

Causal effects of gallstone disease on risk of gastrointestinal cancer in Chinese

Gallstone disease and gastrointestinal cancer

Yuanjie Pang, DPhil¹, Jun Lv, PhD¹, Christiana Kartsonaki, DPhil^{2,3}, Yu Guo, MSc⁴,
Canqing Yu, PhD¹, Yiping Chen, DPhil^{2,3}, Ling Yang, PhD^{2,3}, Zheng Bian, MSc⁴, Iona Y.
Millwood, DPhil^{2,3}, Robin G. Walters, PhD^{2,3}, Xiaojun Li, MBBS⁵, Ju Zou, MBBS⁵,
Michael V. Holmes, PhD^{2,3,6}, Junshi Chen, MD⁷,
Zhengming Chen, DPhil², Liming Li, MD¹

1. Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Centre, 38 Xueyuan Road, Beijing 100191, China
2. Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, Big Data Institute Building, Roosevelt Drive, University of Oxford, UK
3. Medical Research Council Population Health Research Unit (MRC PHRU) at the University of Oxford, Nuffield Department of Population Health, University of Oxford, UK
4. Chinese Academy of Medical Sciences, 9 Dongdan San Tiao, Beijing 100730, China
5. Jili Street Community Health Service Center, 434 Jinshabei Road, Liuyang 410300, China
6. National Institute for Health Research Oxford Biomedical Research Centre, Oxford University Hospital, Old Road, Oxford OX3 7LE, UK
7. National Center for Food Safety Risk Assessment, 37 Guangqu Road, Beijing 100021, China

Address for correspondence:

Prof. Liming Li
Department of Epidemiology and
Biostatistics
School of Public Health
Peking University Health Science Center
38 Xueyuan Road
Beijing, 100191, China
Tel: 86-010-82801530
Fax: 86-010-82801530
75769458@qq.com

Word count: abstract 247, main text 4154

(4 tables, 1 figure and 15 supplementary tables/figures)

Revised on February 2021

Abstract

Background: Gallstone disease (GSD) is associated with higher risk of gastrointestinal (GI) cancer. However, it is unclear whether the associations are causal.

Methods: The prospective China Kadoorie Biobank (CKB) recorded 17,598 cases of GI cancer among 510,137 participants without cancer at baseline during 10 years of follow-up. Cox regression was used to estimate hazard ratios (HRs) for specific cancer by GSD status and duration. Mendelian randomisation was conducted to assess the genetic associations of GSD with specific cancer.

Results: Overall 6% of participants had symptomatic GSD at baseline. Compared with those without GSD, individuals with symptomatic GSD had adjusted HRs of 1.13 (1.01-1.29) for colorectal, 2.01 (1.78-2.26) for liver, 3.70 (2.88-4.87) for gallbladder, 2.31 (1.78-3.07) for biliary tract, and 1.38 (1.18-1.74) for pancreatic cancer. Compared with participants without GSD, the risks of colorectal, liver, gallbladder, biliary tract, and pancreatic cancer were highest during 0 to <5 years following disease diagnosis. There was evidence of genetic associations of GSD with these cancers, with odds ratios per 1-SD genetic score of 1.08 (1.05-1.11) for colorectal, 1.22 (1.19-1.25) for liver, 1.56 (1.49-1.64) for gallbladder, 1.39 (1.31-1.46) for biliary tract, and 1.16 (1.10-1.22) for pancreatic cancer. When meta-analysing the genetic estimates in CKB and UK Biobank, there was evidence of causal associations of GSD on colon cancer, gallbladder and biliary tract cancer (GBTC), and total GI cancer (RR per 1-SD: 1.05 [0.99-1.11], 2.00 [1.91-2.09], and 1.09 [1.05-1.13]).

Conclusions: GSD was associated with higher risks of several GI cancers, warranting future studies on the underlying mechanisms.

Key words: gallstone disease; gastrointestinal cancer; Mendelian randomisation; Chinese

1 **Introduction**

2 Gallstone disease (GSD) is a major type of gallbladder disease, affecting 10-20% of the global
3 adult population.^{1,2} East Asian countries including China have been traditionally considered as
4 low-risk areas of GSD (i.e. prevalence in women <10%).¹ A national survey in China in 2008
5 showed that 7.9% men and 8.2% women had GSD detected by ultrasound,³ but there is limited
6 data available in recent decades. Gastrointestinal (GI) cancer is a major cause of mortality and
7 morbidity and accounts for 3.6 million deaths globally in 2017.⁴ Compared with high-income
8 countries, the age-standardised incidence rates in China were higher for oesophageal,
9 stomach, and liver cancer, and were lower for colorectal and pancreatic cancer.⁴ In China, GI
10 cancer cases attributable to infectious agents (i.e. hepatitis B virus [HBV] and *H. pylori*) and
11 environmental factors (i.e. aflatoxin) have been declining, while GI cancer cases associated with
12 metabolic risk factors (i.e. physical inactivity, adiposity, and diabetes) have been increasing.⁵

13 GSD is associated with higher risks of several GI cancers in Western populations, including
14 colorectal, liver, pancreatic, and gallbladder and biliary tract cancer (GBTC). However, there is
15 limited evidence in the Chinese population where patterns of metabolic risk factors and GI
16 cancer differ from those in the West.^{4,5} Several mechanisms underlying the associations of GSD
17 with GI cancers have been proposed including shared risk factors (e.g. adiposity, diabetes),
18 surveillance bias, and possible aetiological factors (e.g. bile acids, inflammation, altered
19 microbiota).^{1,2} Mendelian randomisation can be used to disentangle causality from confounding
20 and bias, which utilises the random assortment of genes from parents to offspring at conception
21 and uses gene variants associated with the exposure of interest as unconfounded markers⁶.

22 Whether GSD is causally related to GI cancers can be examined using Mendelian
23 randomisation. Assessing the genetic associations of GSD with GI cancer may provide valuable
24 insights into disease aetiology and prevention. A prospective cohort study in China constructed
25 a genetic risk score (GRS) for GSD,⁷ providing the opportunity to assess the genetic
26 associations of GSD with GI cancer.

27 Therefore, the objectives of this study were: (1) to examine the observational associations of
28 GSD with GI cancer and to meta-analyse the observational associations in CKB with
29 prospective studies published mostly in Western countries; (2) to assess the genetic
30 associations of GSD with GI cancer using a GRS (8 SNPs: rs6742078, rs9843304, rs11887534,
31 rs6471717, rs139526437, rs6782601, rs201644871, and rs2547238).

