Liver function and risk of type 2 diabetes: bidirectional Mendelian randomization study

N. Maneka G. De Silva[1,2,3,†], m.de-silva@imperial.ac.uk

Maria Carolina Borges[1,2,†], m.c.borges@bristol.ac.uk

Aroon Hingorani[4,5], a.hingorani@ucl.ac.uk

Jorgen Engmann[4], j.engmann@ucl.ac.uk

Tina Shah[4], t.shah@ucl.ac.uk

Xiaoshuai Zhang[6,7,8], zhxiaoshuai@hotmail.com

Jian'an Luan[6], Jianan.Luan@mrc-epid.cam.ac.uk

Claudia Langenberg[6], Claudia.Langenberg@mrc-epid.cam.ac.uk

Andrew Wong[9], andrew.wong@ucl.ac.uk

Diana Kuh[9], d.kuh@ucl.ac.uk

John C Chambers[3,10-13], john.chambers@imperial.ac.uk

Weihua Zhang[11,12], weihua.zhang@imperial.ac.uk

Marjo-Ritta Jarvelin[11,14,15], m.jarvelin@imperial.ac.uk

Sylvain Sebert[14], sylvain.sebert@oulu.fi

Juha Auvinen[14,16], Juha.auvinen@oulu.fi

UCLEB consortium

Tom R Gaunt[1,2]*, Tom.Gaunt@bristol.ac.uk

Deborah A Lawlor[1,2]*, D.A.Lawlor@bristol.ac.uk

[1] MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK

[2] Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

[3] Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK

Diabetes Publish Ahead of Print, published online May 14, 2019

[4] UCL Institute of Cardiovascular Science, Research Department of Population Science and Experimental Medicine, Centre for Translational Genomics, University College London, London, United Kingdom

[5] Farr Institute, University College London, London, United Kingdom

[6] MRC Epidemiology Unit, University of Cambridge, Cambridge, UK

[7] Department of Epidemiology and Biostatistics, School of Public Health and Management, Binzhou Medical University, Yantai, Shandong, China

[8] Department of Biostatistics, School of Public Health, Shandong University, Jinan, Shandong, China

[9] MRC Unit for Lifelong Health and Ageing at UCL, University College London, London, United Kingdom

[10] Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore 308232, Singapore

[11] Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK

[12] Department of Cardiology, Ealing Hospital, Middlesex UB1 3HW, UK

[13] Imperial College Healthcare NHS Trust, Imperial College London, London W12 0HS, UK

[14] Center for Life Course Health Research, University of Oulu, Oulu, Finland

[15] Biocenter Oulu, Oulu, Finland

[16] Oulunkaari Health Center, Ii, Finland

[†] These authors contributed equally as joint-first authors

[*] These authors contributed equally as joint-senior authors

Running title: Liver dysfunction and type 2 diabetes

Word count: 4,785

Number of tables: 2

Number of figures: 2

Corresponding author: Prof DA Lawlor, MRC IEU at the University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN. Email: d.a.lawlor@bristol.ac.uk. Telephone: +44 (0)117 3310096

For Twitter:

Tweet (280 characters or less):
Bi-directional Mendelian randomization study supports cross-talk between liver function and insulin resistance/type 2 diabetes

Twitter handles and relevant hashtags:
@MCarol_Borges
@tomrgaunt
@DeborahLawlor2
@mrc_ieu
@BristolUni

Figure that best summarizes article:
Figure 2

ABSTRACT

Liver dysfunction and type 2 diabetes (T2D) are consistently associated. However, it is currently unknown whether liver dysfunction contributes to, results from or is merely correlated with T2D due to confounding. We used Mendelian randomization (MR) to investigate the presence and direction of any causal relation between liver function and T2D risk including up to 64,094 T2D cases and 607,012 controls. Several biomarkers were used as proxies of liver function [i.e. alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), and gamma-glutamyl transferase (GGT)]. Genetic variants strongly associated with each liver function marker were used to investigate the effect of liver function on T2D risk. In addition, genetic variants strongly associated with T2D risk and with fasting insulin were used to investigate the effect of predisposition to T2D and insulin resistance, respectively, on liver function. Genetically predicted higher circulating ALT and AST were related to increased risk of T2D. There was a modest negative association of genetically predicted ALP with T2D risk and no evidence of association between GGT and T2D risk. Genetically predisposition to higher fasting insulin, but not to T2D, was related to increased circulating ALT. Since circulating ALT and AST are markers of NAFLD, these findings provide some support for insulin resistance resulting in NAFLD, which in turn increases T2D risk.

INTRODUCTION

Observational studies have repeatedly reported that liver dysfunction and type 2 diabetes (T2D) are associated (1-3). Since the liver plays a core role in the regulation of glucose homeostasis, it is hypothesized that liver dysfunction might increase T2D risk by exacerbating hepatic insulin resistance, leading to overstimulation of hepatic gluconeogenesis (4). Alternatively, it is suggested that insulin resistance and T2D might disturb liver function, possibly via an effect of chronic inflammation and immunological changes (5, 6), as well as by directly upregulating hepatic lipogenesis (4). It is currently unknown whether liver dysfunction contributes to, or results from T2D development, or whether there is a genuine bidirectional relationship, where insulin resistance facilitates fat accumulation in the liver, which in turn leads to hepatic insulin resistance and increased fasting glucose (7-9). As most evidence to date is from observational studies, it is possible that the association between liver dysfunction and T2D reflects underlying common causes, such as obesity or lifestyle characteristics.

Plasma concentration of liver enzymes (i.e. alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), and gamma-glutamyl transferase (GGT)) are routinely measured clinical markers that represent different dimensions of liver dysfunction. ALT, located in the cytosol, and AST, located in the mitochondria, are released from damaged hepatic cells into the blood after hepatocellular injury or death. ALT and AST are potentially useful surrogates for alcohol-induced liver disease and non-alcoholic fatty liver disease (NAFLD), defined as hepatic steatosis in the absence of excessive alcohol consumption. ALP is present in the ducts of the liver and GGT is located on liver cell membranes. The combined elevation of ALP and GGT can indicate obstructive or cholestatic liver disease, where bile is not properly transported from the liver because of obstruction of the bile duct. GGT is also an indicator of alcohol use (10).

Classical observational studies show that plasma concentrations of these enzymes, even within the normal range, are positively associated with type 2 diabetes (1-3). Mendelian randomization (MR), where genetic variants that are strongly associated with a risk factor of interest are used to test its causal effect on an outcome, can help to distinguish causal effects from associations due to confounding or reverse causality (11, 12). Previous MR studies do not support a link between circulating GGT or ALP on T2D risk or glycemic status in Europeans (13, 14) or of ALT on T2D risk or glycemic status in Chinese (15). In contrast, MR studies reported some evidence of a positive relation of circulating GGT on T2D risk in South Koreans (16) and on insulinaemia in Europeans (17) and of circulating ALT on T2D risk in Europeans (14). To the best of our knowledge, no MR study has investigated the effect of predisposition to T2D or to insulin resistance on liver function markers.

We have used the largest available datasets to interrogate the potential effect of liver dysfunction, proxied by multiple biomarkers (ALT, AST, ALP, and GGT), on T2D risk (64,094 T2D cases and 607,012 controls), as well as on related outcomes (fasting glucose, insulin and lipids). In addition, we have investigated whether predisposition to T2D and insulin resistance affect circulating liver function markers (ALT, AST, ALP, and GGT).

RESEARCH DESIGN AND METHODS

Study design

We explored the relationship of four liver function markers (plasma concentration of ALT, AST, ALP, and GGT) with T2D (primary outcome) and with six related metabolic traits (secondary outcomes) reflecting hyperglycemia, assessed by fasting glucose, insulin resistance, assessed by fasting insulin, and dyslipidemia, assessed by LDLc, HDLc, total cholesterol, triglycerides, using two approaches – multivariable regression and MR. We also used MR to investigate whether predisposition to T2D and to insulin resistance are likely to have an impact on circulating ALT, AST, ALP, and GGT. The hypotheses, study design and data sources used are detailed in **Figure 1**.

Data sources

Participant-level data

The **UCL-LSHTM-Edinburgh-Bristol (UCLEB) consortium** consists of 12 prospective observational studies comprising over 30,000 participants (18). For the present study, data from up to seven UCLEB studies were included in multivariable and MR analyses: the British Regional Heart Study (BRHS) (19), British Women's Heart and Health Study (BWHHS) (20), Caerphilly Prospective Study (CaPS) (21), Edinburgh Artery Study (EAS) (22), English Longitudinal Study of Ageing (ELSA) (23), Whitehall II study (WHII) (24), MRC National Survey of Health and Development (NSHD) (25). Full details of the studies included in the UCLEB consortium have been described previously (18). For the multivariable analyses of liver marker - T2D associations, we used data from up to 6,593 individuals (728 T2D cases) from up to five UCLEB studies (BRHS, BWHHS, EAS, CaPS, and NSHD). For

the MR analyses of the effect of liver function on T2D risk, we used up to 11,790 individuals (1,202 T2D cases) from up to seven UCLEB studies (BRHS, BWHHS, CaPS, EAS, ELSA, NSHD, and WHII) where information on genotypes and a liver function marker/s, and/or information on genotypes and outcome measure/s were available.

The **Fenland study** is a UK population-based study based in the East Cambridgeshire and Fenland areas and has been described in detail elsewhere (26). Data from Fenland study participants were included in both the multivariable (up to 9,968 individuals) and the MR analyses of the liver marker-lipid outcome associations (up to 9,982 individuals) except for AST.

Full details of the exposure, outcome and confounder variables available in each UCLEB and Fenland study are given in **Supplementary table 1** and participant characteristics, in **Supplementary tables 2a and 2b**.

Summary-level data

For the multivariable analysis, we pooled our individual participant-level results (from UCLEB and Fenland studies) with those from two published meta-analysis of the association of liver function markers and T2D risk: Kunutsor *et al.*(2) for ALT and AST (up to 60,359 participants including 3,890 incident T2D cases) and Fraser *et al*.(27) for GGT (up to 63,285 including 2,805 incident T2D cases).

For the MR analysis, we also used publicly available summary level data from genome-wide association studies (GWAS) for the association of single nucleotide polymorphisms (SNPs) with exposures and outcomes from the relevant GWAS studies/consortia. Summary statistics for the association between SNP and liver function markers were extracted from Chambers *et al*. (28), including 61,089 individuals with information of ALT, ALP, AST, and GGT plasma concentration. Summary data for the association between SNPs and T2D was

extracted from a consortium (29) including 62,892 cases and 596,424 controls, mostly of European descent. In cases where a liver function SNP could not be found in this latest T2D GWAS, data was extracted from a previous T2D GWAS from the DIAGRAM consortium including 34,840 cases and 114,981 controls (30). Summary statistics for the association of SNPs with fasting glucose and with fasting insulin were obtained from **MAGIC** (the Meta-Analyses of Glucose and Insulin-related traits Consortium) consortium (31, 32), which included up to 133,010 and 108,557 nondiabetic participants of European ancestry, respectively. Summary data for the association of SNPs with LDLc, HDLc, total cholesterol, and triglycerides were extracted from the **GLGC** (Global Lipids Genetics Consortium) consortium (33), including 188,577 individuals mostly of European ancestry. We excluded Fenland participants who had been included in the meta-analyses of the GLGC consortium to avoid having duplicated information from these participants.

Definition of diabetes

In UCLEB studies, T2D definition varied by study and included self-report, medical history review, use of glucose lowering medication, or having a fasting glucose value of ≥ 7 mmol/L(18). For the MR analysis, we used both prevalent and incident cases of type 2 diabetes to maximise power, as the prevalent diabetes cases cannot influence genetic variation, which is fixed at conception.

In the GWAS, criteria for defining T2D differed across studies and included previous diagnosis, fasting glucose ≥ 7 mmol/L, treatment with glucose-lowering medication, or self-reported T2D status (29, 30, 34).

Genotyping and quality control

All studies in UCLEB consortium were genotyped using the Illumina CardioMetabochip (Illumina, San Diego, CA, USA). Details on the genotyping and imputation quality control criteria used have been previously published (18, 35). Genotyping and imputation of missing genotypes in the published GWAS studies used here are described in the original publications (28, 30-33, 36, 37).

Statistical analyses

All continuous variables in the UCLEB and Fenland studies that were not normally distributed were natural log transformed. All continuously measured traits were standardised within each study to allow comparison between studies. In any published GWAS that did not report effects in standard deviation units, effects were standardised based on the GWAS reported standard deviation or, where this was not available, the median standard deviation across UCLEB studies.

For analyses involving fasting glucose and fasting insulin we removed individuals with type 2 diabetes (defined as a clinical diagnosis, fasting glucose values ≥ 7.0 mmol/L or individuals who were taking glucose lowering drugs). We also excluded individuals on lipid lowering medication from the multivariable and MR.

Analyses were performed using Stata/SE version 14.0 (StataCorp, Brownsville, TX) and R version 3.4.4 (R Foundation for Statistical Computing, Vienna, Austria).

