




Risk Assessment for Gastrointestinal Diseases via Clinical Dimension and Genome-Wide Polygenic Risk Scores of Type 2 Diabetes: A Population-Based Cohort Study

Tian Fu, Yuhao Sun, Shiyuan Lu, Jianhui Zhao, Lintao Dan, Wenming Shi, Jie Chen, Yan Chen, and Xue Li

Diabetes Care 2024;47(3):418–426 | <https://doi.org/10.2337/dc23-0978>

Risk assessment for gastrointestinal diseases via clinical dimension and genome-wide polygenic risk scores of type 2 diabetes

Participants	Methods	Results																																																
 <p>Genetic factors</p>  <p>Environmental factors</p> <p>19,719 participants with T2DM</p>  <p>354,406 participants without T2DM</p> <p>2006 to 2010</p>	<p>Risk of incident GI diseases for participants with T2DM VS Non T2DM</p> <p>Risk of GI diseases for participants with intermediate/high genetic risk of T2DM VS Low genetic risk of T2DM</p> <p>Follow-up</p>	<table border="1"> <thead> <tr> <th>Gastrointestinal diseases</th> <th>T2DM</th> <th>High genetic risk of T2DM</th> </tr> </thead> <tbody> <tr> <td>Barrett esophagus</td> <td>↑</td> <td></td> </tr> <tr> <td>Gastroesophageal reflux disease</td> <td></td> <td></td> </tr> <tr> <td>Gastritis and duodenitis</td> <td>↑↑</td> <td>↑↑</td> </tr> <tr> <td>Celiac disease</td> <td></td> <td></td> </tr> <tr> <td>Peptic ulcer</td> <td>↑↑</td> <td></td> </tr> <tr> <td>Crohn disease</td> <td>↑</td> <td></td> </tr> <tr> <td>Ulcerative colitis</td> <td></td> <td></td> </tr> <tr> <td>Irritable bowel syndrome</td> <td></td> <td></td> </tr> <tr> <td>Diverticular disease</td> <td>↑↑</td> <td></td> </tr> <tr> <td>Pancreatitis</td> <td>↑↑</td> <td></td> </tr> <tr> <td>Non alcoholic fatty liver disease</td> <td>↑↑</td> <td>↑↑</td> </tr> <tr> <td>Liver cirrhosis</td> <td>↑↑</td> <td></td> </tr> <tr> <td>Biliary disease</td> <td>↑↑</td> <td>↑</td> </tr> <tr> <td>Gastrointestinal tract cancers</td> <td>↑↑</td> <td></td> </tr> <tr> <td>Hepatobiliary and pancreatic cancers</td> <td>↑↑</td> <td></td> </tr> </tbody> </table> <p>2021 for England and Scotland, 2018 for Wales</p>	Gastrointestinal diseases	T2DM	High genetic risk of T2DM	Barrett esophagus	↑		Gastroesophageal reflux disease			Gastritis and duodenitis	↑↑	↑↑	Celiac disease			Peptic ulcer	↑↑		Crohn disease	↑		Ulcerative colitis			Irritable bowel syndrome			Diverticular disease	↑↑		Pancreatitis	↑↑		Non alcoholic fatty liver disease	↑↑	↑↑	Liver cirrhosis	↑↑		Biliary disease	↑↑	↑	Gastrointestinal tract cancers	↑↑		Hepatobiliary and pancreatic cancers	↑↑	
Gastrointestinal diseases	T2DM	High genetic risk of T2DM																																																
Barrett esophagus	↑																																																	
Gastroesophageal reflux disease																																																		
Gastritis and duodenitis	↑↑	↑↑																																																
Celiac disease																																																		
Peptic ulcer	↑↑																																																	
Crohn disease	↑																																																	
Ulcerative colitis																																																		
Irritable bowel syndrome																																																		
Diverticular disease	↑↑																																																	
Pancreatitis	↑↑																																																	
Non alcoholic fatty liver disease	↑↑	↑↑																																																
Liver cirrhosis	↑↑																																																	
Biliary disease	↑↑	↑																																																
Gastrointestinal tract cancers	↑↑																																																	
Hepatobiliary and pancreatic cancers	↑↑																																																	

↑↑ indicate positively significant association; ↑ indicate positively suggestive association; T2DM, type 2 diabetes mellitus; GI, gastrointestinal diseases.

ARTICLE HIGHLIGHTS

- Why did we undertake this study?**
 Population-based data indicated that type 2 diabetes may serve as a risk factor of a wide range of gastrointestinal diseases.
- What is the specific questions we wanted to answer?**
 We aimed to investigate the associations between type 2 diabetes and the risk of several gastrointestinal diseases and explore possible genetic and nongenetic pathways.
- What did we find?**
 Novel associations of genetic variants of type 2 diabetes with gastrointestinal phenotypes, including gastritis, duodenitis, and nonalcoholic fatty liver disease, were revealed.
- What are the implications of our findings?**
 We suggest the necessity of early detection of gastrointestinal disease in type 2 diabetes and possible research direction on association mechanisms.



Risk Assessment for Gastrointestinal Diseases via Clinical Dimension and Genome-Wide Polygenic Risk Scores of Type 2 Diabetes: A Population-Based Cohort Study

Tian Fu,^{1,2,3} Yuhao Sun,^{1,2} Shiyuan Lu,¹ Jianhui Zhao,² Lintao Dan,² Wenming Shi,⁴ Jie Chen,^{2,3} Yan Chen,¹ and Xue Li²

Diabetes Care 2024;47:418–426 | <https://doi.org/10.2337/dc23-0978>

OBJECTIVE

We aimed to evaluate whether individuals with type 2 diabetes (T2D) were at higher risk of developing a wide range of gastrointestinal diseases based on a population-based cohort study.

RESEARCH DESIGN AND METHODS

This study included 374,125 participants free of gastrointestinal disorders at baseline; of them, 19,719 (5.27%) with T2D were followed-up by linking to multiple medical records to record gastrointestinal disease diagnoses. Multivariable Cox models were used to estimate the hazard ratios (HRs) and CIs. Logistic models were used to examine the associations between polygenic risk scores (PRS) and clinical gastrointestinal phenotypes.

RESULTS

During a median follow-up of 12.0 years, we observed the new onset of 15 gastrointestinal diseases. Compared with nondiabetes, participants with T2D had an increased risk of gastritis and duodenitis (HR 1.58, 95% CI 1.51–1.65), peptic ulcer (HR 1.56, 95% CI 1.43–1.71), diverticular disease (HR 1.19, 95% CI 1.14–1.24), pancreatitis (HR 1.45, 95% CI 1.24–1.71), nonalcoholic fatty liver disease (HR 2.46, 95% CI 2.25–2.69), liver cirrhosis (HR 2.92, 95% CI 2.58–3.30), biliary disease (HR 1.18, 95% CI 1.10–1.26), gastrointestinal tract cancers (HR 1.28, 95% CI 1.17–1.40), and hepatobiliary and pancreatic cancer (HR 2.32, 95% CI 2.01–2.67). Positive associations of PRS of T2D with gastritis, duodenitis, and nonalcoholic fatty liver disease were also observed.