32 **Methods**

33 *Study population*

34 512,715 participants (210,205 men and 302,510 women) aged 30-79 years were recruited into
35 the study from 10 geographically defined localities (5 urban and 5 rural) in China during 2004-
36 2008.⁸ The study areas were selected to provide diversity in risk exposure and disease
37 patterns, while taking into account population stability, quality of mortality and morbidity
38 registries, capacity, and long-term commitment within the areas. The CKB study was approved
39 by the Ethical Review Committee of the Chinese Centre for Disease Control and Prevention
40 (CDC) and the Oxford Tropical Research Ethics Committee, University of Oxford. All
41 participants eligible for this study had completed a written informed consent form. All methods
42 were performed in accordance with relevant guidelines and regulations.

43 At local study assessment clinics, participants completed an interviewer-administered laptop-
44 based questionnaire on sociodemographic characteristics, lifestyle factors (smoking, alcohol
45 consumption, diet, physical activity), personal and family medical history, and current
46 medication. A range of physical measurements were recorded by trained technicians, including
47 height, weight, hip and waist circumference, bio-impedance, lung function, blood pressure, and
48 heart rate, using calibrated instruments with standard protocols. All participants provided a
49 10mL nonfasting (with time since last meal recorded) blood sample for immediate on-site test of
50 random plasma glucose (SureStep Plus meters; Lifescan, Johnson & Johnson) and HBsAg
51 (ACON Biotech).

52 *Assessment of GSD*

53 Previously diagnosed GSD was defined by the question “Has a doctor ever told you that you
54 had gallstone disease?”. Among positive respondents, additional information about age at
55 diagnosis and the presence of cholecystitis complication was collected.

56 *Follow-up for morbidity and mortality*

57 The vital status of each participant was determined periodically through China Centre for
58 Disease Control and Prevention (CDC)’s Disease Surveillance Points (DSP) system and
59 national health insurance system, supplemented by regular checks against local residential and
60 administrative records and by annual active confirmation through street committees or village
61 administrators.⁹ Additional information about major diseases and any episodes of hospitalisation
62 was collected through linkages, via each participant’s unique national identification number, with
63 disease registries (for cancer, ischaemic heart disease, stroke, and diabetes) and national
64 health insurance claims databases, which has almost universal coverage in the study areas. All
65 events were coded using International Classification of Diseases, 10th Revision (ICD-10) by
66 trained staff who were blinded to baseline information, and reviewed centrally for consistency.
67 Information on cancer histological subtypes was also collected for a subset of the cases through
68 cancer registries or reviews of hospital medical notes as part of the ongoing outcome
69 adjudication for major diseases. The classification and distribution of GI cancer are shown in
70 **Supplementary Table 1**. By 1.1.2017 (censoring date for the present analyses), 44,066 (8.6%)
71 participants had died and 4,751 (1.0%) were lost to follow-up. Participants were followed up to
72 the time of diagnosis of the first GI cancer. Participants who died or were lost to follow-up were
73 censored at the last day known to be alive.

74 *Genotyping*

75 Genotyping was conducted in 100,408 individuals using a custom-designed 800K-SNP array
76 (Axiom; Affymetrix) and imputed to 1000 Genomes Phase III (call rates were >99.97% for all

77 variants). Genotyping consisted of a population-based sample of 75,736 participants randomly
78 selected from the total CKB cohort and included in all genetic analyses. The remaining 24,672
79 participants with genotyping data were selected for nested case-control studies of incident
80 cardiovascular disease or chronic obstructive pulmonary disease, and were excluded from the
81 analyses to avoid potential selection bias. The baseline characteristics were generally
82 comparable between the population subset and the whole cohort (**Supplementary Table 2**).

83 *Statistical analyses*

84 For the observational analyses in CKB, we excluded individuals with a prior history of cancer at
85 baseline (n=2578), leaving 510,137 individuals for the main analysis. Mean values and
86 prevalence of baseline characteristics were calculated for GSD status at baseline, standardised
87 to the age (in 5-year groups), sex, and area structure of the CKB population. Cox proportional
88 hazards models were used to estimate adjusted hazard ratios (HRs) of specific disease
89 incidence associated with baseline GSD, stratified by sex and study area (10 areas), and
90 adjusted for age at baseline, education (4 groups: no formal school, primary school, middle/high
91 school, or college/university), smoking (4 groups: never regular, occasional, former regular, or
92 current regular), alcohol (5 groups: abstainers, ex-weekly drinkers, reduced-intake drinkers,
93 occasional drinkers, or weekly drinkers), total physical activity, BMI, and prevalent diabetes.
94 Cox models with a time-updated exposure for GSD duration were used to estimate the
95 association of incident GSD and GSD duration with risk of GI cancer, with the same adjustment.
96 Age was used as the underlying time scale. Among participants without GSD at baseline,
97 incident GSD cases (ICD-10 code K80, n=8515) that occurred during the follow-up were also
98 included, using Cox models with a time-updated exposure counting incident GSD as exposed
99 from their time of diagnosis. Duration was defined as the time interval between GSD diagnosis
100 and time at risk.

101 For the meta-analysis, the CKB estimates for GSD and GI cancer were meta-analysed with
102 estimates from published prospective cohort studies using a fixed effects meta-analysis.

PubMed was searched for studies published in English from January 2000 to October 2019. The precise search terms are provided in **Supplementary Figure 1**. Inclusion criteria were prospective cohort studies, case-cohort studies, or nested case-control studies reporting the association of GSD and cholecystectomy with incidence or mortality of GI cancer. Bibliographies of included studies and related reviews were manually searched for additional eligible articles. Details of the study selection are reported in **Supplementary Figure 1**. YP and JL conducted the search and extracted the study information. Meta-analyses were done separately for symptomatic GSD, screen-detected GSD, and cholecystectomy as previous studies suggested that different mechanisms were involved.^{10,11} Estimates combining both symptomatic or screen-detected GSD and cholecystectomy were not included. The main subgroup analyses included sex, diagnostic method (for symptomatic GSD: self-reported or ICD-diagnosis), and duration of GSD. Heterogeneity between studies was assessed using I^2 and Cochran's Q test.

For the genetic analysis in CKB, we selected eight single nucleotide polymorphisms as candidate instrumental variables for baseline GSD based on a recent Mendelian randomisation study in Chinese (**Supplementary Table 3**)⁷. These eight SNPs were selected from previous GWAS findings in Europeans while taking into account the minor allele frequencies and effect sizes in Chinese. Therefore, four originally-reported SNPs (rs6742078, rs9843304, rs11887534, and rs6471717) and four proxy SNPs (rs139526437, rs6782601, rs201644871, and rs2547238) were selected.⁷ For each variant, the effect allele was defined as the allele associated with higher risk of GSD in the Chinese study. An externally weighted GSD-GRS was constructed by summing the number of effect alleles carried by each participant, weighted by the reported effect size of each variant on GSD reported by that Chinese study. The associations between GSD-GRS and GI cancer were examined using logistic regression adjusting for age, sex, region, the first 10 principal components, education, smoking, and alcohol using individual participant-level data (IPD). The effect of genetically-determined GSD on specific GI cancer was