Multivariable analysis

We estimated the association between each liver function marker and T2D in each UCLEB study using logistic regression. For continuous outcomes we used linear regression models. We adjusted the logistic/linear regression models for age, sex (if relevant), and for as many potential confounders available in each study from the following: BMI, waist

circumference, alcohol consumption, smoking, and social class. We then used DerSimonian and Laird random-effects model meta-analysis to combine these estimates with results from the published meta-analyses of Kunutsor *et al* (2) and Fraser *et al* (27) for the associations between ALT, AST and GGT with T2D, excluding any UCLEB studies that had contributed to the published meta-analyses. In the previously published meta-analyses, all (27) or some (2) of the participating studies expressed T2D results as hazard ratios. Given the overall proportion of T2D cases was lower than 10%, hazard ratios and odds ratios were assumed to approximate to the same measure of T2D relative risk in our meta-analysis.

Mendelian randomization analysis

For the MR analysis, we used multiple genetic variants robustly associated with circulating ALT, AST, ALP, or GGT as genetic instruments for each liver function marker to investigate their effect on T2D and other outcomes. We also applied MR analysis to assess the effect of predisposition to T2D and fasting insulin (a marker of insulin resistance) on circulating ALT, AST, ALP, or GGT by using genetic variants robustly associated with T2D risk and fasting insulin. We used a "two-sample" analysis strategy in which the genetic variant-exposure and genetic variant-outcome associations are estimated from different data sources with comparable populations (38).

Selection of genetic instruments for liver function markers

Genetic instruments for each liver function marker were defined as independent SNPs ($R^2 < 0.3$) associated with each liver function marker at genome-wide levels of significance ($P < 5 \times 10^{-8}$) from GWAS in the NIHR GWAS catalogue (available at www.ebi.ac.uk/gwas/home, last accessed on July 2017). At the time the study was conducted, there was limited GWAS data available for AST. Therefore, we conducted a genome-wide association analysis in up to 6,647 individuals from five UCLEB studies (BRHS, BWHHS, CaPS, EAS, and ET2DS) to identify novel variants for AST, as this biomarker had only been

assessed in one previous GWAS with fewer than 1,000 individuals (37). We selected GWAS studies, which had been primarily conducted in populations of European ancestry. In total, we selected 4, 3 (including the two novel SNPs identified from our own GWAS), 15, and 26 independent SNPs associated with ALT, AST, ALP, and GGT, respectively (**Supplementary table 3**). UCLEB data were excluded from the MR analysis of AST (as an exposure) to avoid any bias from Winner's curse (39).

Selection of genetic instrument for type 2 diabetes and fasting insulin

We used multiple independent SNPs strongly associated ($P < 5 \times 10^{-8}$) with T2D risk (n = 139 SNPs) (29) and with fasting insulin (n = 14 SNPs) (32). T2D SNPs were identified as reported by the original GWAS publication (29). Fasting insulin SNPs were selected from data published by Scott et al (32) using the R package 'TwoSampleMR' excluding any correlated SNPs ($R^2 > 0.001$) (40).

Main analysis

In the main Mendelian randomization analyses, we used the conventional inverse variance weighted (IVW) estimator. The IVW method consists of a weighted regression of the SNP-outcome regression coefficients on the SNP-exposure regression coefficients constraining the intercept to be zero. IVW weights are the inverse of the variance of the SNP-outcome regression coefficients. For a dichotomous outcome such as a T2D status, the regression coefficient of the SNP-outcome association is a log odds ratio from a logistic regression model. The resulting regression coefficient from the IVW regression represents an increase/decrease in the outcome per unit increase in the exposure.

Sensitivity analyses

We performed several sensitivity analyses to test whether the MR IVW estimates are likely to be biased by unbalanced horizontal pleiotropic effects (i.e. due to genetic variants that affect the outcome independently of the exposure of interest). We used MR-Egger regression method (41) and the weighted median estimator (42), both of which are more robust to pleiotropic genetic variants, to test the extent to which any unbalanced pleiotropy may have biased the IVW result. The MR-Egger method is similar to the IVW except that the model allows the intercept to vary. The intercept of the MR-Egger regression will reflect the average pleiotropic effect across genetic variants and the slope coefficient will provide an estimate of the causal effect provided that the InSIDE (Instrument Strength Independent of Direct Effect) assumption holds, which requires that there is no correlation between SNP-exposure association and any direct (pleiotropic) effects of SNP on outcome (41). In contrast, the weighted median estimator gives consistent estimates even if up to 50% of weight in the analysis comes from invalid genetic instruments (42). For the MR analysis of the effect of insulin resistance, proxied by fasting insulin, on liver markers, we have performed an additional sensitivity analysis, in which we used multivariable MR (43) to adjust results by BMI (32). In all analyses with GWAS data, we have excluded SNPs with C/G or A/T genotypes and minor allele frequency > 0.42 due to ambiguity problems they could introduce when harmonising SNP-exposure and SNP-outcome datasets (44).

RESULTS

Genetic instruments for liver function markers, T2D risk and fasting insulin

Results for the association of genetic instruments with the respective liver function marker are given in **Table 1**. Most of the genetic instruments previously identified in GWAS consortia replicated (had consistent direction and magnitude) in UCLEB and Fenland, seven were not available in UCLEB but were replicated in Fenland, seven were null or in the opposite direction in UCLEB but consistent in Fenland, and one was null in Fenland (but consistent in UCLEB). In all cases, confidence intervals in UCLEB and Fenland included the point estimate from the published GWAS.

For the known variant for AST (rs17109512), only P-values were provided by the original GWAS and, therefore, it was not possible to compare effect estimates with that from the previous GWAS. The two novel variants for AST, identified in our GWAS conducted in five UCLEB studies, were replicated in independent data sources (**Table 1** and **Supplementary table 4**). When meta-analysing UCLEB and Fenland studies, 4, 13 and 22 variants were replicated from 4, 15 and 26 variants associated with ALT, ALP and GGT, and the remaining variants were directionally consistent with the previous reports (data only shown for UCLEB and Fenland separately).

Results for the association of the 139 T2D genetic instruments with T2D risk are given in **Supplementary table 5**. Results for the association of the 14 fasting insulin genetic instruments with fasting insulin (in SD log pmol/L) are given in **Supplementary table 6**.

Association of genetic variants related to liver function markers with potential confounders of the exposure-outcome association

Some individual SNPs used as instruments in the MR analysis were associated with potential confounders of the exposure-outcome association (e.g. age, BMI, waist circumference, waist-hip ratio, and other liver function markers). However, when combined into a single instrument using fixed-effect meta-analysis, there was no strong evidence that genetic instruments were associated with potential confounders, except for the ALT instrument that was also associated with AST and GGT, and the AST instrument that was associated with ALT (**Supplementary tables 7a and 7b)**.

Multivariable analysis between liver function markers and T2D and related continuous outcomes

Pooled results from multivariable analyses across the relevant UCLEB and Fenland studies and published meta-analyses are given in **Figure 2A**. Most liver function markers were positively associated with T2D risk. The OR for T2D was 1.27 [95% CI: 1.15, 1.40] for ALT, 1.06 [95% CI: 0.97, 1.17] for AST, 1.25 [95% CI: 1.10, 1.43] for ALP and 1.61 [95% CI: 1.15, 2.26] for GGT (per standard unit increase in the liver function marker). Heterogeneity across studies was low for ALT and ALP ($I^2 = 0\%$) but high for AST ($I^2 = 81\%$) and GGT ($I^2 = 93\%$) (**Figure 2A**).

In meta-analyses of the continuous outcomes, ALT was positively associated with insulin and triglycerides. AST was positively associated with insulin and HDLc. ALP was negatively related with HDLc. GGT was positively associated with all continuous outcomes (**Supplementary table 8**).

Effect of liver function markers on T2D and related continuous outcomes using MR

Pooled results from the MR across UCLEB, Fenland and GWAS studies are given in **Figure 2A** and **Supplementary table 8**. In the main MR analysis (IVW), the OR for T2D was

1.45 [95% CI: 1.10, 1.92] for ALT, 1.25 [95% CI: 1.14, 1.38] for AST, 0.91 [95% CI: 0.86, 0.97] for ALP and 0.92 [95% CI: 0.80, 1.06] for GGT (per each standard unit increase in the liver function marker) (**Figure 2A**). The other MR methods (MR-Egger and weighted median) used as sensitivity analyses were consistent with IVW estimates for ALT, AST and ALP. The inverse point estimate for the IVW association between GGT and T2D changed in direction when using the MR-Egger method. There was no clear evidence of unbalanced horizontal pleiotropy in any liver function marker – T2D associations based on the intercept for MR-Egger method (all P-values ≥0.38). Heterogeneity between UCLEB and GWAS estimates was low for all liver markers. For AST, only GWAS data were available for analysis (**Figure 2A**).

In meta-analyses of the continuous outcomes, there was some evidence across different MR methods that genetically predicted AST and ALP were negatively related to HDLc, LDLc and total cholesterol (**Supplementary table 8**). There was some evidence of unbalanced horizontal pleiotropy of ALT instruments in relation to HDLc, LDLc, total cholesterol and triglycerides (**Supplementary table 9**).

Effect of T2D and insulin resistance on liver function markers using MR

Overall, findings from MR analysis did not consistently support a reverse causal effect of T2D on any of the liver function markers assessed (ALT, AST, ALP, or GGT). In the main MR analysis (IVW), higher predisposition to T2D (each increase in 1 log odds) was related to an increase of 0.06 SD units of ALT [95% CI: 0.02, 0.09], but not of AST [0.01; 95% CI: -0.04, 0.07], ALP [-0.04; 95% CI: -0.11, 0.02] or GGT [0.03; 95% CI: -0.01, 0.06]. Results for ALT were substantially attenuated when using MR methods that are more robust to pleiotropic variants: 0.01 [95% CI: -0.06, 0.08] for MR-Egger and 0.03 [95% CI: -0.02, 0.07] for the weighted median estimator (**Figure 2B**).

On the other hand, there was evidence from the main MR findings that higher insulin resistance, proxied by fasting insulin, increases circulating ALT. Results were consistent with

the weighted median and BMI-adjusted estimates. MR-Egger point estimates substantially differed, but 95% CI were very wide. Results for the other liver markers were less consistent across different MR methods and 95% CI were wide (**Table 2**).

DISCUSSION

More than a century ago a link between liver disease and diabetes was described (45). Since then, multiple observational studies have repeatedly reported that liver dysfunction and T2D are associated (1-3), as broadly replicated by our multivariable analysis using clinical biomarkers (i.e. circulating ALT, AST, ALP, and GGT) as proxies of liver dysfunction. Nevertheless, it is still unclear whether this association reflects causation, and if so, whether liver dysfunction represents a cause or a consequence of T2D.

Our study expands on previous MR analyses investigating the effect of liver dysfunction on T2D risk by including a more comprehensive set of liver function markers routinely used in clinical practise in the largest available datasets and by applying bi-directional MR to investigate whether predisposition to T2D and to insulin resistance might instead lead to liver dysfunction.

Our findings from the MR analyses show evidence that genetic predisposition to higher circulating ALT and AST is related to higher risk of T2D. No strong evidence of a causal effect of genetically predicted GGT on T2D, and evidence of a modest negative effect of genetically predicted ALP on T2D were found. Genetic predisposition to T2D did not appear to influence blood concentration of any of the studied liver function markers (ALT, AST, ALP, and GGT), whereas genetic predisposition to insulin resistance, proxied by fasting insulin, seems to increase ALT (effects on other liver markers are uncertain).

Our results are broadly consistent with two previous MR studies, of largely European origin participants (using a study sample that partially overlaps with ours), which reported strong evidence for a positive effect of ALT on T2D (14), but not for ALP (14) or GGT (13, 14). Results for ALP (14) were directionally consistent with our findings, but we were better powered to test for the association between ALP and T2D risk given the substantially larger number of T2D cases and controls included in our analyses. In non-European populations, MR

studies suggest that ALT does not relate to T2D risk in Chinese adults (15), but that higher GGT increases T2D risk in Koreans (16). However, the latter result might be explained by statistical overfitting since instruments were selected from a GWAS that included the population used in the MR analyses (~ 20% of the GWAS discovery sample).

The different liver biomarkers reflect different aspects of liver dysfunction. High circulating ALT and AST are widely used proxies of NAFLD, while high circulating ALP and GGT (in combination) are more related to obstructive or cholestatic liver disease. The positive association between genetically predicted liver function markers and T2D in MR analyses was robust for ALT and AST, but not apparent for ALP or GGT, which suggests that NAFLD might be the primary type of liver dysfunction driving these associations.

In agreement with that, a genetic variant (rs738409) in perfect linkage disequilibrium (R2 = 1.0 for 1000 Genomes European population - GRCh37) with one of our instruments for AST (rs738408) has been previously reported to be associated with computed tomography (CT) measured hepatic steatosis (46, 47). NAFLD is the most common cause of chronic liver disease in Western countries owing to the rapid increase in obesity prevalence (48, 49). NAFLD affects 70% of patients with type 2 diabetes (T2D) in contrast to around 20% of the general population (48, 49).

To our knowledge previous MR studies have not examined the causal effect of predisposition to T2D or insulin resistance on liver dysfunction. We found supportive evidence that insulin resistance increases circulating ALT. Our combined findings that insulin resistance (but not T2D) may cause elevated ALT (marker of NAFLD), and that ALT and AST are in turn related to increased T2D risk (but not insulin resistance) are consistent with the twin-cycle hypothesis (8), which postulates that there is a vicious cycle between hepatic insulin resistance and β cell dysfunction. According to the twin-cycle hypothesis, elevated insulin (due to insulin resistance) stimulates *de novo* lipogenesis in the liver, which promotes hepatic insulin resistance leading to overstimulation of hepatic gluconeogenesis and increased fasting glucose.