CONCLUSIONS

In this large cohort study, we found that T2D was associated with increased risks of a wide range of gastrointestinal outcomes. We suggest the importance of early detection and prevention of gastrointestinal disorders among patients with T2D.

Type 2 diabetes (T2D) is one of the most common chronic metabolic disorders, affecting ~415 million adults worldwide (1). Approximately 5 million deaths occur annually

¹Department of Gastroenterology, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

²School of Public Health and The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

³Department of Gastroenterology, The Third Xiangya Hospital, Central South University, Changsha, China

⁴School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, Hong Kong

Corresponding authors: Xue Li, xueli157@zju.edu.cn, Yan Chen, chenyan72_72@zju.edu.cn, and Jie Chen, med_chenjie@zju.edu.cn

Received 28 May 2023 and accepted 7 December 2023

This article contains supplementary material online at <https://doi.org/10.2337/figshare.24794355>.

T.F., Y.S., and S.L. are joint first authors.

© 2024 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/journals/pages/license>.

due to diabetes and corresponding complications, posing a major global health threat worldwide (1). Therefore, it is urgent and critical to clarify the associations between T2D and different complications. Traditional complications of T2D are mainly concentrated in vascular complications, such as cardiovascular disease, and are investigated widely (1). Gastrointestinal diseases were found to coexist with T2D in epidemiological studies; however, there is a lack of comprehensive assessment on the subsequent risk of gastrointestinal disorders after the diagnosis of T2D.

A recent Mendelian randomization study examined the associations of genetic liability to T2D with the risk of a wide range of gastrointestinal diseases (2). Given the important roles of nongenetic risk factors on the development and progression of T2D (3), further evidence from population-based cohort studies was crucial. However, the paucity of population-level data on the associations of T2D with gastrointestinal complications is still a major gap in population-level monitoring. Several sporadic epidemiological studies with limited sample sizes and heterogeneous study designs provided inconclusive evidence for the association of T2D with gastroesophageal reflux disease (4), diverticular disease (5), nonalcoholic fatty liver disease (NAFLD) (6), and gastrointestinal cancers (7). In addition, how T2D played a role in the incidence of other gastrointestinal diseases, such as inflammatory bowel diseases, which caused an increasingly great influence globally, was still unclear. Homeostasis loss of metabolic traits, such as BMI, leptin, incretin hormones, and lipids, was reported in T2D and related to the disease progression, as well as gastrointestinal and pathologic disorders, such as gut microbiota, in observational studies. Gastrointestinal enzymes, especially liver and pancreatic enzymes, are optimal markers for detecting changes in the physiological state of the organs. Exploring the effects of genetics on the associations would be interesting.

We therefore tried to address these gaps in knowledge by using a well-designed and administered cohort study with a large sample size. We aimed to assess and compare the associations between T2D and the risk of 15 gastrointestinal diseases and explore possible genetic and nongenetic pathways, and we conducted a series of supplementary analyses to test the robustness of the primary analyses.

RESEARCH DESIGN AND METHODS

Study Population

We leveraged participants enrolled from the UK Biobank to conduct a cohort study. In brief, UK Biobank is a large-scale prospective study that recruited ~0.5 million participants in 2006–2010. Participants were asked to finish a touchscreen questionnaire, physical examination, and sample collection, and were followed-up to document the health-related information. The UK Biobank received ethical approval from the North West Multicenter Research Ethic Committee in Manchester (REC reference: 21/NW/0157). In the current study, we excluded participants with gastrointestinal diseases ($n = 112,509$) or type 1 diabetes ($n = 2,936$) at baseline and participants with new-onset T2D ($n = 12,919$) during follow-up, leaving 374,125 participants for primary analyses (Fig. 1 and Supplementary Table 1). All participants included in this study provided signed informed consent.

Ascertainment of T2D and Gastrointestinal Diseases

Ascertainment of T2D at baseline was based on the following criteria (8): 1) diagnoses recorded in hospital inpatients (*International Classification of Diseases* [ICD]-9 codes [250] and ICD-10 codes [E11-E14]), primary care (mapped Read codes to ICD codes), death registry (ICD-10 codes [E11-E14]), or self-report; 2) random blood glucose level ≥ 11.1 mmol/L; 3) blood glycated hemoglobin (HbA_{1c}) level ≥ 48 mmol/mol; or 4) used antidiabetes drugs at baseline. Read codes are a coded thesaurus of clinical terms used in primary care since 1985, including version 2

(Read v2) and version 3 (Clinical Terms Version 3 or Read v3), providing standard vocabulary for clinical records.

The primary outcomes included 15 gastrointestinal disorders diagnosed from any of the medical records (in-patient, primary care data, death registry, and cancer registry): Barrett esophagus, gastroesophageal reflux disease, gastritis and duodenitis, celiac disease, peptic ulcer, Crohn disease, ulcerative colitis, irritable bowel syndrome, diverticulum disease, pancreatitis, NAFLD, liver cirrhosis, biliary disease (including cholangitis, cholecystitis, and cholelithiasis), and gastrointestinal cancers (including gastrointestinal tract cancers and hepatobiliary and pancreatic cancers), ascertained by ICD-9 codes (150–157) and ICD-10 codes (C15–C20, C22–C25).

Construction of Polygenic Risk Score

We constructed the polygenic risk score (PRS) (9) using genetic variants ascertained to be strongly related to T2D ($P < 5 * 10^{-8}$) from a transethnic genome-wide association study (GWAS) of 228,499 case subjects and 1,178,783 control subjects (included European, African American, Hispanic, and Asian populations) (10), to estimate the genetic susceptibility to T2D for the UK Biobank population. When we removed genetic variants in linkage disequilibrium, 497 independent single-nucleotide polymorphisms (SNP) ($r^2 < 0.001$) were used (Supplementary Table 2), and we constructed the PRS by summing up the identified SNP above weighted by effect size on genetic liability to T2D (Supplementary Fig. 1). The constructed PRS

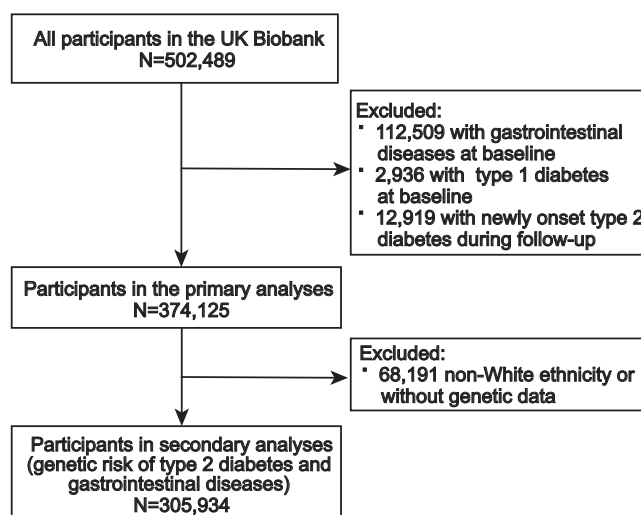


Figure 1—Flowchart of the current study.

