding Medical Research Trust.

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41416-020-01211-x.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

1. Chen, Y.-K., Yeh, J.-H., Lin, C.-L., Peng, C.-L., Sung, F.-C., Hwang, M. et al. Cancer risk in patients with cholelithiasis and after cholecystectomy: a nationwide cohort study. *J. Gastroenterol.* **49**, 923–931 (2014).
2. Ward, H. A., Murphy, N., Weiderpass, E., Leitzmann, M. F., Aglago, E., Gunter, M. J. et al. Gallstones and incident colorectal cancer in a large pan-European cohort study. *Int J. Cancer* **145**, 1510–1516 (2019).
3. World Cancer Research Fund, American Institute for Cancer Research. Diet, nutrition, physical activity and colorectal cancer. https://www.wcrf.org/sites/default/files/Colorectal-Cancer-2017-Report.pdf (2018).
4. Wang, F., Wang, J., Li, Y., Yuan, J., Yao, P., Wei, S. et al. Gallstone disease and type 2 diabetes risk: a mendelian randomization study. *Hepatology* **70**, 610–620 (2019).
5. Temme, E. H. M., Zhang, J., Schouten, E. G. & Kesteloot, H. Serum bilirubin and 10-year mortality risk in a Belgian population. *Cancer Causes Control* **12**, 887–894 (2001).
6. Zucker, S. D., Horn, P. S. & Sherman, K. E. Serum bilirubin levels in the US population: gender effect and inverse correlation with colorectal cancer. *Hepatology* **40**, 827–835 (2004).