The resulting increased output of triglycerides and glucose by the liver to the circulation would impair beta cell function, eventually leading to type 2 diabetes (9).

On the other hand, it is worth emphasising that the relation between ALT/AST and T2D risk might be explained by factors other than NAFLD since circulating ALT/AST are not specific markers of NAFLD and can also increase in response to liver injury from other causes, such as drug toxicity, infection and alcohol consumption (50, 51). In addition, there is some evidence that hepatic triglyceride accumulation by itself may not necessarily cause metabolic changes increasing the risk of cardiometabolic complications (52) and that there might be multiple T2D subtypes that differ in terms of disease presentation and responsiveness to interventions (53), which we were unable to tease out due to the predominant use of summary level data. Finally, given the key roles of ALT and AST in the intermediary metabolism of glucose and amino acids, we cannot fully discard that they might be directly implicated in T2D development.

Although MR can substantially improve causal inference in Epidemiological studies (54), it is important to note that the reliability of MR findings depend on the three core assumptions of instrumental variable (IV) analysis, which require that genetic instruments are strongly associated with the exposure (**IV1**), are not related to exposure-outcome confounders (**IV2**) and only influence the outcome through the exposure (**IV3**).

To avoid violations of IV1, we have selected genetic variants strongly associated with liver function markers ($P < 5*10^{-8}$), that broadly replicated in independent datasets. It should also be noted that our genetic instruments were selected to be strongly associated with each liver function marker in the largest available GWAS ($P < 5*10^{-8}$) and that some variants relevant to specific forms of liver dysfunction might not have met our inclusion criteria, as is the case of variants nearby *TM6SF2*, previously reported to be associated with NAFLD in a GWAS (55) and with alcohol-related cirrhosis in an exome-wide association study (56). To minimise the risk of population stratification, which could violate IV2, we mostly restricted

our analyses to individuals of European ancestry. IV3 could be violated in the presence of horizontal pleiotropy. We have attempted to address that by examining the relation of the genetic variants with several social, behavioural and metabolic phenotypes and by using methods that are more robust to violations of this assumption in sensitivity analyses (i.e. MR-Egger and weighted median estimator). Overall, there was no strong suggestion that horizontal pleiotropy could have biased our results. Importantly, our liver function instruments were not associated with general adiposity measures such as BMI, and therefore it is unlikely that these associations are driven by general adiposity. However, it is important to note that we cannot fully discard that our genetic instruments may be associated with exposure-outcome confounders that we have not tested for and that the sensitivity analyses used (i.e. MR-Egger and weighted median) are of limited use for exposures instrumented by few genetic variants (as in the analyses of ALT and AST as exposures).

Two-sample MR makes the additional assumption that data from two independent (but comparable) populations are used. In our study, there was some overlap between participants used to estimate SNP-exposure and SNP-outcome associations. However, this is unlikely to bias the study results since the overlap is very low in proportion to the overall sample size (< 7% of T2D cases in the main analysis) (57).

Finally, effect estimates for the relation of each liver function marker on T2D risk should be interpreted with caution given these biomarkers are unlikely to be the causal factors for T2D risk, but rather proxies of liver function.

In conclusion, MR findings indicate that increased circulating ALT and AST is related with higher T2D risk, while increased circulating ALP is associated with lower T2D risk. In addition, higher fasting insulin (but not predisposition to T2D) is related to higher circulating ALT. Since circulating ALT and AST are markers of NAFLD, these findings provide some support for insulin resistance resulting in NAFLD, which in turn increases T2D risk.

Funding sources

This work was supported by funds from the UK Medical Research Council (Grant number: G0801456) and a British Heart Foundation grant (AA/18/7/34219). MCB, TRG and DAL work in a unit that receives funding from the UK Medical Research Council (Grant number: MC_UU_00011/1-6 and MC_UU_00011/1-4). MCB is supported by a Skills Development Fellowship from the UK Medical Research Council (Grant number MR/P014054/1). XZ is supported by a scholarship from the China Scholarship Council (Grant number: 201406220101). ADH is an NIHR Senior Investigator. AW and DK are funded by the UK Medical Research Council [MC UU 12019/1]. DAL is a UK National Institute of Health Research Senior Investigator (Grant number: NF-SI-0611-10196).

The **UCLEB** (UCL-LSHTM-Edinburgh-Bristol) consortium is supported by BHF Programme Grant RG/10/12/28456 and the UCL Hospitals NIHR Biomedical Research Centre.

BRHS is supported by British Heart Foundation grants (RG/08/013/25942, RG/13/16/30528). The British Heart Foundation had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. The authors acknowledge the British Regional Heart Study team for data collection.

BWHHS is supported by the British Heart Foundation (PG/13/66/30442). Data on mortality and cancer events were routinely provided from NHS Digital to the BWHHS under data sharing agreement MR104a- Regional Heart Study (Female Cohort). British Women's Heart and Health Study data are available to *bona fide* researchers for research purposes. Please refer to the BWHHS data sharing policy at www.ucl.ac.uk/british-womens-heart-health-study.

CaPS was funded by the Medical Research Council and undertaken by the former MRC Epidemiology Unit (South Wales). The CaPS DNA bank was established with funding from a MRC project grant. The CaPS data archive is maintained by the University of Bristol. MRC Integrative Epidemiology Unit, Bristol is supported by MRC grants (MR_UU_12013/1, MR_UU_12013/5 and MR_UU_12013/8).

The **EAS** is supported by the British Heart Foundation and by the Chief Scientist Office of Scotland.

ELSA is supported by the National Institute on Aging (NIA/NIH) USA (grant number 5 R01 AG017644-16) and a consortium of the United Kingdom government departments coordinated by the Economic and Social Research Council (ESRC).

The **WHII study** is supported by grants from the Medical Research Council (K013351), British Heart Foundation (RG/07/008/23674), Stroke Association, the US National Heart Lung and Blood Institute (5RO1 HL036310), the US National Institute on Aging (5RO1AG13196) the US Agency for Healthcare Research and Quality (HS06516); and the John D. and Catherine T. MacArthur Foundation Research Networks on Successful Midlife Development and Socio-economic Status and Health.

NSHD is funded by the UK Medical Research Council [MC UU 12019/1].

The **Fenland Study** is funded by the Wellcome Trust and the Medical Research Council (Grant number: MC_U106179471 and MC_UU_12015/1).

The funders had no role in the design, analyses, interpretation of results or writing of the paper. The views expressed in this paper are those of the authors and not necessarily any of the funders.

Acknowledgments

The **UCLEB** (UCL-LSHTM-Edinburgh-Bristol) consortium consists of 12 studies: Northwick Park Heart Study II (NPHS II), British Regional Heart Study (BRHS), Whitehall II Study (WHII), English Longitudinal Study of Ageing (ELSA), Medical Research Council National Survey of Health and Development (NSHD), 1958 Birth cohort (1958BC), Caerphilly prospective study (CaPS), British Women's Heart and Health Study (BWHHS), Edinburgh Artery Study (EAS), Edinburgh Heart Disease Prevention Study (EHDPS), Edinburgh Type 2 Diabetes Study (ET2DS) and Asymptomatic Atherosclerosis Aspirin Trial (AAAT). We thank all participants and staff who have contributed to those studies that contributed to the present study (BRHS, BWHHS, CaPS, EAS, ELSA, ET2DS, NSHD, WHII, Fenland study, and NFBC1966). We thank Naveed Sattar for the insightful comments on the interpretation of findings.

Data and Resource Availability

Data on type 2 diabetes have been contributed by the T2D GWAS published by Xue et al (2018) and have been downloaded from http://cnsgenomics.com/data.html. Data on glycaemic traits have been contributed by **MAGIC** investigators and have been downloaded from www.magicinvestigators.org. Data on lipid traits have been contributed by **Global Lipids Genetics Consortium** and have been downloaded from http://csg.sph.umich.edu/abecasis/public/lipids2013/. Data on anthropometric traits have been

contributed by Genetic Investigation of ANthropometric Traits (**GIANT**) consortium and have been downloaded from http://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files. UCLEB data is available upon request.

Author contributions

DAL designed the study and initial analysis plan, with further input from NMGDS, MCB, and TG. DAL, JE, AW, DK, TS, XZ, CL, JL, JCC, WZ, and TRG collected and/or managed data from one or more of the contributing studies, NMGDS and MCB undertook statistical analyses with supervision from DAL and TG. NMGDS, DAL and MCB wrote the first draft of the paper with all other co-authors contributing to revisions.

Conflict of Interest statement

DAL has received support from Medtronic and Roche Diagnostics for biomarker research unrelated to this study. TRG receives support from GlaxoSmithKline, Biogen and Sanofi for research unrelated to this study. All other authors report no competing interests.

Guarantor Statement

NMGDS and MCB are the guarantors of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Ballestri S, Zona S, Targher G, Romagnoli D, Baldelli E, Nascimbeni F, et al. Nonalcoholic fatty liver disease is associated with an almost twofold increased risk of incident type 2 diabetes and metabolic syndrome. Evidence from a systematic review and meta-analysis. J Gastroenterol Hepatol. 2016;31(5):936-44.
2. Kunutsor SK, Apekey TA, Walley J. Liver aminotransferases and risk of incident type 2 diabetes: a systematic review and meta-analysis. Am J Epidemiol. 2013;178(2):159-71.
3. Wannamethee SG, Shaper AG, Lennon L, Whincup PH. Hepatic enzymes, the metabolic syndrome, and the risk of type 2 diabetes in older men. Diabetes Care. 2005;28(12):2913-8.
4. Smith BW, Adams LA. Nonalcoholic fatty liver disease and diabetes mellitus: pathogenesis and treatment. Nat Rev Endocrinol. 2011;7(8):456-65.
5. Day CP, James OF. Steatohepatitis: a tale of two "hits"? Gastroenterology. 1998;114(4):842-5.
6. Utzschneider KM, Kahn SE. Review: The role of insulin resistance in nonalcoholic fatty liver disease. J Clin Endocrinol Metab. 2006;91(12):4753-61.
7. Loria P, Lonardo A, Anania F. Liver and diabetes. A vicious circle. Hepatol Res. 2013;43(1):51-64.
8. Taylor R. Pathogenesis of type 2 diabetes: tracing the reverse route from cure to cause. Diabetologia. 2008;51(10):1781-9.
9. Taylor R, Al-Mrabeh A, Zhyzhneuskaya S, Peters C, Barnes AC, Aribisala BS, et al. Remission of Human Type 2 Diabetes Requires Decrease in Liver and Pancreas Fat Content but Is Dependent upon Capacity for beta Cell Recovery. Cell Metab. 2018;28(4):667.
10. Green RM, Flamm S. AGA technical review on the evaluation of liver chemistry tests. Gastroenterology. 2002;123(4):1367-84.
11. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133-63.
12. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014;23(R1):R89-98.
13. Noordam R, Smit RA, Postmus I, Trompet S, van Heemst D. Assessment of causality between serum gamma-glutamyltransferase and type 2 diabetes mellitus using publicly available data: a Mendelian randomization study. Int J Epidemiol. 2016.
14. Liu J, Au Yeung SL, Lin SL, Leung GM, Schooling CM. Liver Enzymes and Risk of Ischemic Heart Disease and Type 2 Diabetes Mellitus: A Mendelian Randomization Study. Sci Rep. 2016;6:38813.
15. Xu L, Qiang Jiang C, Hing Lam T, Sen Zhang W, Zhu F, Li Jin Y, et al. Mendelian randomization estimates of alanine aminotransferase with cardiovascular disease: Guangzhou Biobank Cohort study. Hum Mol Genet. 2016.
16. Lee YS, Cho Y, Burgess S, Davey Smith G, Relton CL, Shin SY, et al. Serum gamma-glutamyl transferase and risk of type 2 diabetes in the general Korean population: a Mendelian randomization study. Hum Mol Genet. 2016.
17. Conen D, Vollenweider P, Rousson V, Marques-Vidal P, Paccaud F, Waeber G, et al. Use of a Mendelian randomization approach to assess the causal relation of gamma-Glutamyltransferase with blood pressure and serum insulin levels. Am J Epidemiol. 2010;172(12):1431-41.
18. Shah T, Engmann J, Dale C, Shah S, White J, Giambartolomei C, et al. Population genomics of cardiometabolic traits: design of the University College London-London School of Hygiene and Tropical Medicine-Edinburgh-Bristol (UCLEB) Consortium. PLoS One. 2013;8(8):e71345.
19. Shaper AG, Pocock SJ, Walker M, Cohen NM, Wale CJ, Thomson AG. British Regional Heart Study: cardiovascular risk factors in middle-aged men in 24 towns. Br Med J (Clin Res Ed). 1981;283(6285):179-86.
20. Lawlor DA, Bedford C, Taylor M, Ebrahim S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. J Epidemiol Community Health. 2003;57(2):134-40.
21. Bainton D, Miller NE, Bolton CH, Yarnell JW, Sweetnam PM, Baker IA, et al. Plasma triglyceride and high density lipoprotein cholesterol as predictors of ischaemic heart disease in British men. The Caerphilly and Speedwell Collaborative Heart Disease Studies. Br Heart J. 1992;68(1):60-6.
22. Fowkes FG, Housley E, Cawood EH, Macintyre CC, Ruckley CV, Prescott RJ. Edinburgh Artery Study: prevalence of asymptomatic and symptomatic peripheral arterial disease in the general population. Int J Epidemiol. 1991;20(2):384-92.
23. Marmot M, Banks J, Blundell R, Lessof C, Nazroo J. Health, Wealth and Lifestyles of the Older Population in England: The 2002 English Longitudinal Study of Ageing. London: Institute for Fiscal Studies; 2003.
24. Marmot MG, Smith GD, Stansfeld S, Patel C, North F, Head J, et al. Health inequalities among British civil servants: the Whitehall II study. Lancet. 1991;337(8754):1387-93.
25. Wadsworth M, Kuh D, Richards M, Hardy R. Cohort Profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development). Int J Epidemiol. 2006;35(1):49-54.
26. Burgoine T, Forouhi NG, Griffin SJ, Wareham NJ, Monsivais P. Associations between exposure to takeaway food outlets, takeaway food consumption, and body weight in Cambridgeshire, UK: population based, cross sectional study. BMJ. 2014;348:g1464.