predicting T2D in UK Biobank and the current study performed well, with odds ratios of 1.78 (95% CI 1.76–1.80, $P < 0.001$) and 2.06 (95% CI 2.03–2.10, $P < 0.001$) of T2D for per 1-SD increase of PRS. In the current analyses, PRS was divided into three categories as high- (highest quintile), intermediate- (quintiles 2–4), and low- (lowest quintile) genetic-risk groups. In the secondary analysis associated with PRS of T2D and phenotypes of gastrointestinal diseases, 68,191 non-White ethnicities, or without genetic data, were excluded, with 305,934 individuals left for analysis (Fig. 1).

To explore how genetically predicted metabolic traits and gastrointestinal enzymes play roles in gastrointestinal diseases, we additionally calculated the PRS of BMI, leptin, incretin (fasting glucagon-like peptide-1 [GLP-1], GLP-1 2 h, fasting glucose-dependent insulinotropic polypeptide [GIP], and GIP 2 h), lipids (triglyceride, total cholesterol, LDL cholesterol [LDL-C], HDL cholesterol [HDL-C]), liver enzymes (γ -glutamyl transferase, AST, ALT, and alkaline phosphatase), and pancreatic enzymes (trypsin1, trypsin2, trypsin3) to examine the associations of these PRS with the gastrointestinal phenotypes. The independent SNPs of BMI, leptin, incretin, lipids, liver enzymes, and pancreatic enzymes were derived from the present open GWAS studies, and the detailed information of SNPs for each trait are presented in Supplementary Table 3. The corresponding PRS were computed for each individual by summing the product of the allele weighting and the allele dosage across the selected SNPs and divided into low (first quintile), intermediate (2–4 quintile), and high (last quintile) genetic risk according to quintiles.

Assessment of Covariates

We collected information on age, sex, ethnicity, education, Townsend deprivation index (an area-based proxy measure for socioeconomic status), smoking status, alcohol drinking status, and physical activity and diet measured by a healthy diet score, and BMI calculated by height and weight. The healthy diet score was constructed from seven food items, ranging from 0 to 7, with higher scores indicating a healthier diet (11). We also gathered information on medication use (proton pump inhibitors), inflammation (INFLA)-score (12), baseline

cardiovascular diseases, and metabolic abnormalities (13). The INFLA-score contains C-reactive protein, white blood cell and platelet count, and the neutrophil-to-lymphocyte ratio, used to reflect low-grade inflammation, as described in a previous study (12). For missing values of each covariate, the sex-specific median was used to impute for continuous variables, and categorical variables were imputed by plural (<3% missing) and an indicator (>3% missing), respectively. The missing rate of most covariates in the current study was <3%, except for C-reactive protein (6.73%) and the INFLA-score (9.28%) (Supplementary Table 4).

Statistical Analysis

Baseline characteristics of included participants are described as mean (SD) or percentages according to baseline T2D status. Follow-up time was calculated from the baseline date to the date of occurrence of gastrointestinal outcomes of interest, death, loss, or the last data collection for the practice (31 March 2021 for England and Scotland and 28 February 2018 for Wales), whichever occurred first. We used age-scaled Cox regression models to evaluate the hazard ratios (HRs) and CIs. The Schoenfeld residuals method (14) was used to test the proportional hazard assumptions. In primary analysis, we used two age-scaled Cox models separately adjusted for sex, Townsend deprivation index, education, ethnicity, and further adjusted for BMI, smoking status, drinking status, physical activity, and adherence to a healthy diet to assess the associations between T2D and gastrointestinal diseases. Given disease duration, HbA_{1c} levels and medication were common risk factors for T2D-related complications, stratification based on these factors might help identify participants at risk for complications. Thus, we evaluated the risk of gastrointestinal diseases in T2D with different diabetes duration (≤ 5 , 5–10, >10 years), HbA_{1c} levels (≥ 53 mmol/mol or <53 mmol/mol), and diabetes medication use (yes or no) compared with nondiabetes.

In the additional post hoc analysis, we examined whether genetic liability of T2D and the other 17 metabolic traits or enzymes were associated with the 15 gastrointestinal phenotypes. Multivariable logistic regression models were used to evaluate the odds ratio (OR) and 95% CI

for the associations of the constructed PRS (low-, intermediate-, and high-genetic-risk group) of T2D with gastrointestinal outcomes. The associations of T2D-related PRS were also stratified by the status of baseline T2D.

To further evaluate whether the observed association was altered by confounders, we tested the interaction between diabetes and each covariate and recalculated the associations of T2D with gastrointestinal end points stratified by age (≤ 60 , >60 years old), sex (female, male), smoking status (ever, never smoker), drinking status (current, noncurrent drinker), BMI (obesity [≥ 30 kg/m²], nonobesity [< 30 kg/m²]), and proton pump inhibitor use (yes, no).

We also conducted several sensitivity analyses to test the robustness of our main findings: 1) changed the ICD-10 codes (E11–E14) in the diagnoses recorded used in the primary analyses to ICD-10 code E11 (15); defined baseline T2D as diagnosed by medical records or self-report plus any of the following: abnormal blood glucose, abnormal HbA_{1c}, or use of antidiabetes drugs; 2) additionally adjusted for C-reactive protein concentration, INFLA-score, proton pump inhibitor use, baseline cardiovascular diseases, and baseline metabolic abnormalities based on the primary analysis; and 3) adopted the step function analyses for Cox models that did not satisfy the proportional hazards assumption (16). Generally, when the proportional hazards assumption is not fulfilled, modeling time-varying coefficients by step function (split the analysis time into several intervals) is an alternative method (16). All analyses were conducted using R 4.2.1 software. All tests were two-sided, and association with P value <0.05 was considered suggestive, and a P value <0.05/15 (after the Bonferroni test) was deemed significant.

Data and Resource Availability

Data are available from the UK Biobank (www.ukbiobank.ac.uk/).