7. Ioannou, G. N., Liou, I. W. & Weiss, N. S. Serum bilirubin and colorectal cancer risk: a population-based cohort study. *Aliment Pharm. Ther.* **23**, 1637–1642 (2006).
8. Stender, S., Frikke-Schmidt, R., Nordestgaard, B. G. & Tybjærg-Hansen, A. Extreme bilirubin levels as a causal risk factors for sympotomatic gallstone disease. *JAMA Intern. Med.* **173**, 1222–1228 (2013).
9. Kühn, T., Sookthai, D., Graf, M. E., Schübel, R., Freisling, H., Johnson, T. et al. Albumin, bilirubin, uric acid and cancer risk: results from a prospective population-based study. *Br. J. Cancer* **117**, 1572–1579 (2017).
10. Buch, S., Schafmayer, C., Völzke, H., Seeger, M., Miquel, J. F., Sookoian, S. C. et al. Loci from a genome-wide analysis of bilirubin levels are associated with gallstone risk and composition. *Gastroenterology* **139**, 1942–1951 (2010).
11. Stocker, R., Yamamoto, Y., McDonagh, A. F., Glazer, A. N. & Ames, B. N. Bilirubin is an antioxidant of possible physiological importance. *Science* **235**, 1043–1046 (1987).
12. Rao, P., Suzuki, R., Mizobuchi, S., Yamaguchi, T. & Sasaguri, S. Bilirubin exhibits a novel anti-cancer effect on human adenocarcinoma. *Biochem Biophys. Res. Commun.* **342**, 1279–1283 (2006).
13. Vítek, L. The role of bilirubin in diabetes, metabolic syndrome, and cardiovascular diseases. *Front. Pharm.* **3**, 55 (2012).
14. Didelez, V. & Sheehan, N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat. Methods Med. Res.* **16**, 309–330 (2007).
15. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* **37**, 658–665 (2013).
16. Haycock, P. C., Burgess, S., Wade, K. H., Bowden, J., Relton, C. & Davey Smith, G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *Am. J. Clin. Nutr.* **103**, 965–978 (2016).
17. Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D. et al. The MR-base platform supports systematic causal inference across the human phenome. *eLife* **7**, e34408 (2018).
18. Davies, N. M., Holmes, M. V. & Davey Smith, G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* **362**, k601 (2018).
19. Burgess, S., Smith, G. D., Davies, N. M., Dudbridge, F., Gill, D., Glymour, M. M. et al. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res.* **4**, 186 (2019). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7384151/, Published online 2020.
20. Brumpton, B., Sanderson, E., Heilbron, K., Hartwig, F. P., Harrison, S., Vie, G. Å. et al. Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nat. Commun.* **11**, 3519 (2020).
21. Gudbjartsson, D. F., Helgason, H., Gudjonsson, S. A., Zink, F., Oddson, A., Gylfason, A. et al. Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47**, 435–444 (2015).
22. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
23. Ferkingstad, E., Oddsson, A., Gretarsdottir, S., Benonisdottir, S., Thorleifsson, G., Deaton, A. M. et al. Genome-wide association meta-analysis yields 20 loci associated with gallstone disease. *Nat. Commun.* **9**, 5101 (2018).
24. Johnson, A. D., Kavousi, M., Smith, A. V., Chen, M.-H., Dehghan, A., Aspelund, T. et al. Genome-wide association meta-analysis for total serum bilirubin levels. *Hum. Mol. Genet.* **18**, 2700–2710 (2009).
25. Borodulin, K., Tolonen, H., Jousilahti, P., Jula, A., Koskinen, S., Kuulasmaa, K. et al. Cohort profile: the national FINRISK study. *Int. J. Epidemiol.* **47**, 696–696i (2018).
26. Law, P. J., Timofeeva, M., Fernandez-Rozadilla, C., Broderick, P., Studd, J., Fernandez-Tajes, J. et al. Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nat. Commun.* **10**, 2154 (2019).
27. Burgess, S. & Thompson, S. G. Assumptions for causal inference. in *Mendelian Randomization: Methods for Using Genetic Variants in Causal Estimation*, 1st edn. (eds Burgess, S., Thompson, S. G.) 25–44 (Chapman & Hall/CRC Press, New York, 2015).
28. Bowden, J., Del Greco, M. F., Minelli, C., Davey Smith, G., Sheehan, N. & Thompson, J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat. Med.* **36**, 1783–1802 (2017).
29. Thompson, S. G. & Sharp, S. J. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat. Med.* **18**, 2693–2708 (1999).
30. Didelez, V., Meng, S. & Sheehan, N. A. Assumptions of IV methods for observational epidemiology. *Stat. Sci.* **25**, 22–40 (2010).
31. Disney-Hogg, L., Cornish, A. J., Sud, A., Law, P. J., Kinnersley, B., Jacobs, D. I. et al. Impact of atopy on risk of glioma: a Mendelian randomisation study. *BMC Med.* **16**, 1–13 (2018).
32. Richmond, R. C., Davey & Smith, G. Comment on 'orienting causal relationships between two phenotypes using bidirectional Mendelian randomization'. *Int. J. Epidemiol.* **48**, 907–911 (2019).
33. Burgess, S. & Labrecque, J. A. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur. J. Epidemiol.* **33**, 947–952 (2018).
34. Cochran, W. G. The combination of estimates from different experiments. *Biometrics* **10**, 101–129 (1954).
35. Del Greco, M. F., Minelli, C., Sheehan, N. A. & Thompson, J. R. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat. Med.* **34**, 2926–2940 (2015).
36. Higgins, J. P., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *BMJ* **327**, 557–560 (2003).
37. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* **40**, 304–314 (2016).
38. Hartwig, F. P., Smith, G. D. & Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.* **46**, 1985–1998 (2017).
39. Bowden, J., Smith, G. D. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).
40. Bowden, J., Fabiola Del Greco, M., Minelli, C., Davey Smith, G., Sheehan, N. A. & Thompson, J. R. Assessing the suitability of summary data for two-sample mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. *Int. J. Epidemiol.* **45**, 1961–1974 (2016).
41. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* **50**, 693–698 (2018).
42. Brion, M.-J. A., Shakhbazov, K. & Visscher, P. M. mRnd: power calculations for Mendelian randomization. https://cnsgenomics.shinyapps.io/mRnd/ (2020).
43. Brion, M.-J. A., Shakhbazov, K. & Visscher, P. M. Calculating statistical power in Mendelian randomization studies. *Int. J. Epidemiol.* **42**, 1497–1501 (2013).
44. Shim, H., Chasman, D. I., Smith, J. D., Mora, S., Ridker, P. M., Nickerson, D. A. et al. A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. *PLoS ONE* **10**, e0120758 (2015).
45. Staiger, D. O. & Stock, J. H. *Instrumental Variables Regression with Weak Instruments* (National Bereau of Economic Research Inc, Cambridge, MA, 1994).
46. R Core Team. *R: A Language and Environment for Statistical Computing*. http://www.R-project.org/ (R Core Team, Vienna, Austria, 2020).
47. Jorgensen, T. & Rafaelsen, S. Gallstones and colorectal cancer–there is a relationship, but it is hardly due to cholecystectomy. *Dis. Colon Rectum* **35**, 24–28 (1992).
48. Lawlor, D. A. Comment on 'two-sample Mendelian randomization: opportunities and challenges'. *Int. J. Epidemiol.* **45**, 908–915 (2016).