27. Fraser A, Harris R, Sattar N, Ebrahim S, Davey Smith G, Lawlor DA. Alanine aminotransferase, gamma-glutamyltransferase, and incident diabetes: the British Women's Heart and Health Study and meta-analysis. Diabetes Care. 2009;32(4):741-50.
28. Chambers JC, Zhang W, Sehmi J, Li X, Wass MN, Van der Harst P, et al. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet. 2011;43(11):1131-8.
29. Xue A, Wu Y, Zhu Z, Zhang F, Kemper KE, Zheng Z, et al. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. Nat Commun. 2018;9(1):2941.
30. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet. 2012;44(9):981-90.
31. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet. 2010;42(2):105-16.
32. Scott RA, Lagou V, Welch RP, Wheeler E, Montasser ME, Luan J, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat Genet. 2012;44(9):991-1005.
33. Global Lipids Genetics C, Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013;45(11):1274-83.
34. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet. 2010;42(7):579-89.
35. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518(7538):197-206.
36. Yuan X, Waterworth D, Perry JR, Lim N, Song K, Chambers JC, et al. Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. Am J Hum Genet. 2008;83(4):520-8.
37. Shen H, Damcott C, Shuldiner SR, Chai S, Yang R, Hu H, et al. Genome-wide association study identifies genetic variants in GOT1 determining serum aspartate aminotransferase levels. J Hum Genet. 2011;56(11):801-5.
38. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. Int J Epidemiol. 2016;45(3):908-15.
39. Kraft P. Curses--winner's and otherwise--in genetic epidemiology. Epidemiology. 2008;19(5):649-51; discussion 57-8.
40. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife. 2018;7.
41. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44(2):512-25.
42. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016;40(4):304-14.
43. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int J Epidemiol. 2018.
44. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. Int J Epidemiol. 2016;45(6):1717-26.
45. Naunyn B. Glykosurie und Diabetes durch experimentelle Insulte und Krankheiten der Leber. In: Naunyn B, editor. Der Diabetes Mellitus. Vienna: A. Holder; 1898. p. 38–49.
46. Speliotes EK, Yerges-Armstrong LM, Wu J, Hernaez R, Kim LJ, Palmer CD, et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. PLoS Genet. 2011;7(3):e1001324.
47. Lauridsen BK, Stender S, Kristensen TS, Kofoed KF, Kober L, Nordestgaard BG, et al. Liver fat content, non-alcoholic fatty liver disease, and ischaemic heart disease: Mendelian randomization and meta-analysis of 279 013 individuals. Eur Heart J. 2018;39(5):385-93.
48. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, et al. The diagnosis and management of non-alcoholic fatty liver disease: Practice guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Am J Gastroenterol. 2012;107(6):811-26.
49. Sattar N, Forrest E, Preiss D. Non-alcoholic fatty liver disease. BMJ. 2014;349:g4596.
50. Dufour DR, Lott JA, Nolte FS, Gretch DR, Koff RS, Seeff LB. Diagnosis and monitoring of hepatic injury. II. Recommendations for use of laboratory tests in screening, diagnosis, and monitoring. Clin Chem. 2000;46(12):2050-68.
51. Dufour DR, Lott JA, Nolte FS, Gretch DR, Koff RS, Seeff LB. Diagnosis and monitoring of hepatic injury. I. Performance characteristics of laboratory tests. Clin Chem. 2000;46(12):2027-49.

52. Sliz E, Sebert S, Wurtz P, Kangas AJ, Soininen P, Lehtimaki T, et al. NAFLD risk alleles in PNPLA3, TM6SF2, GCKR and LYPLAL1 show divergent metabolic effects. Hum Mol Genet. 2018;27(12):2214-23.
53. Ahlqvist E, Storm P, Karajamaki A, Martinell M, Dorkhan M, Carlsson A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. Lancet Diabetes Endocrinol. 2018;6(5):361-9.
54. Smith GD, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. PLoS Med. 2007;4(12):e352.
55. Kozlitina J, Smagris E, Stender S, Nordestgaard BG, Zhou HH, Tybjaerg-Hansen A, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. Nat Genet. 2014;46(4):352-6.
56. Buch S, Stickel F, Trepo E, Way M, Herrmann A, Nischalke HD, et al. A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. Nat Genet. 2015;47(12):1443-8.
57. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genet Epidemiol. 2016;40(7):597-608.

Table 1. The associations of individual SNPs (used as genetic instruments in Mendelian randomization analyses) with the relevant liver function markers in UCLEB, Fenland and GWAS studies

SNP	Liver function marker	Locus*	Effect size in SD units in UCLEB studies† (95% CI)	Effect size in SD units in Fenland study† (95% CI)	Effect size in SD units from published GWAS‡ (95% CI)
rs6834314	ALT	*HSD17B13, MAPK10*	0.05 (0.01, 0.09)	0.05 (0.02, 0.07)	0.06 (0.04, 0.08)
rs11597390	ALT	*CPN1*	0.09 (0.05, 0.13)	0.03 (0.01, 0.06)	0.04 (0.03, 0.06)
rs2143571	ALT	*SAMM50*	0.14 (0.09, 0.19)	0.07 (0.04, 0.12)	0.09 (0.07, 0.11)
rs2954021	ALT	*TRIB1n*	0.03 (-0.01, 0.07)	0.02 (-0.00, 0.04)	0.04 (0.02, 0.05)
rs17109512	AST	*GOT1*	-0.04 (-0.18, 0.11)	N/A	0.03 (-0.05, 0.1)
rs738407	AST	*PNPLA3*	0.13 (0.08, 0.19)	N/A	0.09 (0.06, 0.12)
rs738408	AST	*PNPLA3*	0.12 (0.08, 0.17)	N/A	0.19 (0.16, 0.22)
rs1780324	ALP	*NBPF3-ALPL*	0.07 (0.02, 0.12)	0.08 (0.05, 0.10)	0.09 (0.07, 0.11)
rs16856332	ALP	*ABCB11*	-0.02 (-0.20, 0.15)	0.06 (-0.01, 0.13)	0.14 (0.09, 0.19)
rs9467160	ALP	*GPLD1*	0.06 (0.01, 0.12)	0.10 (0.07, 0.13)	0.1 (0.08, 0.13)
rs514708	ALP	*ABO*	0.06 (0.00, 0.12)	0.08 (0.05, 0.11)	0.1 (0.08, 0.13)
rs2236653	ALP	*ST3GAL4*	0.01 (-0.05, 0.07)	0.07 (0.04, 0.10)	0.06 (0.04, 0.08)
rs7186908	ALP	*HPR, PMFBP1*	-0.02 (-0.09, 0.05)	0.02 (-0.01, 0.05)	0.07 (0.05, 0.1)
rs7267979	ALP	*ABHD12, GINS1, PYGB*	0.03 (-0.02, 0.08)	0.07 (0.04, 0.10)	0.05 (0.03, 0.07)
rs281377	ALP	*ABO*	0.00 (-0.05, 0.06)	0.06 (0.04, 0.09)	0.07 (0.05, 0.09)
rs314253	ALP	*ASGR1o, DLG4n*	0.02 (-0.04, 0.07)	0.04 (0.01, 0.07)	0.08 (0.06, 0.1)
rs579459	ALP	*ABO*	0.08 (0.02, 0.14)	0.25 (0.22, 0.28)	0.31 (0.28, 0.34)
rs6984305	ALP	*PPP1R3B*	-0.03 (-0.11, 0.05)	0.08 (0.04, 0.12)	0.1 (0.07, 0.13)
rs7923609	ALP	*JMJD1Cnce, NRBF2e*	0.05 (0.00, 0.10)	0.09 (0.06, 0.11)	0.08 (0.06, 0.1)
rs174601	ALP	*C11orf10e, FADS1e, FADS2ne*	-0.00 (-0.06, 0.05)	0.06 (0.03, 0.08)	0.06 (0.04, 0.09)
rs2954021	ALP	*TRIB1*	0.07 (0.01, 0.12)	0.04 (0.01, 0.06)	0.05 (0.03, 0.07)
rs10819937	ALP	*ALDOBo, C9orf125n*	Not in UCLEB	0.06 (0.02, 0.09)	0.09 (0.06, 0.12)
rs1497406	GGT	*RSG1, EPHA2*	0.05 (0.01, 0.09)	0.07 (0.04, 0.09)	0.06 (0.04, 0.07)
rs12145922	GGT	*CCBL2, PKN2*	-0.02 (-0.05, 0.02)	0.01 (-0.02, 0.03)	0.04 (0.03, 0.06)

rs1335645	GGT	*CEPT1, DENND2D*	Not in UCLEB	0.03 (-0.01, 0.07)	0.07 (0.04, 0.09)
rs10908458	GGT	*DPM3, EFNA1, PKLR*	0.04 (0.01, 0.07)	0.05 (0.02, 0.07)	0.06 (0.04, 0.07)
rs13030978	GGT	*MYO1B, STAT4*	0.07 (0.02, 0.12)	0.03 (0.01, 0.06)	0.06 (0.04, 0.08)
rs2140773	GGT	*EFHD1, LOC100129166*	-0.01 (-0.04, 0.03)	0.02 (-0.01, 0.04)	0.05 (0.03, 0.06)
rs4547811	GGT	*ZNF827*	0.08 (0.03, 0.13)	0.09 (0.06, 0.12)	0.1 (0.08, 0.12)
rs4074793	GGT	*ITGA1*	0.14 (0.07, 0.20)	0.07 (0.02,0.11)	0.08 (0.05, 0.12)
rs9296736	GGT	*MLIP*	0.03 (0.00, 0.07)	0.04 (0.02, 0.07)	0.05 (0.03, 0.06)
rs754466	GGT	*DLG5*	Not in UCLEB	0.07 (0.04, 0.10)	0.05 (0.03, 0.07)
rs8038465	GGT	*NPTN -CD276*	0.04 (-0.00, 0.09)	0.00 (-0.02, 0.03)	0.04 (0.02, 0.05)
rs4581712	GGT	*DYNLRB2*	0.09 (-0.00, 0.18)	0.07 (0.04, 0.10)	0.05 (0.03, 0.07)
rs1076540	GGT	*MICAL3*	0.05 (0.01, 0.09)	0.06 (0.03, 0.09)	0.07 (0.06, 0.09)
rs2739330	GGT	*DDT, DDTL, GSTT1, GSTT2BMIF*	0.06 (0.02, 0.10)	0.03 (0.01, 0.06)	0.06 (0.04, 0.08)
rs4820599	GGT	*GGT1*	0.13 (0.10, 0.17)	0.09 (0.06, 0.12)	0.13 (0.11, 0.15)
rs10513686	GGT	*SLC2A2nc*	0.06 (0.01, 0.11)	0.04 (0.01, 0.08)	0.08 (0.05, 0.1)
rs1260326	GGT	*C2orf16e, GCKRnc*	0.03 (-0.00, 0.06)	0.05 (0.02, 0.07)	0.05 (0.03, 0.07)
rs17145750	GGT	*MLXIPLnce*	0.08 (0.03, 0.12)	0.07 (0.03, 0.10)	0.07 (0.04, 0.1)
rs339969	GGT	*RORAn*	0.05 (0.02, 0.09)	0.05 (0.03, 0.08)	0.07 (0.05, 0.09)
rs516246	GGT	*FUT2*	0.04 (0.01, 0.07)	0.04 (0.01, 0.06)	0.04 (0.02, 0.05)
rs7310409	GGT	*HNF1Anc, C12orf27e*	0.11 (0.08, 0.15)	0.10 (0.08, 0.13)	0.1 (0.09, 0.12)
rs944002	GGT	*C14orf73nc*	0.12 (0.08, 0.15)	0.09 (0.06, 0.12)	0.1 (0.08, 0.12)
rs6888304	GGT	*CDH6n*	Not in UCLEB	0.05 (0.02, 0.08)	0.04 (0.02, 0.06)
rs9913711	GGT	*FLJ37644e, SOX9n*	Not in UCLEB	0.05 (0.03, 0.08)	0.04 (0.02, 0.05)
rs12968116	GGT	*ATP8B1ncg*	Not in UCLEB	0.05 (0.01, 0.09)	0.07 (0.04, 0.1)
rs4503880	GGT	*NEDD4Ln*	Not in UCLEB	0.07 (0.04, 0.10)	0.06 (0.03, 0.08)

***** Mapped gene for each SNP is given as reported in the original GWAS publication.

† All SNP-liver function marker associations in the UCLEB studies and the Fenland study were adjusted for age and sex (if relevant). Individual study estimates in UCLEB were combined using fixed effect meta-analyses.