129 estimated using the ratio method by scaling the natural logOR of specific GI cancer per natural
130 logOR of GSD. To enable comparison with observational estimates, genetic estimates of the
131 odds of specific cancer associated with GSD were estimated using the formula $OR = 1 + (1 -$
132 $\exp(\beta)) / (\exp(\beta) - \exp(1)) \times A$ where β is the regression coefficient of the GRS association with
133 specific cancer as a function of the GRS association with GSD, and A is the prevalence of GSD
134 in CKB.¹² Heterogeneity between observational and genetic risk estimates were assessed using
135 Cochran's Q test. Associations of each SNP and GRS with potential confounders were
136 examined by linear regression for continuous variables and logistic regression for categorical
137 variables. Genetic associations in CKB were meta-analysed with those calculated from the UK
138 Biobank (UKB) using two-sample Mendelian randomisation and summary-level statistics. 32
139 SNPs for GSD were reported by a recent GWAS paper in Europeans and SNPs for GI cancer
140 were extracted from a UKB GWAS by Canela-Xandri et al.^{13,14}

141 We conducted several sensitivity analyses. First, we excluded the first five years of follow-up in
142 the observational analyses to address reverse causality. Second, weighted median and MR-
143 Egger tests were conducted for the causal estimates using summary data from multiple genetic
144 variants. We also compared the genetic associations of GSD with GI cancer using IPD analysis
145 and two-sample Mendelian randomisation (i.e. IVW estimates).

146 **Results**

147 *Baseline characteristics and incidence of GI cancer*

148 Of the 510,137 participants included, mean age was 52 years and 59.2% were women. 30,783
149 participants had prevalent GSD at baseline. The mean BMI was 23.8 (3.4) kg/m². During 10
150 years of follow-up, there were 8515 incident cases of symptomatic GSD among those without
151 prevalent disease at baseline. Participants with GSD were more likely to be older and female,
152 and more likely to have higher levels of household income (**Table 1**). Participants with GSD
153 were more likely to have higher levels of adiposity, and they were more likely to have prevalent
154 diabetes, CHD, and hypertension. They were also more likely to have family history of diabetes

155 and cancer. No differences in female reproductive factors were observed across disease status
156 (**Table 1**). Among those with previously diagnosed GSD, the median age at first diagnosis was
157 45 years and the median duration since diagnosis was 6.2 years. The median (interquartile
158 range) of follow-up was 10.1 (9.2, 11.1) years. During 10 years of follow-up, there were 17,598
159 cases of GI cancer (**Supplementary Table 1**).

160 *Observational associations of GSD with GI cancer*

161 Symptomatic GSD was associated with increased risks of colon, rectal, liver, gallbladder, biliary
162 tract, and pancreatic cancer, and not associated with risks of oesophageal, stomach, and small
163 intestine cancer. The corresponding adjusted HRs were 0.97 (95% CI 0.78-1.21) for
164 oesophageal, 0.99 (0.88-1.14) for stomach, 1.01 (0.58-1.90) for small intestine, 1.15 (1.00-1.37)
165 for colon, 1.23 (1.08-1.46) for rectal, 2.01 (1.78-2.26) for liver, 3.70 (2.88-4.87) for gallbladder,
166 2.31 (1.78-3.07) for biliary tract, and 1.38 (1.18-1.74) for pancreatic cancer (**Table 2**). The HRs
167 were similar in men and women, with no heterogeneity by sex (p -values for heterogeneity 0.08-
168 0.92, **Supplementary Table 4**). Compared with participants without GSD, the risks of
169 colorectal, liver, gallbladder, biliary tract, and pancreatic cancer were highest during 0 to <5
170 years following disease diagnosis, with the HRs of 1.35 (0.87-2.09), 3.85 (3.08-4.80), 9.60
171 (6.19-14.90), 7.34 (4.78-11.27), and 3.25 (2.14-4.93), respectively (**Table 3**). The excess risks
172 decreased with increasing duration of GSD, but remain elevated even ≥ 10 years after GSD
173 diagnosis (HR 1.18 [0.98-1.42], 2.02 [1.77-2.31], 3.53 [2.65-4.69], 2.50 [1.85-3.37], and 1.35
174 [1.04-1.74]). There were no associations between duration of GSD with risks of oesophageal
175 and stomach cancer. The number of small intestinal cancer cases was too small to permit such
176 an analysis.

177 *Meta-analysis of observational associations of gallstone disease with GI cancer*

178 The meta-analysis involved 16 prospective studies, with 56,227 cases of GI cancer
179 (**Supplementary Table 5**).^{10,11, 15-28} In the meta-analysis, symptomatic GSD was associated
180 with increased risks of stomach, small intestine, colon, liver, gallbladder, biliary tract, and

pancreatic cancer (**Figure 1**). The corresponding pooled RRs were 1.15 (1.08-1.22) for stomach, 1.24 (1.07-1.43) for small intestine, 1.12 (1.02-1.23) for colon, 1.94 (1.84-2.05) for liver, 3.41 (3.09-3.75) for gallbladder, 2.15 (1.68-2.74) for biliary tract, and 1.27 (1.20-1.34) for pancreatic cancer. There were no associations for oesophageal, proximal colon, and colorectal cancer (RR: 1.05 [0.96-1.15], 1.00 [0.95-1.05], and 1.00 [0.96-1.03]), while there were inverse associations for distal colon and rectal cancer (0.94 [0.89-1.00] and 0.88 [0.84-0.94]). Compared with symptomatic GSD, the associations were similar for cholecystectomy, except for liver cancer where a weaker association was observed for cholecystectomy (p -value for heterogeneity <0.001 , **Supplementary Table 6**). For symptomatic GSD, there were similar associations by diagnostic methods except for colorectal and rectal cancer (**Supplementary Table 7**). For colorectal cancer, there was no association for GSD ascertained by ICD-diagnosis (HR 0.97 [0.93-1.01]), whereas there was a positive association for self-reported GSD (1.09 [1.02-1.17]). For rectal cancer, there was an inverse association for GSD ascertained by ICD-diagnosis (0.83 [0.78-0.89]), whereas there was a positive trend for self-reported GSD (1.07 [0.96-1.20]). Only one study ascertained screen-detected GSD and showed a positive association for distal colon cancer (HR 2.22 [1.11-4.46]) and no associations for proximal colon and rectal cancer (2.73 [0.95-7.87] and 0.74 [0.27-2.06]). Of individual studies reporting sex-specific associations, there were no differences in the GSD-cancer associations by sex (**Supplementary Table 8**). Meta-analysis for duration was not conducted as the associations of duration with GI cancer were not assessed uniformly by previous studies.