RESULTS

This study included 374,125 individuals, among whom 19,719 (5.27%) had T2D, with a mean age of 55.90 (SD 8.14) years. Participants with T2D are more likely to be older ($P < 0.001$), male ($P < 0.001$), with higher BMI ($P < 0.001$) (Table 1). During a median follow-up of

Table 1—Baseline characteristics of study participants by baseline T2D status

Characteristic	Overall (N = 374,125)	Non-T2D (n = 354,406)	T2D (n = 19,719)	P values
Age, mean (SD), years	55.90 (8.14)	55.72 (8.15)	59.10 (7.36)	<0.001
Female sex	201,493 (53.9)	194,229 (54.8)	7,264 (36.8)	<0.001
Townsend deprivation index, mean (SD)	−1.38 (3.05)	−1.43 (3.02)	−0.44 (3.42)	<0.001
BMI, mean (SD), kg/m ²	27.06 (4.58)	26.82 (4.38)	31.28 (5.83)	<0.001
College degree and above	128,693 (34.4)	123,769 (34.9)	4,924 (25.0)	<0.001
White	352,271 (94.2)	335,566 (94.7)	16,705 (84.7)	<0.001
Ever smoking	161,979 (43.3)	151,728 (42.8)	10,251 (52.0)	<0.001
Current drinker	347,544 (92.9)	330,801 (93.3)	16,743 (84.9)	<0.001
Regular physical activity	299,842 (80.1)	285,632 (80.6)	14,210 (72.1)	<0.001
Adherence to a healthy diet	261,608 (69.9)	248,379 (70.1)	13,229 (67.1)	<0.001
C-reactive protein, mean (SD), mg/L	2.30 (3.95)	2.24 (3.88)	3.30 (4.94)	<0.001
INFLA-score, mean (SD)	−0.39 (5.75)	−0.49 (5.73)	1.33 (5.84)	<0.001
Use of proton pump inhibitors	13,289 (3.6)	11,524 (3.3)	1,765 (9.0)	<0.001
Baseline cardiovascular diseases	104,005 (27.8)	91,189 (25.7)	12,816 (65.0)	<0.001
Baseline metabolic abnormalities	164,874 (44.1)	147,447 (41.6)	17,427 (88.4)	<0.001

Data are presented as *n* (%), unless shown otherwise as mean (SD).

12.0 years, patients with new onset of gastrointestinal disease were documented, including 3,246 with Barrett esophagus, 26,982 with gastroesophageal reflux disease, 24,310 with gastritis and duodenitis, 1,116 with celiac disease, 5,302 with peptic ulcer, 729 with Crohn disease, 1,466 with ulcerative colitis, 5,469 with irritable bowel syndrome, 32,630 with diverticulum diseases, 1,613 with pancreatitis, 3,723 with NAFLD, 1,870 with cirrhosis, 11,915 with biliary disease, 6,570 with gastrointestinal tract cancers, and 1,700 with hepatobiliary and pancreatic cancers.

We observed positive associations between T2D and subsequently increased risk of 11 gastrointestinal diseases. Specifically, compared with nondiabetes, participants with T2D had higher risks of gastritis and duodenitis (HR 1.58, 95% CI 1.51–1.65, $P < 0.001$), peptic ulcer (HR 1.56, 95% CI 1.43–1.71, $P < 0.001$), diverticular disease (HR 1.19, 95% CI 1.14–1.24, $P < 0.001$), pancreatitis (HR 1.45, 95% CI 1.24–1.71, $P < 0.001$), NAFLD (HR 2.46, 95% CI 2.25–2.69, $P < 0.001$), liver cirrhosis (HR 2.92, 95% CI 2.58–3.30, $P < 0.001$), biliary disease (HR 1.18, 95% CI 1.10–1.26, $P < 0.001$), gastrointestinal tract cancers (HR 1.28, 95% CI 1.17–1.40, $P < 0.001$), and hepatobiliary and pancreatic cancer (HR 2.32, 95% CI 2.01–2.67, $P < 0.001$). T2D was suggestively associated with higher risks of Barrett esophagus (HR 1.20, 95% CI

1.05–1.36, $P = 0.006$) and Crohn disease (HR 1.36, 95% CI 1.02–1.80, $P = 0.036$). In contrast, we did not observe any significant associations with gastroesophageal reflux disease, celiac disease, ulcerative colitis, or irritable bowel syndrome (Fig. 2). For Cox models for risk of gastritis and duodenitis ($P < 0.05$), diverticular disease ($P < 0.05$), and NAFLD ($P < 0.05$) not satisfying the proportional hazard assumption, we adopted the step function analyses for a series analyses. We observed consistent positive associations of T2D with gastritis and duodenitis, diverticular disease, and NAFLD, with almost higher strengths (Supplementary Table 14).

Among the 15 gastrointestinal phenotypes examined associations with the PRS of T2D, we found significant associations with gastritis and duodenitis (OR 1.08, 95% CI 1.03–1.13, $P = 0.001$) and NAFLD (OR 1.42, 95% CI 1.26–1.59, $P < 0.001$), and a suggestive association with biliary disease (OR 1.08, 95% CI 1.01–1.15, $P = 0.022$) in the high-genetic-risk compared with the low-genetic-risk group. We did not observe any associations between PRS of T2D and other investigated gastrointestinal diseases (Fig. 3 and Supplementary Table 5). The above associations for NAFLD remained significant in participants without baseline T2D (OR 1.23, 95% CI 1.08–1.40, $P = 0.002$) (Supplementary Table 6). As for

other genetically predicted traits, we found that higher PRS of BMI was associated with increased risk of pancreatitis, NAFLD, and biliary disease, while higher genetically predicted levels of leptin, incretin (GIP 2 h), and lipids (triglyceride, total cholesterol, LDL-C) showed inverse associations with NAFLD, respectively. As for gastrointestinal enzymes, higher PRS of liver enzymes showed higher risk of NAFLD and liver cirrhosis, but showed inconsistent findings for risk of biliary disease. No significant associations with gastrointestinal diseases for genetically predicted trypsins were found (Supplementary Table 7).

When we grouped T2D according to diabetes duration, HbA_{1c} level, or diabetes medication, the association remained consistent between T2D and the above nine gastrointestinal diseases. We also observed P trend for almost all diseases subtyped by disease duration, except for ulcerative colitis and irritable bowel syndrome. In most related diseases, participants with longer disease duration (5–10 years or >10 years) seemed to have higher estimates than general T2D individuals. Considering the dichotomy, we did not conduct the trend analyses according to HbA_{1c} and antidiabetes medication. Furthermore, compared with nondiabetes, we found a higher risk of Barrett esophagus in T2D with 5–10 years