‡ Effect sizes extracted from Chambers *et al*., 2011 (28). GWAS effect sizes for ALT, ALP and GGT were converted to SD units using the median SD from the UCLEB and Fenland studies.

ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; GGT: gamma-glutamyl transferase; GWAS: genome-wide association study; N/A: not applicable; SNP: single nucleotide polymorphism; SD: standard deviation; UCLEB consortium: UCL-LSHTM-Edinburgh-Bristol consortium

Table 2. Mendelian randomization analysis of the effect of insulin resistance (proxied by circulating fasting insulin) on liver function markers

Outcome	Method	Beta	95%CI	P-value
ALT	IVW	0.46	0.22; 0.69	0.0001
	MR Egger	0.14	-1.15; 1.42	0.84
	Weighted median	0.43	0.10; 0.76	0.01
	IVW adjusted by BMI	0.47	0.21; 0.72	0.0004
AST	IVW	0.31	-0.1; 0.71	0.14
	MR Egger	-1.01	-3.31; 1.29	0.41
	Weighted median	0.32	-0.26; 0.89	0.28
	IVW adjusted by BMI	0.25	-0.22; 0.72	0.30
ALP	IVW	-0.14	-0.78; 0.49	0.66
	MR Egger	3.39	0.37; 6.41	0.05
	Weighted median	-0.40	-0.88; 0.07	0.09
	IVW adjusted by BMI	-0.21	-0.6; 0.18	0.29
GGT	IVW	0.28	-0.24; 0.8	0.3
	MR Egger	-2.35	-4.95; 0.25	0.10
	Weighted median	0.51	0.18; 0.84	0.003
	IVW adjusted by BMI	0.22	-0.1; 0.54	0.18

Results correspond to mean difference (in S.D. units) of log10 liver function marker (95% CI) per one S.D increase in fasting insulin (in pmol/L). ALT: alanine aminotransferase; AST: aspartate aminotransferase; ALP: alkaline phosphatase; GGT: gamma-glutamyl transferase; IVW: inverse variance weighted method; S.D.: standard deviation.

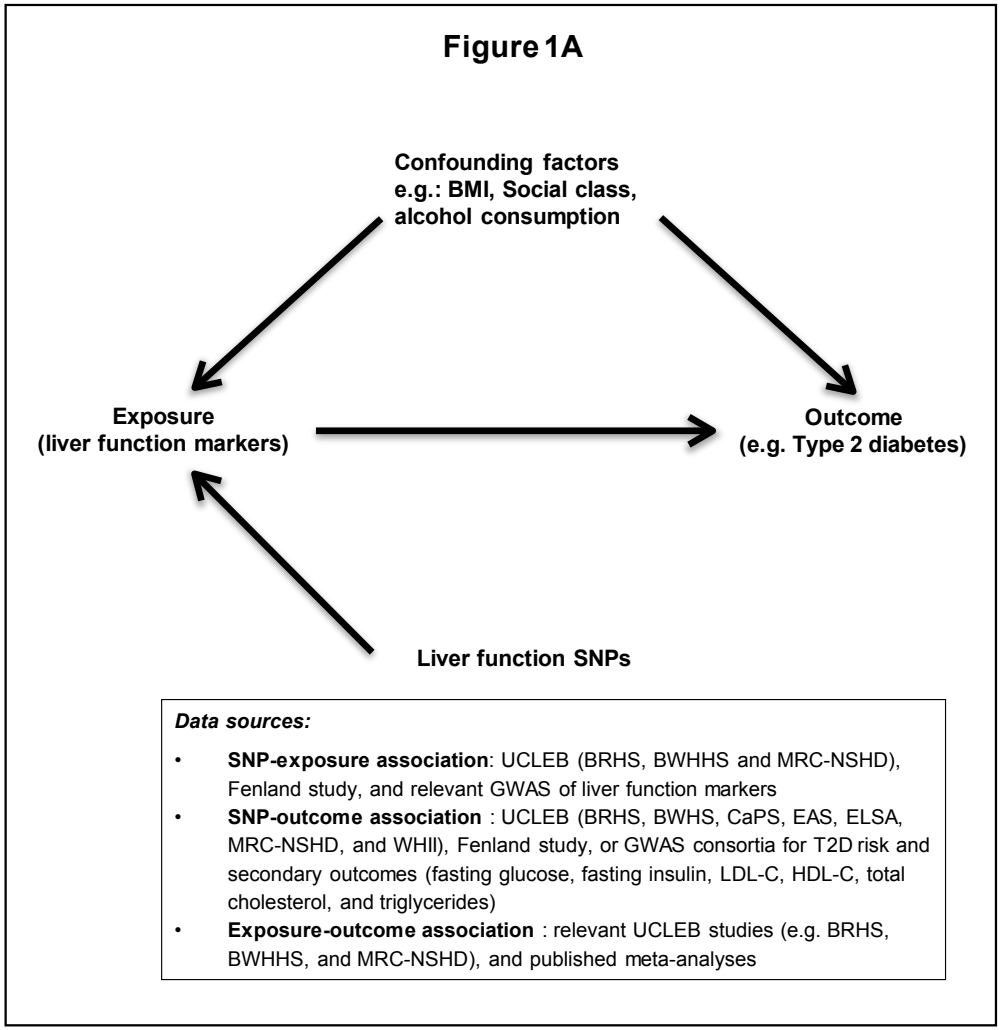
Figure legends

Figure 1. Study design and data sources used to investigate the effect of liver dysfunction (proxied by biomarkers: ALT, AST, ALP, and GGT) on type 2 diabetes or secondary outcomes (fasting glucose, fasting insulin, LDL-C, HDL-C, total cholesterol, and triglycerides) (**A**) and the effect of predisposition to type 2 diabetes or insulin resistance on circulating liver function biomarkers (**B**)

As shown in **figure 1A**, the multivariable association of liver function markers with T2D risk (or related outcomes) was estimated by meta-analysing results from each data source using logistic regression models (or linear regression models in the case of secondary outcomes) with participant-level data from relevant studies within UCLEB consortium (BRHS, BWHHS, MRC-NHSD) and summary-level data from the published meta-analyses of Kunutsor *et al* (2013) and Fraser *et al* (2009). We also estimated the association of liver function markers with T2D risk (or secondary outcomes) using a Mendelian randomization approach. In Mendelian randomization analysis, we used different data sources to estimate the SNP-liver function marker association (UCLEB consortium — BRHS, BWHHS and MRC-NHSD —, Fenland study, and GWAS of liver function markers — Chambers *et al* (2011)) —, and SNP-T2D risk association (UCLEB — BRHS, BWHS, CaPS, EAS, ELSA, MRC-NHSD, and WHII —, and GWAS consortium) or SNP-secondary outcomes. As shown in **figure 1B**, the summary-level data for the association of SNP-T2D risk and SNP-fasting insulin for the reverse MR was extracted from GWAS consortia, and the association of SNP-liver function marker was extracted from Chambers *et al* (2011). ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; BMI: body mass index; BRHS: British Regional Heart Study; BWHHS: British Women's Heart and Health Study; CaPS: Caerphilly Prospective Study; DIAGRAM consortium: Diabetes Genetics Replication And Meta-analysis consortium; EAS: Edinburgh Artery Study; ELSA: English Longitudinal Study of Ageing; GGT: gamma-glutamyl transferase; GWAS: genome-wide association study; HDL-C: high density lipoprotein-cholesterol; LDL-C: low density lipoprotein-cholesterol; MRC-NSHD: National Survey of Health and Development; SNPs: single nucleotide polymorphisms; T2D: type 2 diabetes; UCLEB consortium: UCL-LSHTM-Edinburgh-Bristol consortium; WHII: Whitehall II study.

Figure 2. Multivariable and Mendelian randomization analysis of the effect of liver function on T2D (**A**) and Mendelian randomization analysis of the effect of T2D on liver function markers (**B**).

Results from Figure 1A correspond to odds ratio of T2D per unit increase in standardized liver function markers (and 95% confidence interval). Results from Figure 1B correspond to change in standardized liver function markers per unit increase in log odds of T2D (and 95% confidence interval). I-squared indicates between-study heterogeneity and is only presented when estimates for more than one study were available. ALT: alanine aminotransferase; AST: aspartate aminotransferase; ALP: alkaline phosphatase; GGT: gamma-glutamyl transferase; SNPs: single nucleotide polymorphisms; T2D: type 2 diabetes.



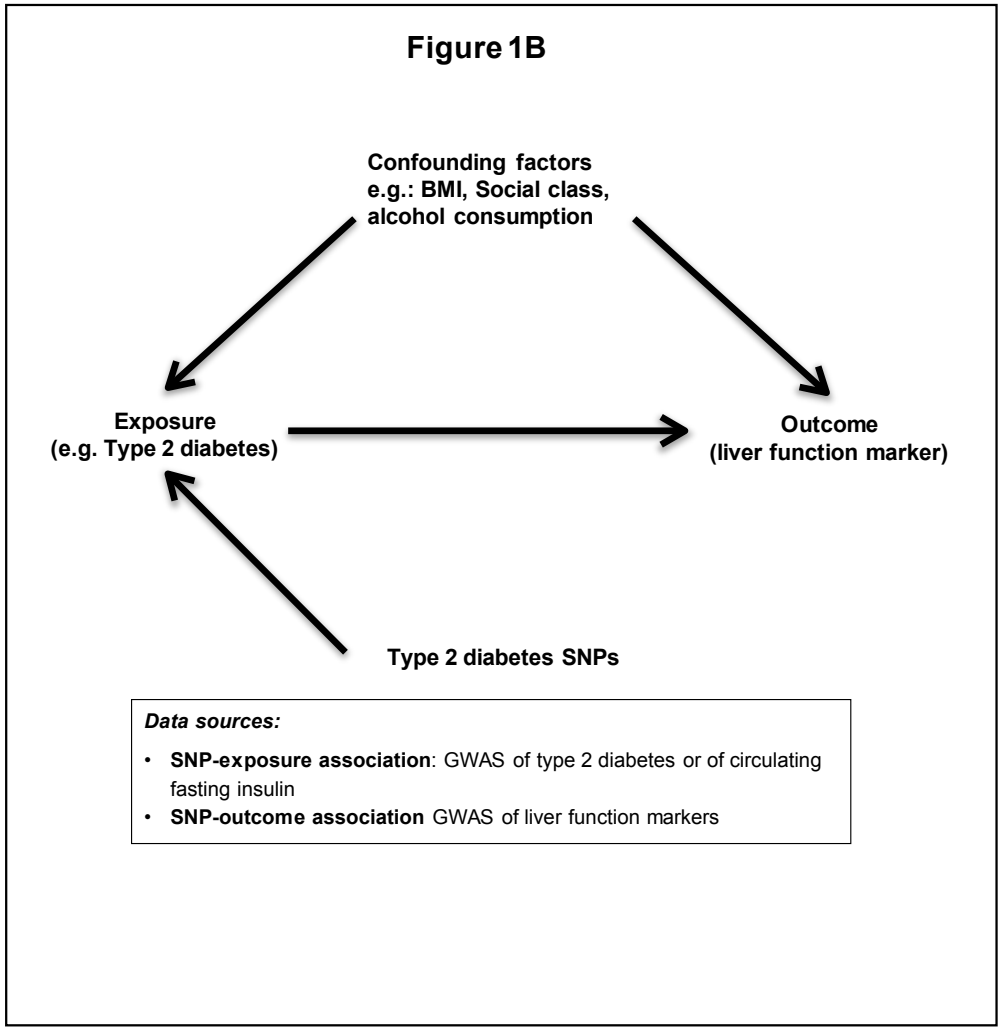


Figure 1. Study design and data sources used to investigate the effect of liver dysfunction (proxied by biomarkers: ALT, AST, ALP, and GGT) on type 2 diabetes or secondary outcomes (fasting glucose, fasting insulin, LDL-C, HDL-C, total cholesterol, and triglycerides) (**Figure 1A**) and the effect of predisposition to type 2 diabetes or insulin resistance on circulating liver function biomarkers (**Figure 1B**)

As shown in **figure 1A**, the multivariable association of liver function markers with T2D risk (or related outcomes) was estimated by meta-analysing results from each data source using logistic regression models (or linear regression models in the case of secondary outcomes) with participant-level data from relevant studies within UCLEB consortium (BRHS, BWHHS, MRC-NHSD) and summary-level data from the published meta-analyses of Kunutsor *et al* (2013) and Fraser *et al* (2009). We also estimated the association of liver function markers with T2D risk (or secondary outcomes) using a Mendelian randomization approach. In Mendelian randomization analysis, we used different data sources to estimate the SNP-liver function marker association (UCLEB consortium — BRHS, BWHHS and MRC-NHSD —, Fenland study, and GWAS of liver function markers — Chambers *et al* (2011)) —, and SNP-T2D risk association (UCLEB — BRHS, BWHS, CaPS, EAS, ELSA, MRC-NHSD, and WHII —,

For Peer Review Only

and GWAS consortium) or SNP-secondary outcomes. As shown in **figure 1B**, the summary-level data for the association of SNP-T2D risk and SNP-fasting insulin for the reverse MR was extracted from GWAS consortia, and the association of SNP-liver function marker was extracted from Chambers *et al* (2011). ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; BMI: body mass index; BRHS: British Regional Heart Study; BWHHS: British Women's Heart and Health Study; CaPS: Caerphilly Prospective Study; DIAGRAM consortium: Diabetes Genetics Replication And Meta-analysis consortium; EAS: Edinburgh Artery Study; ELSA: English Longitudinal Study of Ageing; GGT: gamma-glutamyl transferase; GWAS: genome-wide association study; HDL-C: high density lipoprotein-cholesterol; LDL-C: low density lipoprotein-cholesterol; MRC-NSHD: National Survey of Health and Development; SNPs: single nucleotide polymorphisms; T2D: type 2 diabetes; UCLEB consortium: UCL-LSHTM-Edinburgh-Bristol consortium; WHII: Whitehall II study.