Genetic associations of GSD with GI cancer

The mean weighted GRS was 1.24 (SD 0.20) in the present study. Each 1-SD higher GRS was associated with a 1.04-fold increased risk of symptomatic GSD (OR 1.04 [1.01-1.07]). Weighted GRS was associated with increased risks of colorectal, liver, gallbladder, biliary tract, and pancreatic cancer (OR: 1.08 [1.05-1.11], 1.22 [1.19-1.25], 1.56 [1.49-1.64], 1.39 [1.31-1.46], and 1.16 [1.10-1.22], **Table 2**). There was no evidence of genetic associations of GSD with risks of

oesophageal, stomach, and small intestine cancer (**Table 2**). There was no evidence of heterogeneity between observational and genetic associations, except for liver, gallbladder, and biliary tract cancer (p -value for heterogeneity <0.001). For these three cancers, the genetic estimates were weaker than the observational estimates (**Table 2**). The HRs of GI cancer associated with 1-SD higher weighted GRS were reported in **Supplementary Table 9**. The associations of individual SNPs with gallstone disease and individual cancer were reported in **Supplementary Figure 2**. When meta-analysing the genetic estimates in CKB and UKB, there was evidence of causal associations of GSD on colon cancer, gallbladder and biliary tract cancer (GBTC), and total GI cancer (RR per 1-SD: 1.05 [0.99-1.11], 2.00 [1.91-2.09], and 1.09 [1.05-1.13], **Supplementary Table 10**). However, large heterogeneity between CKB and UKB was observed for rectal cancer and GBTC.

Observational and genetic associations of GSD with non-GI cancer

For non-GI cancer, symptomatic GSD was associated with a lower risk of lung cancer (HR 0.98 [0.96-0.99]) and higher risks of breast and cervix cancer (1.63 [1.17-2.27] and 1.53 [1.15-2.04]); GSD was not associated with risks of endometrial and ovarian cancer (1.10 [0.95-1.28] and 1.15 [0.91-1.45], **Table 4**). For lung cancer, the observational association became non-significant when restricting to never regular smokers. For female cancer, the positive associations attenuated towards the null but were significant for breast and cervix cancer when additionally adjusting for total physical activity, BMI, and diabetes. In contrast to the observational associations, there was no evidence of genetic associations of GSD with any of these non-GI cancers (**Table 4**).

Sensitivity analyses

For the observational analyses, the associations attenuated slightly towards the null and remained almost unchanged when excluding the first five years of follow-up (**Supplementary Table 11**). For the genetic analyses, two-sample MR (i.e. IVW estimator) yielded similar associations compared with IPD analysis (**Supplementary Table 9**). Similarly, weighted median

test showed a similar pattern of associations, whereas the MR-Egger estimates were more extreme. For the MR-Egger estimates, p -value of intercept was non-significant except for biliary tract cancer (p -value=0.03).

Discussion

In this Chinese population, symptomatic GSD was associated with higher risks of several GI cancers. The observational associations in our study were consistent with those reported by previous prospective studies primarily conducted in Western populations, as shown in the meta-analysis. We showed that there is evidence of causal effects of symptomatic GSD on these GI cancers. Future studies are warranted to understand the mechanisms linking GSD to GI cancer. Our findings suggest that GSD might be a causal risk factor for specific GI cancer and provide insights into the aetiology of GI cancer.

The observational associations of symptomatic GSD with GI cancer in CKB are largely consistent with previous prospective studies, as shown in the meta-analysis. For six GI cancer (oesophageal, colon, liver, gallbladder, biliary tract, and pancreatic cancer), the effect estimates in CKB agreed with those reported in previous studies. However, it should be noted that large between-study heterogeneity was observed for liver, gallbladder, and pancreatic cancer (**Figure 1 and Supplementary Table 6**). In contrast to the positive association reported in previous studies, there was no evidence of association for stomach cancer in CKB, which might be due to the different subtypes. The commonest subtype is non-cardiac stomach cancer in Western populations, while the commonest subtype is cardiac stomach cancer in Chinese³³. However, there is limited evidence on the mechanisms between GSD and different subtypes of stomach cancer. For rectal and colorectal cancer, the CKB estimates were different from those reported by previous studies, where large between-study heterogeneity was observed (**Figure 1**). However, when stratifying on diagnostic methods, there were positive associations of self-reported GSD with rectal and colorectal cancer (**Supplementary Table 7**), consistent with the CKB estimates. Studies using ICD-diagnosed GSD were record-linkage studies and mostly

adjusted for only age and sex. These studies also had large heterogeneity, and therefore the associations might be biased.^{14, 15}

The observational associations of symptomatic GSD with GI cancer might be explained by confounding and surveillance bias. Symptomatic GSD and GI cancer share many risk factors, including physical inactivity, adiposity, and diabetes.^{1,2,20-22} However, we carefully adjusted for a wide range of potential confounders, minimising residual confounding. For breast cancer, the weak association for symptomatic GSD was most likely due to common factors and through the effects of endogenous hormones on lithogenesis.²³ On the other hand, surveillance bias may occur as patients with newly diagnosed GSD are under increased medical surveillance and thus might be more likely to be diagnosed with cancer or vice versa. Indeed, there were 0.5- to 11-fold higher risks of several GI cancers in the first five years following GSD diagnosis, consistent with previous studies.^{18-20,31,32} Nonetheless, we showed that participants with ≥ 10 years of GSD had elevated risks of GI cancer.

In line with the observational associations, Mendelian randomisation suggested evidence of causal effects of symptomatic GSD on colorectal, liver, gallbladder, biliary tract, and pancreatic cancer. The genetic associations of GSD with gallbladder and biliary tract cancers are consistent with two previous studies using variants in ABCG8 and TRAF3 genes as predictors^{34,35}. Several mechanisms have been proposed. First, bile acids have been implicated in the pathogenesis of certain GI cancer, including colorectal, liver, gallbladder, biliary tract, and pancreatic cancer.^{36,37} Case-control studies have shown higher levels of bile acids (especially secondary bile acids) in cases of colorectal, liver, and gallbladder cancer than controls.³⁸⁻⁴³ For pancreatic cancer, increased secretion of bile acids may cause bile acid reflux into the pancreatic duct and to the epithelial or acinar cells, where pancreatic adenocarcinoma (the most common type of pancreatic cancer) originates.^{36,43} Second, GSD can cause repeated trauma to the gallbladder mucosa, leading to chronic inflammation in the gallbladder and adjacent organs.³⁹⁻⁴¹ Prospective studies have suggested a positive association of inflammatory

cytokines with colorectal, gallbladder, and pancreatic cancer.^{44,45} However, inflammation may
 be a marker rather than an aetiological factor of cancer. In the current study, we did not observe
 genetic associations between GSD and inflammation markers including fibrinogen and C-
 reactive protein. Third, GSD can lead to pancreatitis via blockage of the pancreatic duct by
 gallstones, which is a known risk factor for pancreatic cancer.⁴⁷ Indeed, our study showed that
 the weighted GRS was associated with a higher risk of acute pancreatitis in CKB, a strong risk
 factor for GI cancer.⁴⁷ Lastly, experimental and clinical studies have suggested a key role of
 altered microbiota (particularly *Fucibacterium nucleatum*) underlying the association between
 GSD and colon cancer.¹¹