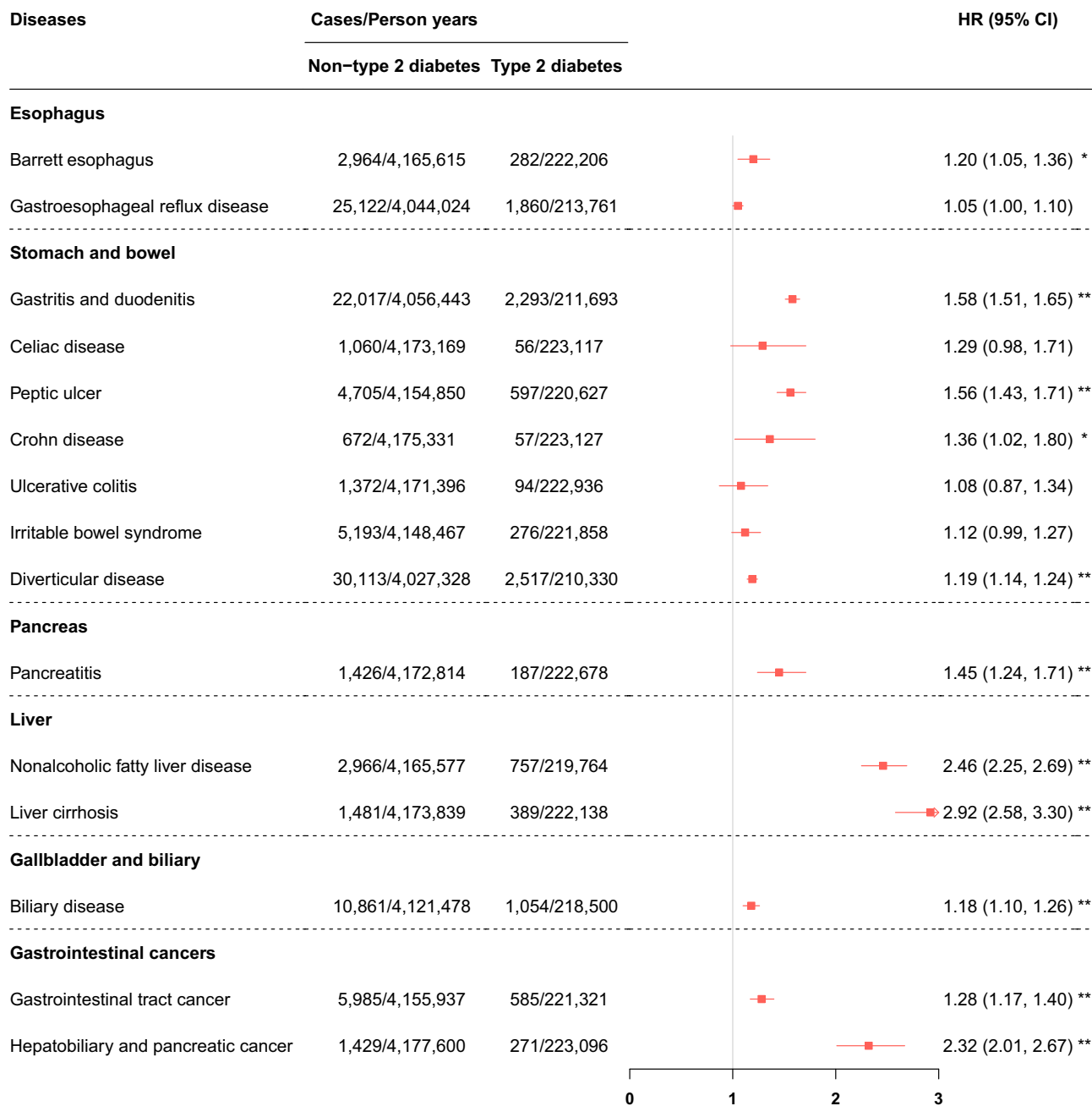


Figure 2—Adjusted HRs for incident gastrointestinal diseases according to baseline T2D. HRs (95% CIs) were calculated by age-scaled Cox models adjusted for sex, Townsend deprivation index, education, ethnicity, BMI, smoking status, drinking status, physical activity, and adherence to a healthy diet. The non-T2D group was set as the reference. $P < 0.05/15$ after Bonferroni correction was considered significant. * $P < 0.05$ and $P > 0.05/15$ and ** $P < 0.05/15$ in the model 2.

or use of diabetes medication, and an increased risk of gastroesophageal reflux disease in T2D with 5–10 years or >10 years, which were not observed in the primary analyses (Supplementary Tables 8–10).

In further analyses, the associations still existed in different subgroups, while we found that interaction between diabetes and each covariate, especially BMI, when associated with different gastrointestinal diseases (P interaction < 0.05)

(Supplementary Table 11). Overall, the risk of gastrointestinal outcomes in T2D seemed to be higher in participants who were men, nonobese, ever smoked, and not currently drinkers, although the effects of these covariates differed in direction and magnitude (Supplementary Table 11). In addition, the primary findings were robust in a series of sensitivity analyses. HRs and significance also did not differ substantially when the definition of T2D was changed, additionally

adjusted for C-reactive protein, INFLA-score, proton pump inhibitors use, baseline cardiovascular diseases, or metabolic abnormalities based on the primary Cox models (Supplementary Tables 12 and 13).

CONCLUSIONS

This population-based cohort study provided strong epidemiological evidence for the positive independent associations between T2D and a wide range of