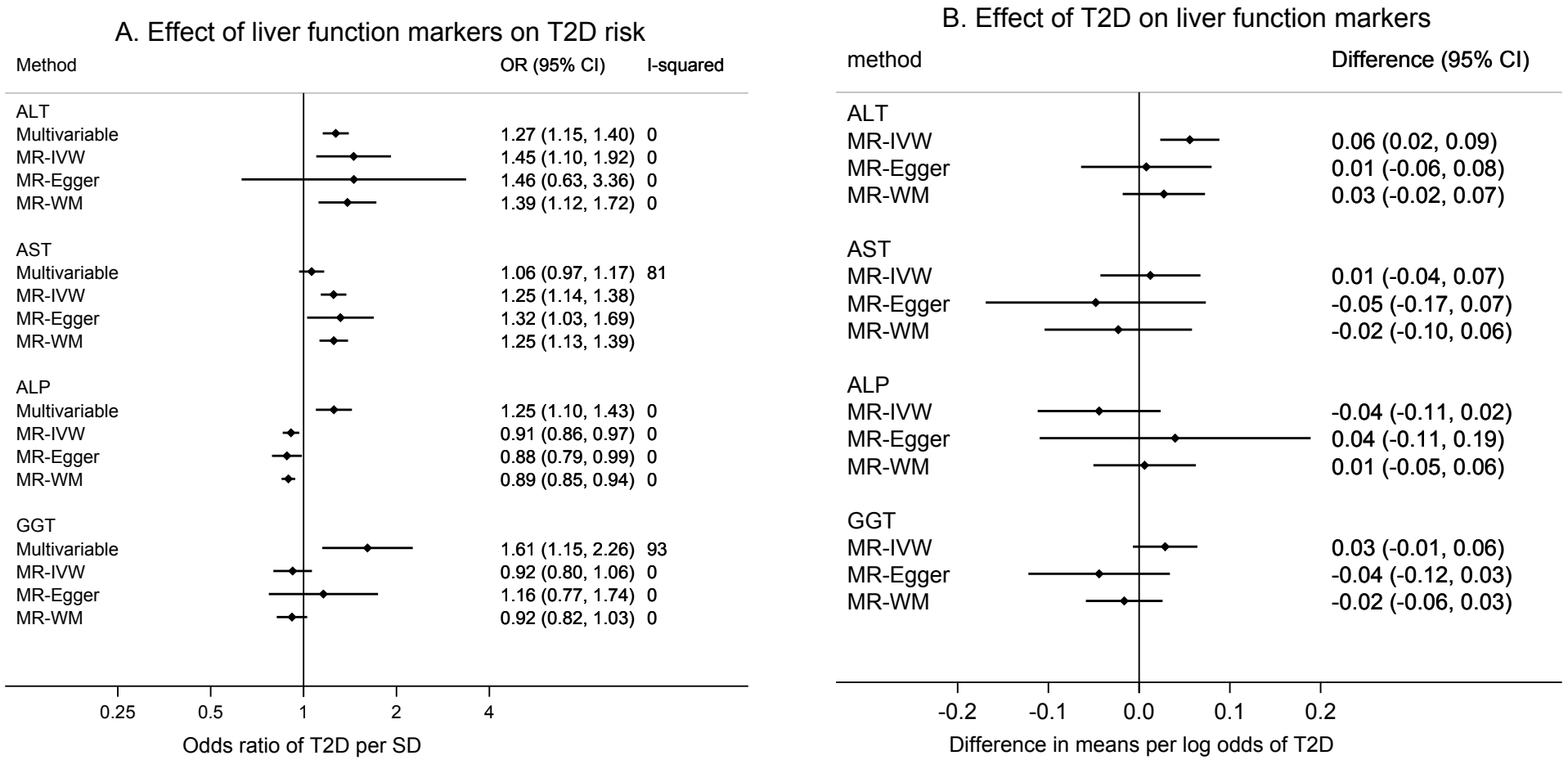


Figure 2. Multivariable and Mendelian randomization analysis of the effect of liver function on T2D (**A**) and Mendelian randomization analysis of the effect of T2D on liver function markers (**B**).

Results from Figure 1A correspond to odds ratio of T2D per unit increase in standardized liver function markers (and 95% confidence interval). Results from Figure 1B correspond to change in standardized liver function markers per unit increase in log odds of T2D (and 95% confidence interval). I-squared indicates between-study heterogeneity and is only presented when estimates for more than one study were available. ALT: alanine aminotransferase; AST: aspartate aminotransferase; ALP: alkaline phosphatase; GGT: gamma-glutamyl transferase; SNPs: single nucleotide polymorphisms; T2D: type 2 diabetes.

Supplementary table 1. Number of participants with data on exposures, outcomes and covariates/potential confounders in UCLEB and Fenland studies used in the multivariable and Mendelian randomization analyses

Trait	Variable type*	Available studies	Total N**
ALT (U/L)	Exposure	BRHS, BWHHS, NSHD & Fenland	15371
ALP (U/L)	Exposure	CAPS, NSHD & Fenland	12761
AST (U/L)	Exposure	BRHS, BWHHS, CAPS & NSHD	6624
GGT (U/L)	Exposure	BRHS, BWHHS, CAPS, EAS, NSHD & Fenland	17336
Fasting glucose (mmol/l)	Continuous outcome	BRHS, BWHHS, CAPS, EAS, ELSA, NSHD & WHII	9689
Fasting insulin (µIU/ml)	Continuous outcome	BRHS, BWHHS, CAPS, NSHD & WHII	7024
LDL cholesterol (mmol/L)	Continuous outcome	BRHS, BWHHS, CAPS, EAS, ELSA, WHII, NSHD & Fenland	21296
HDL cholesterol (mmol/L)	Continuous outcome	BRHS, BWHHS, CAPS, EAS, ELSA, WHII, NSHD & Fenland	21464
Total cholesterol (mmol/L)	Continuous outcome	BRHS, BWHHS, CAPS, EAS, ELSA, WHII, NSHD & Fenland	21639
Triglycerides (mmol/l)	Continuous outcome	BRHS, BWHHS CAPS, EAS, ELSA, WHII, NSHD & Fenland	20856
Type 2 diabetes cases	Disease outcome	BRHS, BWHHS, CAPS, EAS, ELSA, WHII, NSHD	1202
Type 2 diabetes controls			10588
Age (years)	Covariate/Potential Confounder	BRHS, BWHHS CAPS, EAS, ELSA, WHII, NSHD & Fenland	21772
BMI (kg/m^2)$^{\$}$	Potential confounder	BRHS, BWHHS CAPS, EAS, ELSA, WHII, NSHD & Fenland	21513
Waist circumference (cm)$^{\$}$	Potential confounder	BRHS, BWHHS, ELSA, WHII, NSHD & Fenland	19595
Sex	Covariate/Potential confounder	BRHS, BWHHS CAPS, EAS, ELSA, WHII, NSHD & Fenland	21772
Smoking$^{\$}$	Potential confounder	BRHS, BWHHS, CAPS, EAS, ELSA, WHII, NSHD & Fenland	21631
Alcohol consumption$^{\$}$	Potential confounder	BRHS, BWHHS, CAPS, EAS, ELSA, WHII, NSHD & Fenland	21501
Social class$^{\$}$	Potential confounder	BRHS, BWHHS, CAPS, EAS, ELSA, NSHD & Fenland	8563

*Variable types as used in the multivariable/Mendelian randomization study.

$^{\$}$Only the multivariable analyses were adjusted for these confounders.

**Number of individuals with non-missing data for each trait in the analysis sample. For each UCLEB study, we included individuals who had complete data for all available genetic variants, age, sex, the liver function markers measured in the study, and were not on lipid lowering medication. For those studies that were only used to estimate the gene-outcome association (i.e.: studies where none of the liver function markers were available), we included individuals who had complete data for all available genetic variants, age and sex only.

ALT: alanine aminotransferase; ALP: alkaline phosphatase; AST: aspartate aminotransferase; GGT: gamma-glutamyl transferase; BMI: body mass index; UCLEB consortium: UCL-LSHTM-Edinburgh-Bristol consortium – BRHS: British Regional Heart Study; BWHHS: British Women's Heart and Health Study; CaPS: Caerphilly Prospective Study; EAS: Edinburgh Artery Study; ELSA: English Longitudinal Study of Ageing; NSHD: MRC National Survey of Health and Development, WHII: Whitehall II study.

Supplementary table 2a. Distribution of participant characteristics (continuously measured variables) within each UCLEB study and the Fenland study

Variable	BRHS	BWHHS	CAPS	EAS	ELSA	WHII	NSHD	Fenland
N[1]	2090	1733	1221	762	1688	2716	1580	9982
Age (years)	69.01 (5.66)	71.55 (5.31)	56.74 (4.47)	64.51 (5.64)	73.93 (9.60)	60.39 (5.91)	53 (N/A)	48.14 (7.44)
BMI (kg/m2) [2]	26.48 (24.47, 28.83) [2077]	27 (24, 30) [1714]	26.44 (24.40, 28.67) [1202]	25.21 (23.15, 27.40) [762]	26.91 (24.36, 29.96) [1571]	26.00 (23.80, 28.60) [2708]	26.32 (24.15, 29.34) [1574]	26.16 (29.31-23.56) [9905]
Waist circumference (cm)	96.71 (10.22) [2071]	86.35 (12.28) [1707]	Not available	Not available	95.91 (12.78) [1627]	91.29 (12.18) [2713]	90.48 (12.70) [1578]	90.86 (13.43) [9899]
Physical activity (m/s2) [3]	Not available	Not available	Not available	Not available	Not available	Not available	Not available	0.11 (0.09, 0.15)
ALT (U/L) [2]	15 (12, 20)	12 (9, 15)	Not available	Not available	Not available	Not available	26 (21, 34)	25 (19, 34) [9968]
ALP (U/L) [2]	Not available	Not available	89 (75, 105)	Not available	Not available	Not available	79 (67, 94)	83 (69, 98) [9960]
AST (U/L) [2]	23 (19, 27)	22 (19, 25)	19 (16, 23)	Not available	Not available	Not available	24 (21, 29)	Not available
GGT (U/L)[2]	26 (19, 38)	20 (14, 31)	30 (22, 44)	22.50 (16.13, 33.29)	Not available	Not available	29.6 (21.6, 42.3)	26 (20, 38) [9950]
Fasting glucose (mmol/L) [2,4]	5.30 (0.65) [1820]	5.70 (0.50) [1498]	5.10 (0.49) [1055]	5.52 (0.55) [685]	4.90 (0.49) [702]	5.26 (0.49) [2509]	5.48 (0.56) [1420]	Not available
Fasting insulin (μIU/ml) [2,4]	7.70 (5.20, 11.30) [1813]	6.10 (4.40, 9.20) [1498]	4.82 (3.17, 7.10) [535]	Not available	Not available	6.7 (4.6, 10.1) [2438]	42 (28, 60) [740]	Not available
LDL cholesterol (mmol/L)	3.96 (0.98) [2060]	4.25 (1.07) [1687]	3.72 (0.88) [1178]	5.35 (1.24) [759]	3.49 (1.05) [1645]	3.62 (0.87) [2683]	3.53 (0.94) [1374]	3.43 (0.90) [9910]
HDL cholesterol (mmol/L)	1.15 (0.25) [2004]	1.62 (0.45) [1731]	1.02 (0.25) [1214]	1.45 (0.37) [759]	1.50 (0.39) [1677]	1.57 (0.45) [2715]	1.70 (0.51) [1382]	1.53 (0.41)
Total cholesterol (mmol/L)	6.29 (0.99) [2082]	6.73 (1.20) [1733]	5.60 (0.99) [1214]	7.10 (1.33) [762]	5.79 (1.27) [1678]	5.80 (0.95) [2715]	6.10 (1.03) [1473]	5.45 (1.02)
Triglycerides (mmol/L) [2]	1.71 (1.2, 2.48) [1302]	1.62 (1.21, 2.30) [1733]	1.64 (1.18, 2.34) [1214]	1.35 (1.01, 1.78) [762]	1.50 (1.1, 2.2) [1678]	1.10 (0.80, 1.60) [2715]	1.7 (1.2, 2.5) [1470]	1.0 (0.7, 1.4)

[1] Number of individuals in the analysis sample of the Mendelian randomization study. The analysis sample for each UCLEB study included individuals who had complete data for all available SNPs, age, sex, the liver function markers measured in the study, and were not on lipid lowering medication. Note that the analysis sample in UCLEB was not defined based on the availability of complete data for the potential confounders and the outcome variables. Therefore, the number of individuals included in analyses involving the confounders and the outcome variables depends on the number of individuals with non-missing data for each outcome /confounder variable in a given study and is given in brackets if it differed from the N in the analysis sample. Analysis sample in the Fenland study included individuals who were not on lipid lowering medication and had complete data on age, sex and at least 80% of the SNPs available.