In the present Mendelian randomisation analysis, the plausibility of three assumptions was
 assessed to ensure the validity of our analyses. First, our data showed that the individual SNPs
 and the weighted GRS were associated with GSD (F statistic 532), suggesting that the GRS is a
 strong instrumental variable. Indeed, previous GWAS have reported plausible biological
 mechanisms for several SNPs⁴⁸⁻⁵⁰. For example, rs11887534 in ABCG8 gene may facilitate
 efflux of cholesterol from enterocytes and hepatocytes in to the intestine and bile, promoting the
 formation of cholesterol gallstones⁴⁹. Besides, rs6471717 resides in the intergenic region of
 CYP7A1 and UBXN2B genes, both of which play a major role in biliary cholesterol secretion⁵⁰.
 Second, the instrument should be independent of measured and unmeasured confounders. As
 shown in **Supplementary Table 12**, the weighted GRS were not associated with these
 confounders (all *p*-values >0.05). Two SNPs (rs6471717 and rs2547238) were positively
 associated with random plasma glucose. However, the findings were similar when constructing
 a new GRS excluding these SNPs (**Supplementary Table 13**). Third, the instrument should not
 affect the outcome by pathways other than the exposure of interest. *GCKR* was a locus for type
 2 diabetes and may have pleiotropic effect on the outcomes. However, the findings were similar
 when excluding the SNP in *GCKR* in the weighted GRS (**Supplementary Table 13**).

The strengths of CKB include a prospective design, a large and diverse study population, a

311 large number of cancer cases, and extensive adjustment for risk factors for GSD and GI cancer.

312 In particular, we ascertained GI cancer through linkage to hospital records in addition to death

313 and cancer registries, which allowed us to examine different subtypes of GI cancer. This study

314 also has limitations. First, baseline GSD was primarily ascertained by self-report, which might

315 lead to under-diagnosis. For GSD, the prevalence of baseline GSD was 6% in CKB, which was

316 comparable with the prevalence of symptomatic GSD in regional surveys in China.³ In addition,

317 misclassification of GSD is likely to be non-differential, and therefore would lead to

318 underestimation of the observational estimates. Nonetheless, we showed that our estimates for

319 symptomatic GSD were comparable with pooled RRs in the meta-analysis. Second, large

320 between-study heterogeneity was observed in the meta-analyses, particularly for stomach,

321 colorectal, liver, and gallbladder cancer. Although anatomical subtypes may explain the

322 between-study heterogeneity for colorectal cancer (**Supplementary Table 6**), such information

323 was not provided by all included studies. As shown in ours and previous meta-analyses⁵¹⁻⁵³,

324 sex, regions, and diagnostic methods did not explain the large heterogeneity between studies,

325 but it is not known whether subtypes might explain the between-study heterogeneity. Third,

326 GWAS data was only available for ~20% of all participants, and therefore there is limited power

327 to conduct analysis by cancer subtypes (i.e. proximal and distal colon cancer).

328 In conclusion, symptomatic GSD was associated with higher risks of a diverse range of GI

329 cancer, and the evidence was most robust for colorectal, liver, gallbladder, biliary tract, and

330 pancreatic cancer. Our study showed causal effects of GSD on these GI cancers. Our findings

331 suggest that GSD might be a causal risk factor for specific GI cancer. As GSD and GI cancer

332 share metabolic risk factors (i.e. physical inactivity, adiposity, diabetes), lifestyle intervention

333 may be beneficial in lowering risk of both diseases. Future studies are warranted to investigate

334 the mechanisms underlying the causal effects of GSD on GI cancer.

Acknowledgment

The chief acknowledgment is to the participants, the project staff, and the China National Centre for Disease Control and Prevention (CDC) and its regional offices for access to death and disease registries. The Chinese National Health Insurance scheme provides electronic linkage to all hospital admission data.

Author contributions

LL and ZC had full access to the data. YP, JL, CK, ZC, and LL conducted data analysis and are responsible for accuracy of the results and the decision to submit for publication. YG, CY, YC, ZB, IYM, RGW, XL, JZ, MVH, and JC were involved in study design, conduct, long-term follow-up, review and coding of disease events, and/or interpretation of the results. All authors were involved in drafting and revising the manuscript and approved the final version of the manuscript.

Ethical approval and consent to participate

The CKB study was approved by the Ethical Review Committee of the Chinese Center for Disease Control and Prevention and the Oxford Tropical Research Ethics Committee, University of Oxford. The study was performed in accordance with the Declaration of Helsinki. All participants eligible for this study had completed a written informed consent form.

Data availability

CKB investigators are committed to sharing this important resource with the wider scientific community, so that the potential value of the CKB resource can be maximised. Open access to the CKB resource has begun in a phased approach. To facilitate the process a Data Access Committee (see <http://www.ckbiobank.org/site/Data+Access>) has been established, comprising not only senior CKB scientists but also external experts in related fields. For any external data

access requests, an outline proposal defining the purpose of the investigation, the data/samples required and the time-scale for the analysis needs to be completed and submitted for review by the study executive committee. The access request review will assess the scientific merit of the proposal to ensure that research questions are legitimate and that there is no duplication of effort. Only proposals complying with the activities listed in the participant's original consent and with the study's ethical approval will be considered.

Competing interests

None declared.

Funding information

Dr Pang acknowledges support from the China Postdoctoral Science Foundation (2019TQ0008 and 2020M670071). This work was supported by grants (2016YFC0900500, 2016YFC0900501, 2016YFC0900504, 2016YFC1303904) from the National Key R&D Program of China. The CKB baseline survey and the first re-survey were supported by a grant from the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up is supported by grants from National Natural Science Foundation of China (91846303, 91843302, 81390540, 81390541, 81390544), and Chinese Ministry of Science and Technology (2011BAI09B01). Dr Holmes is supported by a British Heart Foundation Intermediate Clinical Research Fellowship (FS/18/23/33512) and the National Institute for Health Research Oxford Biomedical Research Centre.