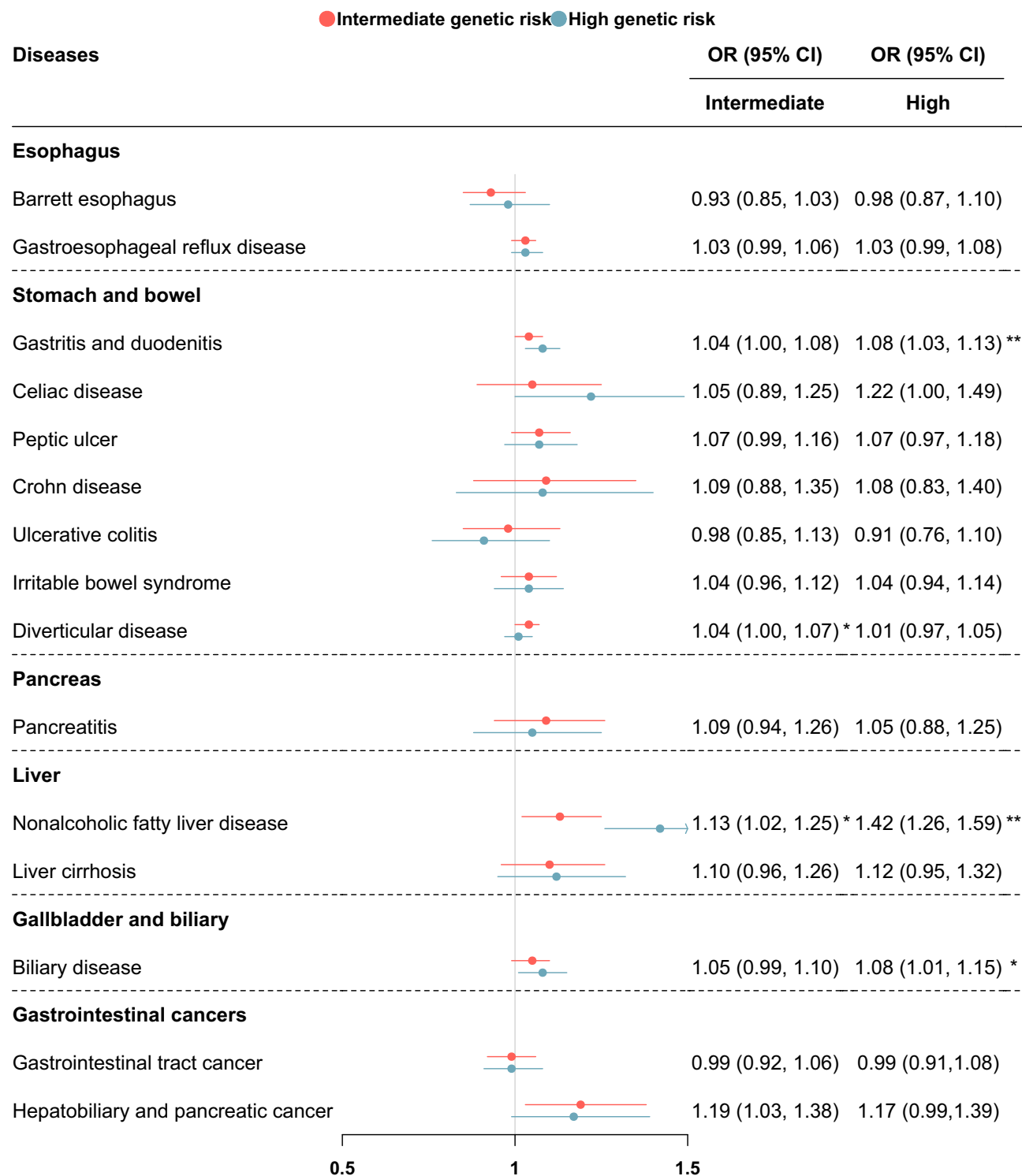


Figure 3—ORs and CIs for the association between the PRS of T2D and gastrointestinal diseases. OR (95% CI) was calculated by logistic models adjusted for sex, Townsend deprivation index, education, ethnicity, BMI, smoking status, drinking status, physical activity, and adherence to a healthy diet. The low-genetic-risk group was the reference group. $P < 0.05/15$ after Bonferroni correction was considered significant. * $P < 0.05$ and $P > 0.05/15$ and ** $P < 0.05/15$.

gastrointestinal diseases, which have not been comprehensively investigated before. We also provided evidence for the novel association between genetic liability of T2D and the risk of gastritis and duodenitis and NAFLD, as well as the

roles of genetically predicted traits of BMI, leptin, incretin, lipids, and liver and pancreas enzymes on gastrointestinal disease. In brief, our findings draw attention to the high risks of incident gastritis and duodenitis, diverticular disease, peptic

ulcer, hepatobiliary disease, and gastrointestinal cancers in T2D, with genetically related mechanisms worth noting and exploring.

Several studies have put forward the positive associations of T2D with Barrett

esophagus (17), gastroesophageal reflux disease (4), diverticular disease (5), pancreatitis (18), NAFLD (6) and advanced liver diseases (19), and biliary disease (18), although with small sample sizes and inconsistent study design. Evidence from a majority of Mendelian studies (2) reported the relationships between T2D and gastroesophageal reflux disease (20), pancreatitis (21), cholelithiasis (22), NAFLD (23), and diverticular disease (24). Our study, based on a large-scale cohort, complemented findings from the previous studies, as well as provided new insight into the increased risk of gastritis and duodenitis and peptic ulcer in T2D. Although there was weak evidence of the positive association between T2D and gastroesophageal reflux disease here, we found significant associations among diabetes duration-specific (5–10 years and >10 years) individuals. Studies showed that higher BMI, smoking, and alcohol drinking, are important independent risk factors for gastroesophageal reflux disease (20,25). In the current study, the association existed when we adjusted sex, Townsend deprivation index, education, and ethnicity, but was not significant when we further adjusted BMI and other lifestyle risk factors. Therefore, combined with the previous evidence, obesity and poor lifestyle, instead of T2D, may play an important role in the development of gastroesophageal reflux disease. The disparity in the risk of gastroesophageal reflux disease in T2D among studies could be attributed to the different characteristics for diseases, study design, or real differences in the epidemiology in these populations. We evaluated the risk of celiac disease, Crohn disease, ulcerative colitis, and irritable bowel syndrome and found null associations, which have not been investigated before.

Possible risk of cancer among individuals with T2D has long been speculated, and reports from multiple epidemiological and Mendelian studies have claimed an increased risk of esophagus (26), liver (27), pancreas (26), and colorectal cancer (28). A recent more rigorous meta-analysis also found positive associations of T2D with intrahepatic cholangiocarcinoma cancer and colorectal cancer after critical appraisal (7). However, the authors pointed out that although most included studies have strong claims of significance for the associations, only a minority have robust supporting evidence without hints of bias. We separately

examined the risk of gastrointestinal tract cancer and hepatobiliary and pancreatic cancer among individuals with T2D and provided additional epidemiological evidence for the positive associations. This study, based on the large sample size and a comprehensive diagnosis of T2D, may better characterize the associations between T2D and risk of the gastrointestinal cancer diagnosis.

Several existing mechanisms also support our findings for the association of T2D and multiple gastrointestinal disorders. Individuals with T2D had defective insulin secretion and glucose homeostasis, which can lead to a metabolic imbalance responsible for the development of multiple diseases (29). Several investigators have assessed the potential pathways by which T2D can increase hepatobiliary disease, especially NAFLD risk, and insulin resistance seems to be one of the key events and plays a mediating role (30). The presence of metabolic disorders represented by insulin resistance and triglyceride metabolism, leading to fat accumulation, oxidative stress, and release of inflammatory mediators, were exactly one of the core pathogenetic causes of intestinal dysbiosis and numerous gastrointestinal disorders (31), especially gastrointestinal cancers (32) and NAFLD (19). Leptin was taken as insulin mimetic to normalize glucose levels in states of insulin deficiency through multiple mechanisms (33). Similarly, GIP receptor activation has been considered as a promising therapeutic concept, given GIP can improve insulin sensitivity and lipid homeostasis through specific metabolic effects (34). In our analyses, we found that genetically predicted higher levels of leptin and GIP 2 h were associated with lower risk of NAFLD and biliary diseases. The protective effects of genetically predicted leptin and GLP against hepatobiliary and pancreatic diseases provided additional evidence for impaired insulin metabolism as a major mechanism underlying the association between T2D and the development of gastrointestinal diseases. NAFLD risk was associated with hyperglycemia and higher insulin levels in several previous studies (35). Findings from our study are a logical extension of the previous studies, providing a more clinically relevant and easily measured state of disease. Additionally, experimental data showed that diabetes

aggravated pancreatitis by inhibiting the regeneration of exocrine tissue, leading to strong atrophy of the pancreas (36). Although trypsin activation can be interpreted as the main step of pancreatitis onset, we did not observe any significant association between the PRS of trypsin and pancreatitis.