[2] Distributions for these variables are median (interquartile range; IQR) because the distributions were right skewed; all other results are means (standard deviation; SD).

[3] Physical activity was assessed as average acceleration trunk longitudinal axis (m/s2).

[4] For fasting glucose and fasting insulin we removed individuals with type 2 diabetes and/or fasting glucose values ≥ 7.0 mol/L and/or individuals who were taking glucose lowering drugs.

ALT: alanine aminotransferase; ALP: alkaline phosphatase; AST: aspartate aminotransferase; GGT: gamma-glutamyl transferase BMI: body mass index; UCLEB consortium: UCL-LSHTM-Edinburgh-Bristol consortium – BRHS: British Regional Heart Study; BWHHS: British Women's Heart and Health Study; CaPS: Caerphilly Prospective Study; EAS: Edinburgh Artery Study; ELSA: English Longitudinal Study of Ageing; NSHD: MRC National Survey of Health and Development, WHII: Whitehall II study; HDL-C: high density lipoprotein-cholesterol; LDL-C: low density lipoprotein-cholesterol

Supplementary table 2b. Distribution of participant characteristics (categorical variables) within each UCLEB study and the Fenland study

Variable		BRHS	BWHHS	CaPS	EAS	ELSA	WHII	NSHD	Fenland
N*		2090	1733	1221	762	1688	2716	1580	9982
Sex	Men	2090	0	1221	369	882	2073	747	4606
	Women	0	1733	0	393	806	643	833	5376
	Missing[a]	0	0	0	0	0	0	0	0
Smoking	Never smokers	585	933	240	292	573	1345	472	5329
	Ever smokers	1503	800	978	456	1115	1368	1033	4609
	Missing[a]	2	0	3	14	0	3	75	44
Alcohol consumption	Never	78	306	70	83	173	67	309	1556
	Ever	1953	1258	1147	679	1515	2623	1271	8413
	Missing	59	169	4	0	0	26	0	13
Social class	Unskilled	62	135	59	25	344		52	
	Semi-skilled	168	171	224	56	325		147	
	Manual skilled	850	432	707	187	350		232	
	Non-manual skilled	192	214	173	139	365	Not available	359	Not available
	Managerial and lower professional	567	282	57	263	290		619	
	Professional	191	124	0	88	0		114	
	Missing	60	375	1	4	14	N/A	57	N/A
Type 2 Diabetes	Yes	251	230	149	76	210	159	127	Not available
	No	1839	1503	1072	686	1478	2557	1453	
	Missing	0	0	0	0	0	0	0	N/A

* Number of individuals in the analysis sample of the Mendelian randomization study. The analysis sample for each study included individuals who had complete data for all available SNPs, age, sex, the liver function markers measured in the study, and were not on lipid lowering medication. N/A: not applicable. UCLEB consortium: UCL-LSHTM-Edinburgh-Bristol consortium – BRHS: British Regional Heart Study; BWHHS: British Women's Heart and Health Study; CaPS: Caerphilly Prospective Study; EAS: Edinburgh Artery Study; ELSA: English Longitudinal Study of Ageing; NSHD: MRC National Survey of Health and Development, WHII: Whitehall II study.

Supplementary table 3. List of index SNPs associated with ALT, AST, ALP and GGT at genome-wide levels of significance ($P<5 \times 10^{-8}$) selected from published GWAS studies[1]

Index SNP from GWAS	Associated liver function marker	Chr	Mapped gene/s[2]	Publication Index SNP selected from	Index SNP in all relevant UCLEB studies	Genotyped/ imputed in UCLEB studies	Index SNP in Fenland study	Genotyped/ imputed in Fenland	Proxy SNP used if applicable	R^2 of proxy with index SNP
rs6834314	ALT	4	*HSD17B13, MAPK10*	Chambers et al_2011	Yes	imputed	Yes	imputed	N/A	N/A
rs2954021	ALP/ALT	8	*TRIB1*	Chambers et al_2011	Yes	imputed	Yes	imputed	N/A	N/A
rs11597390	ALT	10	*CPN1*	Yuan et al_2008	Yes	genotyped	Yes	genotyped	N/A	N/A
rs2143571	ALT	22	*SAMM50*	Yuan et al_2008	Yes	imputed	Yes	genotyped	N/A	N/A
rs17109512	AST	10	*GOT1*	Shen et al_2011	Yes except in CAPS and EAS [3]	imputed	Yes	genotyped	rs7073497	1.00
rs738407	AST	22	*PNPLA3*	Novel	Yes	imputed	Yes	imputed	N/A	N/A
rs738408	AST	22	*PNPLA3*	Novel	Yes	imputed	Yes	imputed	N/A	N/A
rs1780324	ALP	1	*NBPF3-ALPL*	Yuan et al_2008	Yes	genotyped	Yes	genotyped	N/A	N/A
rs16856332	ALP	2	*ABCB11*	Chambers et al_2011	No	N/A	Yes	genotyped	rs11892966	1.00
rs9467160	ALP	6	*GPLD1*	Yuan et al_2008	Yes	imputed	Yes	genotyped	N/A	N/A
rs6984305	ALP	8	*PPP1R3B*	Chambers et al_2011	Yes	genotyped	Yes	genotyped	N/A	N/A
rs514708	ALP	9	*ABO*	Yuan et al_2008	Yes	genotyped	No	N/A	rs8176714	1.00
rs10819937	ALP	9	*ALDOB, C9orf125*	Chambers et al_2011	No	N/A	Yes	genotyped	No suitable proxy SNP ($R^2 \geq 0.8$) in UCLEB	N/A
rs579459	ALP	9	*ABO*	Chambers et al_2011	Yes	genotyped	Yes	genotyped	N/A	N/A
rs7923609	ALP	10	*JMJD1C, NRBF2*	Chambers et al_2011	Yes	genotyped	Yes	genotyped	N/A	N/A
rs174601	ALP	11	*C11orf10, FADS1, FADS2*	Chambers et al_2011	Yes	genotyped	Yes	genotyped	N/A	N/A
rs2236653	ALP	11	*ST3GAL4*	Chambers et al_2011	Yes	imputed	Yes	imputed	N/A	N/A
rs7186908	ALP	16	*HPR, PMFBP1*	Chambers et al_2011	Yes	imputed	Yes	imputed	N/A	N/A
rs314253	ALP	17	*ASGR1, DLG4*	Chambers et al_2011	Yes	genotyped	Yes	genotyped	N/A	N/A
rs281377	ALP	19	*FUT2*	Chambers et al_2011	Yes	imputed	Yes	genotyped	N/A	N/A
rs7267979	ALP	20	*ABHD12, GINS1, PYGB*	Chambers et al_2011	Yes	genotyped	Yes	genotyped	N/A	N/A
rs1497406	GGT	1	*RSG1, EPHA2*	Chambers et al_2011	Yes	imputed	Yes	genotyped	N/A	N/A
rs12145922	GGT	1	*CCBL2, PKN2*	Chambers et al_2011	Yes	imputed	Yes	genotyped	N/A	N/A
rs1335645	GGT	1	*CEPT1, DENNDD*	Chambers et al_2011	No	N/A	Yes	genotyped	No suitable proxy SNP ($R^2 \geq 0.8$) in UCLEB	N/A
rs10908458	GGT	1	*DPM3, EFNA1, PKLR*	Chambers et al_2011	Yes	imputed	Yes	imputed	N/A	N/A
rs1260326	GGT	2	*C2orf16, GCKR*	Chambers et al_2011	Yes	genotyped	Yes	genotyped	N/A	N/A
rs13030978	GGT	2	*MYO1B, STAT4*	Chambers et al_2011	Yes	imputed	Yes	genotyped	N/A	N/A
rs2140773	GGT	2	*EFHD1, LOC100129166*	Chambers et al_2011	Yes except in the EAS [4]	genotyped	Yes	genotyped	No suitable proxy SNP ($R^2 \geq 0.8$) in EAS	N/A
rs10513686	GGT	3	*SLC2A2*	Chambers et al_2011	Yes	genotyped	Yes	genotyped	N/A	N/A
rs4547811	GGT	4	*ZNF827*	Chambers et al_2011	Yes	imputed	Yes	genotyped	N/A	N/A
rs6888304	GGT	5	*CDH6*	Chambers et al_2011	No	N/A	Yes	genotyped	No suitable proxy SNP ($R^2 \geq 0.8$) in UCLEB	N/A
rs4074793	GGT	5	*ITGA1*	Chambers et al_2011	Yes	imputed	Yes	genotyped	N/A	N/A
rs9296736	GGT	6	*MLIP*	Chambers et al_2011	Yes	genotyped	Yes	imputed	N/A	N/A
rs17145750	GGT	7	*MLXIPL*	Chambers et al_2011	Yes	genotyped	Yes	genotyped	N/A	N/A
rs754466	GGT	10	*DLG5*	Chambers et al_2011	No	N/A	Yes	imputed	No suitable proxy SNP ($R^2 \geq 0.8$) in UCLEB	N/A

rs7310409	GGT	12	*HNF1A, C12orf7*	Chambers et al_2011	Yes	genotyped	Yes	genotyped	N/A	N/A
rs944002	GGT	14	*C14orf73*	Chambers et al_2011	Yes	imputed	Yes	genotyped	N/A	N/A
rs339969	GGT	15	*RORA*	Chambers et al_2011	Yes	imputed	Yes	genotyped	N/A	N/A
rs8038465	GGT	15	*CD276*	Chambers et al_2011	No except in the BRHS study [5]	imputed	Yes	genotyped	rs62004612	0.87
rs4581712	GGT	16	*DYNLRB2*	Chambers et al_2011	Yes except in BWHHS, EAS and CAPS [6]	imputed	Yes	genotyped	No suitable proxy SNP present in BWHHS, EAS and CAPS stuies	N/A
rs9913711	GGT	17	*FLJ37644, SOX9*	Chambers et al_2011	No	N/A	Yes	genotyped	No suitable proxy SNP ($R^2 \geq 0.8$) in UCLEB	N/A
rs12968116	GGT	18	*ATP8B1*	Chambers et al_2011	No	N/A	Yes	genotyped	No suitable proxy SNP ($R^2 \geq 0.8$) in UCLEB	N/A
rs4503880	GGT	18	*NEDD4L*	Chambers et al_2011	No	N/A	Yes	imputed	No suitable proxy SNP ($R^2 \geq 0.8$) in UCLEB	N/A
rs516246	GGT	19	*FUT2*	Chambers et al_2011	Yes	genotyped	Yes	genotyped	N/A	N/A
rs4820599	GGT	22	*GGT1*	Yuan et al_2008	Yes	genotyped	Yes	genotyped	N/A	N/A
rs1076540	GGT	22	*MICAL3*	Chambers et al_2011	Yes	imputed	Yes	genotyped	N/A	N/A
rs2739330	GGT	22	*DDT, DDTL, GSTT1, GSTT 2B, MIF*	Chambers et al_2011	Yes	imputed	Yes	imputed	N/A	N/A

[1] SNPs mapped to the same gene, as reported in the original GWAS publications, were considered as part of the same locus. For loci where the index SNP reported for the same locus was different in independent publications, we selected the SNP that was genotyped in the UCLEB studies. If the different index SNPs reported in independent publications for the same locus were all imputed or genotyped in UCLEB, we selected the one that showed the strongest association in the largest GWAS study. We used a proxy SNP ($R^2 \geq 0.8$) if the index SNP was not present across all or some of relevant the UCLEB studies. If there was more than one eligible proxy, we selected the proxy that was genotyped in the UCLEB studies, and if all eligible proxies were genotyped or imputed we selected a proxy SNP in random. There were seven index SNPs in total (one SNP associated with ALP and six SNPs associated with GGT) that were not genotyped/imputed in any of the UCLEB studies, and did not have a suitable proxy available. These SNPs were therefore not included as genetic instruments in the UCLEB studies. However, all SNPs (or a suitable proxy) associated with a given liver function marker were available in the GWAS and Fenland studies, and were used as genetic instruments in those studies. One further GWAS has been published since we completed our search (July 2017), which found a novel rare variant associated with ALT (rs28929474; MAF=2%). Given that this variant is rare and is not available in T2D GWAS, which is the largest dataset contributing to the MR estimates, we have not included it in our analyses as it would add little information.

[2] Mapped gene for each SNP is given as reported in the original publication.

[3] Index SNP was absent in CAPS and EAS studies therefore we used the SNP rs7073497 in UCLEB studies, which is perfectly correlated with the index SNP and was available across all UCLEB studies.

[4] This SNP was not present in EAS study and there was no suitable proxy SNP present across all studies. Therefore, we did not use the EAS study to estimate the SNP-exposure and SNP-outcome for this locus.

[5] This SNP was used instead of the index SNP as it was present across all the relevant UCLEB studies and is correlated at R^2=0.87 with the index SNP.