References

1. Lammert, F., Gurusamy, K., Ko, C.W., Miquel, J.F., Méndez-Sánchez, N., Portincasa, P., et al. Gallstones. *Nat Rev Dis Primers* 2:16024 (2016).
2. Stinton, L.M., Shaffer, E.A. Epidemiology of gallbladder disease: cholelithiasis and cancer. *Gut Liver* 6(2):172-187 (2012).
3. Zeng, Q., He, Y., Qiang, D.C., Wu, L.X. Prevalence and epidemiological pattern of gallstones in urban residents in China. *Eur J Gastroenterol Hepatol* 24(12):1459-1460 (2012).
4. Global Burden of Disease Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 392(10159):1736-1788 (2018).
5. Zhou, M., Wang, H., Zeng, X., Yin, P., Zhu, J., Chen, W., et al. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 394(10204):1145-1158 (2019).
6. Holmes, M.V., Ala-Korpela, M., Davey Smith, G. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat Rev Cardiol* 14(10):577-590 (2017).
7. 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(2):610-620 (2019).
8. Chen, Z., Chen, J., Collins, R., Guo, Y., Peto, R., Wu, F., et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* 40(6):1652-1666 (2011).
9. Yang, G., Rao, C., Ma, J., Wang, L., Wan, X., Dubrovsky, G., et al. Validation of verbal autopsy procedures for adult deaths in China. *Int J Epidemiol* 35(3):741-748 (2006).
10. Shabanzadeh, D.M., Sorensen, L.T., Jorgensen, T. Association between screen-detected gallstone disease and cancer in a cohort study. *Gastroenterology* 152(8):1965-1974 (2017).
11. Ryu, S., Chang, Y., Yun, K.E., Jung, H.S., Shin, J.H., Shin, H. Gallstones and the risk of gallbladder cancer mortality: a cohort study. *Am J Gastroenterol* 111(10):1476-1487 (2016).
12. Gan, W., Bragg, F., Walters, R.G., Millwood, I.Y., Lin, K., Chen, Y., et al. Genetic predisposition to type 2 diabetes and risk of subclinical atherosclerosis and cardiovascular diseases among 160,000 Chinese adults. *Diabetes* 68(11):2155-2164 (2019).
13. 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(1): 5101 (2018).
14. Canela-Xandri, O., Rawlik, K., Tenesa, A. An atlas of genetic associations in UK Biobank. *Nat Genet* 50(11):1593-1599 (2018).
15. Ward, H.A., Murphy, N., Weiderpass, E., Leizmann, 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(6):1510-1516 (2019).
16. Lee, P.C., Hu, Y.W., Hu, L.Y., Chen, S.C., Chien, S.H., Shen, C.C., et al. Risk of cancer in patients with cholecystitis: a nationwide population-based study. *Am J Med* 128(2):185-191 (2015).
17. Vogtmann, E., Shu, X.O., Li, H.L., Chow, W.H., Yang, G., Ji, B.T., et al. Cholelithiasis and the risk of liver cancer: results from cohort studies of 134,546 Chinese men and women. *J Epidemiol Community Health* 68(6):565-570 (2014).
18. Nogueira, L., Freedman, N.D., Engels, E.A., Warren, J.L., Castro, F., Koshiol, J. Gallstones, cholecystectomy, and risk of digestive system cancers. *Am J Epidemiol* 179(6):731-739 (2014).

19. Chen, Y.K., Yeh, J.H., Lin, C.L., Peng, C.L., Sung, F.C., Hwang, I.M., et al. Cancer risk in patients with cholelithiasis and after cholecystectomy: a nationwide cohort study. *J Gastroenterol* 49(5):923-931 (2014).
20. Tavani, A., Rosato, V., Di Palma, F., Bosetti, C., Talamini, R., Maso, L.D., et al. History of cholelithiasis and cancer risk in a network of case-control studies. *Ann Oncol* 23(8):2173-2178 (2012).
21. Hartz, A., He, T., Ross, J.J. Risk factors for colon cancer in 150,912 postmenopausal women. *Cancer Causes Control* 23(10):1599-1605 (2012).
22. Arnold, L.D., Patel, A.V., Yan, Y., Jacobs, E.J., Thun, M.J., Calle, E.E., et al. Are racial disparities in pancreatic cancer explained by smoking and overweight/obesity? *Cancer Epidemiol Biomarkers Prev* 18(9):2397-2405 (2009).
23. Shao, T., Yang, Y.X. Cholecystectomy and the risk of colorectal cancer. *Am J Gastroenterol* 100(8):1813-1820 (2005).
24. Goldacre, M.J., Abisgold, J.D., Seagroatt, V., Yeates, D. Cancer after cholecystectomy: record-linkage cohort study. *Br J Cancer* 92(7):1307-1309 (2005).
25. Schernhammer, E.S., Michaud, D.S., Leitzmann, M.F., Giovannucci, E., Colditz, G.A., Fuchs, C.S. Gallstones, cholecystectomy, and the risk for developing pancreatic cancer. *Br J Cancer* 86(7):1081-1084 (2002).
26. Shibata, A., Mack, T.M., Paganini-Hill, A., Ross, R.K., Henderson, B.E. A prospective study of pancreatic cancer in the elderly. *Int J Cancer* 58(1):46-49 (1994).
27. Goldbohm, R.A., van den Brandt, P.A., van 't Veer, P., Dorant, E., Sturmans, F., Hermus, R.J. Cholecystectomy and colorectal cancer: evidence from a cohort study on diet and cancer. *Int J Cancer* 53(5):735-739 (1993).
28. Nielsen, G.P., Theodors, A., Tulinius, H., Sigvaldason, H. Cholecystectomy and colorectal carcinoma: a total-population historical prospective study. *Am J Gastroenterol* 86(10):1486-1490 (1991).
29. Murphy, N., Jenab, M., Gunter, M.J. Adiposity and gastrointestinal cancers: epidemiology, mechanisms and future directions. *Nat Rev Gastroenterol Hepatol* 15(11):659-670 (2018).
30. Leitzmann, M., Powers, H., Anderson, A.S., Scoccianti, C., Berrino, F., Boutron-Ruault, M.C., et al. European Code against Cancer 4th Edition: Physical activity and cancer. *Cancer Epidemiol* 39(Suppl 1):46-55 (2015).
31. Giovannucci, E., Harlan, D.M., Archer, M.C., Bergenstal, R.M., Gapstur, S.M., Habel, L.A., et al. Diabetes and cancer: a consensus report. *Diabetes Care* 33(7):1674-1685 (2010).
32. Wang, H.H., Liu, M., Clegg, D.J., Portincasa, P., Wang, D.Q. New insights into the molecular mechanisms underlying effects of estrogen on cholesterol gallstone formation. *Biochim Biophys Acta* 1791(11):1037-1047 (2009).
33. Colquhoun, A., Arnold, M., Ferlay, J., Goodman, K.J., Forman, D., Soerjomataram, I., et al. Global patterns of cardia and non-cardia gastric cancer incidence in 2012. *Gut* 64(12):1881-1888 (2015).
34. Stender, S., Frikke-Schmidt, R., Nordestgaard, B.G., Tybjaerg-Hansen, A. Sterol transporter adenosine triphosphate-binding cassette transporter G8, gallstones, and biliary cancer in 62,000 individuals from the general population. *Hepatology* 53(2):640-648 (2011).
35. Bustos, B.I., Pérez-Palma, E., Buch, S., Azócar, L., Riveras, E., Ugarte, G.D., et al. Variants in ABCG8 and TRAF3 genes confer risk for gallstone disease in admixed Latinos with Mapuche native American ancestry. *Sci Rep* 9(1):772 (2019). doi:10.1038/s41598-018-35852-z.
36. Di Ciaula, A., Wang, D.Q., Molina-Molina, E., Baccetto, R.L., Calamita, G., Palmieri, V.O., et al. Bile acids and cancer: direct and environmental-dependent effects. *Ann Hepatol* 16(Suppl 1):87-105 (2015).
37. Li, T., Apte, U. Bile acid metabolism and signaling in cholestasis, inflammation, and cancer. *Adv Pharmacol* 74:263-302 (2017).