Another possibility is the role of obesity. As a common comorbidity of T2D, obesity also served as a risk factor for a series of gastrointestinal disorders (37) and cancers (38). Thus, the potential mediating effect of obesity in T2D and gastrointestinal diseases deserved to be explored. In genetic analyses of the PRS of BMI, we found a positive association between the PRS of BMI and the increased risk of pancreatitis, NAFLD, and biliary disease, reinforcing the explanation that obesity plays an important role in the risk of gastrointestinal diseases. However, when we controlled the confounding effects of obesity in the model by adjusting for BMI, a common measure of general obesity, the associations existed, without great change. Therefore, whether and how (e.g., by abdominal visceral fat) obesity plays roles on the effects of T2D on the development of gastrointestinal disorders needs to be supported by more genetic and observational evidence. A prior study also revealed that visceral obesity increases the potential nonmechanical pathway for esophageal injury in diabetes (17). Interestingly, when grouping T2D according to diabetes duration, HbA_{1c} level, or antidiabetes medication, we found that longer disease duration seems to show a trend toward higher point estimates as well as novel associations that individuals with T2D at longer disease duration or under treatment were at higher risk of Barrett esophagus and gastroesophageal reflux disease, indicating specific effects of disease features on progression and complications, and deserved more precise categorization for better management. Overall, metabolic pathways associated with T2D should be further explored in the pathogenesis of gastrointestinal diseases.

In the genetic analysis, we revealed the relationship between genetic variants of diabetes and clinically diagnosed gastritis and duodenitis and NAFLD, consistent with the positive associations we observed in the primary cohort study. And our additional genetic study might offer new insight into possible common

genetic pathways leading to the three diseases. Previous studies put forward that genetic polymorphism in the human leptin receptor (LEPR) gene associated with T2D and NAFLD (39), possibly through pathways of regulating lipid metabolism and insulin sensitivity. Correspondingly, our findings on the associations with gastrointestinal disease risk of genetically predicted lipids indicated the potential roles of lipid metabolism to some extent. Our genetic analyses also revealed the effects of liver enzymes on liver diseases. However, there are no such studies for gastritis and duodenitis, and how genetic factors play roles in the correlation between T2D and gastritis and duodenitis deserves more exploration. Overall, our study with the large sample size added to the literature demonstrating the value of studying genetics, diabetes, and gastrointestinal diseases, especially gastritis, duodenitis, and NAFLD. Noteworthy, the area under the curves for PRS in the UK Biobank and the current analysis predicting T2D were 0.645 and 0.681, respectively. Participants with higher PRS scores constructed by the GWAS we used were reported with higher risk for T2D (10), in line with our findings.

Of note, there are strengths of our study, including the large sample size and the cohort study design. Most importantly, we compared the risk of 15 gastrointestinal end points in T2D in the same study, which has not been considered before.

Several limitations also exist in the current study. As an observational study, we cannot avoid the bias of confounder and reverse causation, although we tried to adjust for multiple covariates that may cause effects on the association of T2D and gastrointestinal outcomes. We further explored whether the risk of gastrointestinal outcomes changed in diabetes with different HbA_{1c} levels, disease duration, or diabetes medication use, considering a previous study indicated that the duration of diabetes is also an important factor for gastrointestinal symptoms (40).

Second, heterogeneity for the diagnosis of T2D was relatively large in previous studies. We adopted the most common diagnostic methods in the current study and changed to a new diagnostic method in the sensitivity analyses, with consistent associations observed.

Finally, the participants leveraged in our study were from the UK Biobank,

most of whom were of European descent at an older age, limiting the generalization of our findings to other populations but not affecting the valid measures of association.

Conclusion

Our findings revealed T2D may serve as a risk factor for a wide range of gastrointestinal diseases, and observed associations between genetic variants of T2D and gastrointestinal phenotypes. Overall, considering the great influence caused by diabetes-related complications, this study suggested the necessity of early detection of gastrointestinal disease in T2D and possible research direction on association mechanisms although more causal evidence was needed.

Acknowledgments. This research was conducted using data from the UK Biobank study under application number 66354. The authors appreciate the participants and team of the UK Biobank for their participation and assistance.

Funding. This work was supported by the National Nature Science Foundation of China (grant no. 82204019) and the Natural Science Fund for Distinguished Young Scholars of Zhejiang Province (LR22H260001) to X.L.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. T.F. contributed to methodology, formal analysis, and to writing the manuscript. Y.S. contributed to methodology, formal analysis, and to writing, reviewing, and editing the manuscript. S.Y.L. contributed to methodology, formal analysis, and to writing the manuscript. J.Z. contributed to formal analysis and to reviewing and editing the manuscript. L.D. contributed to formal analysis and to writing, reviewing, and editing the manuscript. W.S. contributed to methodology and to reviewing and editing the manuscript. J.C. contributed to conceptualizing the study, methodology, formal analysis, and to reviewing and editing the manuscript. Y.C. contributed to conceptualizing the study, methodology, formal analysis, and to writing, reviewing, and editing the manuscript. X.L. contributed to conceptualizing the study, methodology, formal analysis, and to writing, reviewing, and editing the manuscript. All authors approved the final version of the manuscript. X.L. and J.C. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the analysis.

References

- Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nat Rev Endocrinol* 2018;14:88–98
- Chen J, Yuan S, Fu T, et al. Gastrointestinal consequences of type 2 diabetes mellitus and impaired glycemic homeostasis: a Mendelian