38. Costarelli, V., Key, T.J., Appleby, P.N., Allen, D.S., Fentiman, I.S., Sanders, T.A. A prospective study of serum bile acid concentrations and colorectal cancer risk in post-menopausal women on the island of Guernsey. *Br J Cancer* 86(11):1741-1744 (2002).
39. Cross, A.J., Moore, S.C., Boca, S., Huang, W.Y., Xiong, X., Stolzenberg-Solomon, R., et al. A prospective study of serum metabolites and colorectal cancer risk. *Cancer* 120(19):3049-3057 (2014).
40. Jain, K., Sreenivas, V., Velpandian, T., Kapil, U., Garg, P.K. Risk factors for gallbladder cancer: a case-control study. *Int J Cancer* 132(7):1660-1666 (2013).
41. Kuhn, T., Stepien, M., Lopez-Noguerolles, M., Damms-Machado, A., Sookthai, D., Johnson, T., et al. Pre-diagnostic plasma bile acid levels and colon cancer risk: A prospective study. *J Natl Cancer Inst* 112(5):516-524 (2020).
42. Zhang, W., Zhou, L., Yin, P., Wang, J., Lu, X., Wang, X., et al. A weighted relative difference accumulation algorithm for dynamic metabolomics data: long-term elevated bile acids are risk factors for hepatocellular carcinoma. *Sci Rep* (5):8984 (2015). doi:10.1038/srep08984.
43. Feng, H.Y., Chen, Y.C. Role of bile acids in carcinogenesis of pancreatic cancer: An old topic with new perspective. *World J Gastroenterol* 22(33):7463-7477 (2016).
44. Shaffer, E.A. Gallbladder cancer: the basics. *Gastroenterol Hepatol (N Y)* 4(10):737-741 (2008).
45. Shrikhande, S.V., Barreto, S.G., Singh, S., Udwadia, T.E., Agarwal, A.K. Cholelithiasis in gallbladder cancer: coincidence, cofactor, or cause! *Eur J Surg Oncol* 36(6):514-519 (2010).
46. Grivnickov, S.I., Greten, F.R., Karin, M. Immunity, inflammation, and cancer. *Cell* 140(6):883-899 (2010).
47. Pang, Y., Kartsonaki, C., Turnbull, I., Guo, Y., Yang, L., Bian, Z., et al. Metabolic and lifestyle risk factors for acute pancreatitis in Chinese adults: A prospective cohort study of 0.5 million people. *PLoS Med* 15(8):e1002618 (2018). doi:10.1371/journal.pmed.1002618.
48. Gellert-Kristensen, H., Dalila, N., Nielsen, S.F., Nordestgaard, B.G., Tybjaerg-Hansen, A., Stender, S. Identification and replication of six loci associated with gallstone disease. *Hepatology* 70(2):597-609 (2019).
49. Joshi, A.D., Andersson, C., Buch, S., Stender, S., Noordam, R., Weng, L.C., et al. Four susceptibility loci for gallstone disease identified in a meta-analysis of genome-wide association studies. *Gastroenterology* 151(2):351-363 (2016).
50. 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(1):5101 (2018). doi:10.1038/s41467-018-07460-y.
51. Liu, T., Siyin, S.T., Yao, N., Xu, G., Chen, Y.T., Duan, N., et al. Risk of primary liver cancer associated with gallstones and cholecystectomy. *Medicine (Baltimore)* 99(39):e22428 (2020). doi:10.1097/MD.00000000000022428.
52. Fan, Y., Hu, J., Feng, B., Yao, G., Zhai, J., Li, X. Increased risk of pancreatic cancer related to gallstones and cholecystectomy. *Pancreas* 45(4):503-509 (2016).
53. Zhang, Y., Liu, H., Li, L., Gong, Z., He, Y., Dong, Y., et al. Cholecystectomy can increase the risk of colorectal cancer: A meta-analysis of 10 cohort studies. *PLoS One* 12(8):e0181852 (2017). doi:10.1371/journal.pone.0181852.

Figure legends

Figure 1. Meta-analysis of observational associations of symptomatic GSD and cholecystectomy with GI cancer

Boxes represent the hazard ratios (HRs) of specific GI cancer associated with symptomatic GSD or cholecystectomy, separately, with the size of the box inversely proportional to the variance of the logHR.

Table 1. Baseline characteristics of participants by GSD disease status

Variable*	No disease (n=470,839)	Prevalent GSD (n=30,783)	Incident GSD (n=8515)
Age (SD), year	51.8 (10.5)	54.7 (10.1)	53.4 (9.8)
Female, %	57.8	73.7	72.3
Socioeconomic and lifestyle factors			
Urban region, %	44.0	51.4	23.4
≥9 years of education, %	20.7	24.9	20.5
Household income ≥35 000 RMB/year, %	17.8	21.5	18.2
Ever regular smoking, %			
Male	67.8	66.2	66.1
Female	2.8	3.0	3.3
Weekly drinking, %			
Male	33.6	28.2	32.2
Female	2.1	1.9	2.4
Total physical activity (SD), MET-h/day	21.2 (12.7)	20.1 (12.5)	20.4 (12.5)
Sedentary leisure time (SD), h/day	3.0 (1.5)	3.1 (1.7)	3.1 (1.6)
Blood pressure and anthropometry			
SBP (SD), mmHg	131.2 (22.0)	129.7 (21.3)	132.2 (21.9)
RPG (SD), mmol/L	6.1 (2.3)	6.1 (2.4)	6.2 (2.4)
BMI (SD), kg/m ²	23.6 (3.4)	24.0 (3.5)	24.7 (3.5)
Waist circumference (SD), cm	80.1 (9.5)	81.6 (9.8)	83.2 (9.6)
Hip circumference (SD), cm	90.9 (6.9)	91.7 (7.1)	92.5 (6.8)
Waist-to-hip ratio (SD)	0.88 (0.07)	0.89 (0.07)	0.90 (0.07)
Percent body fat (SD), %	27.9 (7.1)	28.7 (7.3)	30.1 (7.3)
Height (SD), cm	158.7 (6.0)	159.1 (6.0)	158.9 (5.9)
Female reproductive factors			
Age at menarche (SD), year	15.5 (2.0)	15.4 (2.0)	15.4 (2.0)
Menopause, %	52.2	54.5	52.7
Number of pregnancies (SD)	3.3 (1.3)	3.4 (1.4)	3.3 (1.4)
Number of liver births (SD)	2.2 (1.7)	2.2 (1.7)	2.3 (1.7)
HBsAg positive, %	3.0	3.6	3.1
Prior disease history, %			
Diabetes	5.7	7.3	7.1
Coronary heart disease	2.8	5.2	3.4
Stroke or TIA	1.7	1.8	1.9
Hypertension	11.4	14.1	12.8
Family history of diabetes	4.8	6.0	5.2
Family history of cancer	13.8	16.2	14.6

* Results were standardised by age, sex, and area (where appropriate). Values are means unless otherwise stated.