- randomization study. *Diabetes Care* 2023;46:828–835
- Talmud PJ, Hingorani AD, Cooper JA, et al. Utility of genetic and non-genetic risk factors in prediction of type 2 diabetes: Whitehall II prospective cohort study. *BMJ* 2010;340:b4838
- Sun XM, Tan JC, Zhu Y, Lin L. Association between diabetes mellitus and gastroesophageal reflux disease: a meta-analysis. *World J Gastroenterol* 2015;21:3085–3092
- Lin X, Li J, Ying M, Wei F, Xie X. Diabetes increases morbidities of colonic diverticular disease and colonic diverticular hemorrhage: a systematic review and meta-analysis. *Am J Ther* 2017;24:e2113–e221
- Younossi ZM, Gramlich T, Matteoni CA, Boparai N, McCullough AJ. Nonalcoholic fatty liver disease in patients with type 2 diabetes. *Clin Gastroenterol Hepatol* 2004;2:262–265
- Tsilidis KK, Kasimis JC, Lopez DS, Ntzani EE, Ioannidis JP. Type 2 diabetes and cancer: umbrella review of meta-analyses of observational studies. *BMJ* 2015;350:g7607
- Orliac EJ, Trejo Banos D, Ojavee SE, et al. Improving GWAS discovery and genomic prediction accuracy in biobank data. *Proc Natl Acad Sci U S A* 2022;119:e2121279119
- Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. *Genome Med* 2020;12:44
- Vujkovic M, Keaton JM, Lynch JA, et al.; HPAP Consortium; Regeneron Genetics Center; VA Million Veteran Program. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nat Genet* 2020;52:680–691
- Fu T, Ye S, Sun Y, Dan L, Wang X, Chen J. Greater adherence to cardioprotective diet can reduce inflammatory bowel disease risk: a longitudinal cohort study. *Nutrients* 2022;14:4058
- Shi H, Schwenen LJS, Ter Horst R, et al. Low-grade inflammation as mediator between diet and behavioral disinhibition: a UK Biobank study. *Brain Behav Immun* 2022;106:100–110
- Grundy SM, Cleeman JI, Daniels SR, et al.; American Heart Association; National Heart, Lung, and Blood Institute. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. *Circulation* 2005;112:2735–2752
- Xue Y, Schifano ED. Diagnostics for the Cox model. *Commun Stat Appl Methods* 2017;24:583–604
- Eastwood SV, Mathur R, Atkinson M, et al. Algorithms for the capture and adjudication of prevalent and incident diabetes in UK Biobank. *PLoS One* 2016;11:e0162388
- Zhang Z, Reinikainen J, Adeleke KA, Pieterse ME, Groothuis-Oudshoorn CGM. Time-varying covariates and coefficients in Cox regression models. *Ann Transl Med* 2018;6:121
- Iyer PG, Borah BJ, Heien HC, Das A, Cooper GS, Chak A. Association of Barrett's esophagus with type II diabetes mellitus: results from a large population-based case-control study. *Clin Gastroenterol Hepatol* 2013;11:1108–1114.e5
- Noel RA, Braun DK, Patterson RE, Bloomgren GL. Increased risk of acute pancreatitis and biliary disease observed in patients with type 2 diabetes:

- a retrospective cohort study. *Diabetes Care* 2009;32:834–838
19. Stefan N, Cusi K. A global view of the interplay between non-alcoholic fatty liver disease and diabetes. *Lancet Diabetes Endocrinol* 2022;10:284–296
20. Yuan S, Larsson SC. Adiposity, diabetes, lifestyle factors and risk of gastroesophageal reflux disease: a Mendelian randomization study. *Eur J Epidemiol* 2022;37:747–754
21. Yuan S, Giovannucci EL, Larsson SC. Gallstone disease, diabetes, calcium, triglycerides, smoking and alcohol consumption and pancreatitis risk: Mendelian randomization study. *NPJ Genom Med* 2021;6:27
22. Yuan S, Gill D, Giovannucci EL, Larsson SC. Obesity, type 2 diabetes, lifestyle factors, and risk of gallstone disease: a Mendelian randomization investigation. *Clin Gastroenterol Hepatol* 2022;20:e529–e537
23. Yuan S, Chen J, Li X, et al. Lifestyle and metabolic factors for nonalcoholic fatty liver disease: Mendelian randomization study. *Eur J Epidemiol* 2022;37:723–733
24. Yuan S, Larsson SC. Genetically predicted adiposity, diabetes, and lifestyle factors in relation to diverticular disease. *Clin Gastroenterol Hepatol* 2022;20:1077–1084
25. Cheng Y, Kou F, Liu J, Dai Y, Li X, Li J. Systematic assessment of environmental factors for gastroesophageal reflux disease: an umbrella review of systematic reviews and meta-analyses. *Dig Liver Dis* 2021;53:566–573
26. Yuan S, Kar S, Carter P, et al. Is type 2 diabetes causally associated with cancer risk? Evidence from a two-sample Mendelian randomization study. *Diabetes* 2020;69:1588–1596
27. Pang Y, Kartsonaki C, Turnbull I, et al. Diabetes, plasma glucose, and incidence of fatty liver, cirrhosis, and liver cancer: a prospective study of 0.5 million people. *Hepatology* 2018;68:1308–1318
28. Ma Y, Yang W, Song M, et al. Type 2 diabetes and risk of colorectal cancer in two large U.S. prospective cohorts. *Br J Cancer* 2018;119:1436–1442
29. Galicia-Garcia U, Benito-Vicente A, Jebari S, et al. Pathophysiology of type 2 diabetes mellitus. *Int J Mol Sci* 2020;21:6275
30. Hazlehurst JM, Woods C, Marjot T, Cobbold JF, Tomlinson JW. Non-alcoholic fatty liver disease and diabetes. *Metabolism* 2016;65:1096–1108
31. Bhattacharyya A, Chattopadhyay R, Mitra S, Crowe SE. Oxidative stress: an essential factor in the pathogenesis of gastrointestinal mucosal diseases. *Physiol Rev* 2014;94:329–354
32. Lin S, Li Y, Zamyatnin AA Jr, Werner J, Bazhin AV. Reactive oxygen species and colorectal cancer. *J Cell Physiol* 2018;233:5119–5132
33. Marino JS, Xu Y, Hill JW. Central insulin and leptin-mediated autonomic control of glucose homeostasis. *Trends Endocrinol Metab* 2011;22:275–285
34. Karagiannis T, Avgerinos I, Liakos A, et al. Management of type 2 diabetes with the dual GIP/GLP-1 receptor agonist tirzepatide: a systematic review and meta-analysis. *Diabetologia* 2022;65:1251–1261
35. Murphy N, Song M, Papadimitriou N, et al. Associations between glycemic traits and colorectal cancer: a Mendelian randomization analysis. *J Natl Cancer Inst* 2022;114:740–752
36. Zechner D, Spitzner M, Bobrowski A, Knapp N, Kuhla A, Vollmar B. Diabetes aggravates acute pancreatitis and inhibits pancreas regeneration in mice. *Diabetologia* 2012;55:1526–1534
37. Eslick GD. Gastrointestinal symptoms and obesity: a meta-analysis. *Obes Rev* 2012;13:469–479
38. Park J, Morley TS, Kim M, Clegg DJ, Scherer PE. Obesity and cancer—mechanisms underlying tumour progression and recurrence. *Nat Rev Endocrinol* 2014;10:455–465
39. Lu H, Sun J, Sun L, Shu X, Xu Y, Xie D. Polymorphism of human leptin receptor gene is associated with type 2 diabetic patients complicated with non-alcoholic fatty liver disease in China. *J Gastroenterol Hepatol* 2009;24:228–232
40. Bytzer P, Talley NJ, Leemon M, Young LJ, Jones MP, Horowitz M. Prevalence of gastrointestinal symptoms associated with diabetes mellitus: a population-based survey of 15,000 adults. *Arch Intern Med* 2001;161:1989–1996