Abbreviations: BMI = body mass index, CHD = coronary heart disease, MET = metabolic equivalent of task, RPG = random plasma glucose, SBP = systolic blood pressure, TIA = transient ischaemic attack.

P-values of baseline characteristics between participants with and without GSD: all <0.05.

Table 2. Observational and genetic associations of GSD with GI cancer

Cancer	Observational	Genetic	<i>p</i> -value for heterogeneity
	HR (95% CI)	HR (95% CI)	
Oesophageal (C15)	0.97 (0.78, 1.21)	0.97 (0.91, 1.03)	0.85
Stomach (C16)	0.99 (0.88, 1.14)	1.00 (0.97, 1.04)	0.65
Small intestine (C17)	1.01 (0.58, 1.90)	1.01 (0.81, 1.18)	0.98
Colon (C18)	1.15 (1.00, 1.37)	1.06 (1.02, 1.10)	0.23
Rectal (C20)	1.23 (1.08, 1.46)	1.09 (1.05, 1.14)	0.14
Colorectal (C18-20)	1.13 (1.01, 1.29)	1.08 (1.05, 1.11)	0.29
Liver (C22)	2.01 (1.78, 2.26)	1.22 (1.19, 1.25)	<0.001
Gallbladder (C23)	3.70 (2.88, 4.87)	1.56 (1.49, 1.64)	<0.001
Biliary tract (C24)	2.31 (1.78, 3.07)	1.39 (1.31, 1.46)	<0.001
Pancreatic (C25)	1.38 (1.18, 1.74)	1.16 (1.10, 1.22)	0.11

The observational model was stratified by sex and region, and adjusted for age at baseline, education, smoking, alcohol, total physical activity, BMI, and diabetes. The genetic model adjusted for age, age squared, sex, region, the first 10 principal components, education, smoking, and alcohol.

* *P*-value for heterogeneity between observational and genetic associations.

Table 3. Observational associations of GSD duration with GI cancer

Cancer	Duration of gallstone (years)				Incident
	None	0 to 5	5 to 10	10+	
Oesophageal (C15)					
No. cases	2248	25	27	50	
HR (95% CI)	1.00 (0.95, 1.06)	1.46 (0.92, 2.31)	0.95 (0.66, 1.38)	0.82 (0.62, 1.07)	0.83 (0.47, 1.47)
Stomach (C16)					
No. cases	3087	53	48	161	
HR (95% CI)	1.00 (0.96, 1.04)	154 (1.12, 2.10)	0.88 (0.68, 1.14)	0.98 (0.84, 1.16)	1.00 (0.87, 1.154)
Colorectal (C18-20)					
No. cases	2773	42	71	175	
HR (95% CI)	1.00 (0.95, 1.06)	1.35 (0.87, 2.09)	1.20 (0.91, 1.59)	1.18 (0.98, 1.42)	1.15 (1.01, 1.31)
Colon (C18)					
No. cases	1570	30	31	129	
HR (95% CI)	1.00 (0.94, 1.06)	1.55 (1.02, 2.35)	1.07 (0.79, 1.46)	1.28 (1.06, 1.55)	1.19 (1.01, 1.40)
Rectal (C20)					
No. cases	1551	33	29	132	
HR (95% CI)	1.00 (0.96, 1.04)	1.42 (1.02, 1.97)	1.15 (0.92, 1.43)	1.15 (0.99, 1.33)	1.23 (1.04, 1.45)
Liver (C22)					
No. cases	2492	112	83	217	
HR (95% CI)	1.00 (0.95, 1.05)	3.85 (3.08, 4.80)	1.48 (1.18, 1.84)	2.02 (1.77, 2.31)	1.92 (1.70, 2.16)
Gallbladder (C23)					
No. cases	205	35	21	48	
HR (95% CI)	1.00 (0.86, 1.16)	9.60 (6.19, 14.90)	3.62 (2.38, 5.50)	3.53 (2.65, 4.69)	3.47 (2.63, 4.56)
Biliary tract (C24)					
No. cases	300	38	17	42	
HR (95% CI)	1.00 (0.88, 1.14)	7.34 (4.78, 11.27)	2.06 (1.26, 3.37)	2.50 (1.85, 3.37)	2.08 (1.56, 2.76)
Pancreatic (C25)					
No. cases	692	31	21	60	
HR (95% CI)	1.00 (0.92, 1.09)	3.25 (2.14, 4.93)	1.15 (0.76, 1.74)	1.35 (1.04, 1.74)	1.38 (1.10, 1.47)

Models were stratified by sex and region, and adjusted for age at baseline, education, smoking, alcohol, total physical activity, BMI, and diabetes. The analysis for incident GSD further adjusted for prevalent GSD.

Table 4. Observational and genetic associations of GSD with non-GI cancer

Cancer	Model 1	Model 2	Per 1-SD GRS
	HR (95% CI)	HR (95% CI)	OR (95% CI)
Lung (C34)	0.98 (0.96, 0.99)	1.05 (0.91, 1.22)	1.03 (0.95, 1.12)
Breast (C50)	1.63 (1.17, 2.27)	1.59 (1.14, 2.21)	0.98 (0.78, 1.22)
Cervix (C53)	1.53 (1.15, 2.04)	1.52 (1.14, 2.03)	0.95 (0.77, 1.16)
Endometrial (C54)	1.10 (0.95, 1.28)	1.08 (0.93, 1.26)	0.96 (0.87, 1.05)
Ovary (C56)	1.15 (0.91, 1.45)	1.14 (0.90, 1.44)	0.96 (0.84, 1.10)

Model 1 was stratified by sex and region, and adjusted for age at baseline, education, smoking, and alcohol.

Model 2 further adjusted for total physical activity, BMI, and diabetes (for female cancer) and restricted to never smokers (for lung cancer).

The genetic model adjusted for age, age squared, sex, region, the first 10 principal components, education, smoking, and alcohol.

Figure 1. Meta-analysis of observational associations of symptomatic GSD and cholecystectomy with GI cancer