[6] This SNP was absent in BWHHS, EAS and CAPS and there was no suitable proxy that was available across all studies. Therefore, we did not use the BWHHS, EAS and the CAPS study to estimate the SNP-exposure and SNP-outcome for this locus.

ALT: alanine aminotransferase; ALP: alkaline phosphatase; AST: aspartate aminotransferase; GGT: gamma-glutamyl transferase; SNPs: single nucleotide polymorphisms; UCLEB consortium: UCL-LSHTM-Edinburgh-Bristol consortium; GWAS: Genome-wide association study

Supplementary table 4. Association of novel variants with AST plasma concentration in the replication sample (NFBC1966 and NSHD studies)

SNP	Study	N	Effect size (SD units)	SE	P-value	MAF
rs738407	NFBC1966	3614	0.043	0.022	0.049	0.48
rs738407	NSHD	1580	0.088	0.057	0.120	0.34
	Overall	**5194**	**0.049**	**0.020**	**0.017**	
rs738408	NFBC1966	3614	0.071	0.026	0.006	0.23
rs738408	NSHD	1580	0.087	0.050	0.080	0.16
	Overall	**5194**	**0.075**	**0.023**	**0.001**	

The genome-wide association analyses were carried out in each individual study (BRHS, BWHHS, CaPS, EAS, and ET2DS) using the SnpStats software in R. The individual study estimates were combined using the inverse variance- fixed effect meta-analysis using the METAL software. Monomorphic SNPs, SNPs that were not sufficiently well imputed (Rsq < 0.3), and SNPs with MAF <0.001 were filtered out prior to meta-analyses. Genome-wide significance was inferred at $p< 5 \times 10^{-8}$. Two novel variants (rs738407 and rs738408) near *PNPLA3* gene were associated with AST at genome-wide levels of significance ($P<5 \times 10^{-8}$) in the discovery stage and were replicated in the MRC NHSD study (n = 1,580) (1) and 1966 Northern Finland Cohort (NFBC1966) (n = 3,614) (2). AST: aspartate aminotransferase; BRHS: British Regional Heart Study; BWHHS: British Women's Heart and Health Study; CaPS: Caerphilly Prospective Study; EAS: Edinburgh Artery Study; ET2DS: Edinburgh Type 2 Diabetes Study; MAF: minor allele frequency; NSHD: MRC National Survey of Health and Development; NFBC1966: Northern Finland Birth Cohort 1966; SD: standard deviation; SNP: single nucleotide polymorphism; UCLEB consortium: UCL-LSHTM-Edinburgh-Bristol consortium.

Supplementary table 5. The associations of individual SNPs (used as genetic instruments for type 2 diabetes risk in Mendelian randomization analyses) with type 2 diabetes

SNP	Effect allele	Other allele	Beta	SE	P-value
rs2296173	G	A	0.065	0.009	7.7E-14
rs12088739	G	A	-0.088	0.013	9.8E-12
rs1127655	C	T	0.044	0.008	2.5E-08
rs2493394	G	A	0.073	0.011	1.2E-10
rs340874	T	C	-0.063	0.007	8.4E-18
rs2820426	A	G	-0.052	0.007	1.3E-12
rs348330	G	A	0.049	0.008	1.9E-09
rs2867125	T	C	-0.060	0.010	4.3E-10
rs780094	T	C	-0.069	0.007	5.2E-21
rs17334919	T	C	-0.140	0.013	6.7E-28
rs243019	C	T	0.057	0.007	2.3E-15
rs1009358	C	T	-0.055	0.008	9.8E-12
rs10169613	T	C	-0.043	0.008	3.6E-08
rs12617659	T	C	-0.069	0.010	2.8E-11
rs7572970	A	G	-0.059	0.009	1.4E-11
rs13389219	T	C	-0.072	0.007	2.1E-22
rs2972144	A	G	-0.091	0.008	2.6E-34
rs7561798	G	A	0.040	0.007	2.8E-08
rs1899951	T	C	-0.112	0.011	1.6E-24
rs1496653	G	A	-0.077	0.009	2.6E-18
rs11926707	T	C	-0.046	0.008	1.7E-08
rs2292662	T	C	-0.063	0.011	1.2E-08
rs6795735	T	C	-0.056	0.007	1.6E-14
rs11708067	G	A	-0.097	0.009	5.9E-29
rs9844972	C	G	0.096	0.015	1.0E-10
rs4472028	T	C	0.045	0.007	2.1E-10
rs11925227	A	G	-0.053	0.010	2.3E-08
rs7651090	G	A	0.120	0.008	3.9E-57
rs3887925	C	T	-0.047	0.008	2.5E-09
rs6808574	T	C	-0.055	0.008	4.4E-13
rs1801214	C	T	-0.090	0.007	5.5E-34
rs17086692	T	G	-0.047	0.008	2.5E-08
rs993380	A	G	0.051	0.008	4.6E-10
rs7674212	T	G	-0.047	0.008	6.2E-10
rs11098676	T	C	-0.054	0.010	2.0E-08
rs7685296	T	C	-0.051	0.008	2.3E-10
rs735949	C	T	-0.071	0.011	1.9E-11
rs1061813	G	A	0.043	0.007	3.4E-09
rs4865796	G	A	-0.053	0.008	1.3E-11
rs459193	A	G	-0.071	0.008	8.8E-18
rs3900856	A	G	0.114	0.019	7.4E-10
rs2307111	C	T	-0.041	0.007	3.0E-08
rs7729395	T	C	0.137	0.016	1.1E-17
rs10077431	A	C	-0.049	0.009	4.8E-08
rs1050226	G	A	-0.049	0.007	3.3E-11
rs7756992	G	A	0.130	0.008	6.0E-62
rs2857605	C	T	-0.067	0.009	5.9E-14
rs2071479	T	C	0.147	0.023	6.6E-11
rs9369425	G	A	0.055	0.009	1.1E-10

For Peer Review Only

rs72892910	T	G	0.065	0.010	6.4E-11
rs853974	T	C	0.060	0.009	7.9E-12
rs2246012	C	T	0.053	0.009	2.4E-08
rs622217	C	T	-0.049	0.008	3.1E-10
rs17168486	T	C	0.074	0.009	2.2E-15
rs2191348	G	T	-0.065	0.007	3.4E-19
rs849135	G	A	0.100	0.007	1.0E-43
rs2908282	A	G	0.055	0.009	4.3E-09
rs2299383	T	C	0.041	0.007	1.5E-08
rs13239186	T	C	0.054	0.009	2.7E-10
rs13234269	A	T	-0.058	0.008	7.0E-14
rs7786095	G	A	-0.074	0.013	9.6E-09
rs7841082	T	C	-0.042	0.008	4.9E-08
rs11774915	T	C	0.050	0.009	8.7E-09
rs10100265	A	C	0.049	0.008	6.3E-10
rs17411031	G	C	-0.045	0.008	3.0E-08
rs10087241	G	A	0.048	0.008	2.8E-09
rs12681990	C	T	0.063	0.010	3.6E-11
rs516946	T	C	-0.082	0.009	3.2E-22
rs7845219	C	T	-0.042	0.007	4.5E-09
rs3802177	A	G	-0.122	0.008	2.3E-52
rs2294120	G	A	-0.044	0.008	1.6E-08
rs10974438	C	A	0.059	0.008	3.0E-15
rs1063192	G	A	-0.063	0.007	3.3E-18
rs10811661	C	T	-0.157	0.010	4.1E-58
rs1758632	C	G	-0.049	0.008	1.4E-09
rs17791483	G	A	-0.102	0.015	3.4E-12
rs2796441	A	G	-0.072	0.007	2.0E-22
rs10114341	C	T	-0.041	0.007	1.2E-08
rs687621	G	A	0.043	0.008	1.4E-08
rs11257655	T	C	0.074	0.009	2.0E-17
rs2616132	A	G	0.046	0.008	6.6E-09
rs2633310	T	G	-0.044	0.008	2.4E-08
rs753270	T	C	-0.053	0.008	2.7E-11
rs7923866	T	C	-0.097	0.007	9.3E-40
rs11591741	C	G	-0.048	0.008	1.2E-09
rs7903146	T	C	0.306	0.008	1.0E-200
rs4918796	C	T	0.062	0.009	4.0E-13
rs2421016	T	C	-0.046	0.007	1.5E-10
rs2237892	T	C	-0.096	0.016	8.7E-10
rs5215	C	T	0.068	0.007	2.1E-20
rs7929543	C	A	0.083	0.014	2.2E-09
rs1552224	C	A	-0.103	0.010	8.6E-25
rs10830963	G	C	0.091	0.008	5.8E-30
rs7931302	C	A	0.046	0.008	7.7E-09
rs67232546	T	C	0.060	0.010	4.7E-10
rs11048456	C	T	0.049	0.008	3.0E-09
rs10842994	T	C	-0.076	0.009	1.0E-16
rs2261181	T	C	0.099	0.012	9.2E-17
rs1480474	G	A	0.041	0.007	1.7E-08
rs7138300	C	T	0.044	0.007	5.6E-10
rs11107116	T	G	0.047	0.009	3.8E-08
rs61953351	T	G	-0.070	0.009	2.0E-14

rs940904	G	A	-0.050	0.008	2.1E-09
rs825476	C	T	-0.052	0.007	6.8E-13
rs576674	G	A	0.065	0.010	1.8E-11
rs963740	T	A	-0.048	0.009	2.2E-08
rs1359790	A	G	-0.080	0.008	2.8E-23
rs7144011	T	G	0.048	0.009	1.6E-08
rs4502156	C	T	-0.041	0.007	1.7E-08
rs982077	A	G	0.045	0.007	2.6E-10
rs7177055	G	A	-0.065	0.008	2.7E-16
rs4932143	G	C	0.057	0.009	5.5E-11
rs12910825	G	A	0.052	0.007	2.2E-12
rs9940149	A	G	-0.058	0.010	9.3E-10
rs7185735	G	A	0.106	0.007	1.6E-47
rs244415	A	G	-0.047	0.008	3.9E-09
rs77258096	A	C	-0.117	0.013	1.8E-18
rs2925979	T	C	0.053	0.008	9.1E-12
rs8068804	A	G	0.059	0.008	4.4E-14
rs12945601	T	C	0.048	0.008	1.7E-09
rs17405722	A	G	0.087	0.015	2.3E-09
rs9911983	C	T	-0.040	0.007	4.8E-08
rs9894220	G	A	-0.059	0.008	1.5E-13
rs302864	A	G	0.071	0.013	2.5E-08
rs17631783	T	C	-0.049	0.009	3.9E-08
rs7240767	C	T	0.045	0.008	2.2E-08
rs12970134	A	G	0.056	0.008	5.3E-12
rs10401969	C	T	0.092	0.013	4.1E-12
rs8108269	G	T	0.064	0.008	3.1E-16
rs6515236	C	A	-0.050	0.009	3.3E-08
rs6059662	A	G	-0.045	0.008	1.5E-08
rs4810426	T	C	0.073	0.013	2.1E-08
rs6066138	A	G	-0.049	0.008	1.9E-09
rs16988333	G	A	-0.075	0.013	9.2E-09
rs4823182	G	A	0.048	0.008	3.4E-10
rs6878122	G	A	0.056	0.008	1.2E-12
rs11651755	C	T	0.074	0.008	9.0E-22
rs12299509	G	A	0.047	0.007	2.1E-10
rs1063355	T	G	-0.071	0.008	3.7E-19

SNP: single nucleotide polymorphism; SE: standard deviation. Beta and SE expressed as change in log odds ratio of T2D per effect allele.

Supplementary table 6. The associations of individual SNPs (used as genetic instruments for fasting insulin in Mendelian randomization analyses) with fasting insulin

SNP	Effect allele	Other allele	Effect allele frequency	Beta	SE	P-value
rs10195252	C	T	0.44	-0.02025	0.003291	4.87E-10
rs1121980	G	A	0.48	-0.02532	0.003165	8.36E-16
rs1167800	A	G	0.45	0.020253	0.003291	2.61E-09
rs1260326	C	T	0.42	0.024051	0.003291	3.84E-14
rs1530559	A	G	0.40	0.017722	0.003291	3.37E-08
rs2745353	T	C	0.45	0.017722	0.003165	5.48E-09
rs2820436	C	A	0.32	0.018987	0.003291	4.36E-09
rs2972143	G	A	0.37	0.017722	0.003291	3.15E-08
rs4865796	A	G	0.29	0.018987	0.003291	2.09E-08
rs731839	A	G	0.34	-0.01772	0.003291	1.72E-08
rs7903146	C	T	0.28	0.022785	0.003544	6.13E-11
rs860598	A	G	0.14	0.022785	0.004051	1.64E-08
rs983309	G	T	0.10	-0.03671	0.00481	3.81E-14
rs9884482	T	C	0.35	-0.02025	0.003038	1.40E-11

SNP: single nucleotide polymorphism; SE: standard deviation. Beta and SE expressed as change in fasting insulin (in SD log pmol/L) per effect allele.

Supplementary table 7a. Association of liver function markers-related SNPs with continuous potential confounders of the exposure-outcome associations meta-analysed across the relevant UCLEB studies and GIANT consortium data

	Potential confounders (Effect size per copy of the liver enzyme raising allele (95% CI))							
SNP/ instrument	Age (years)	BMI (SD)*	Waist circumference (SD)*	Waist-hip ratio (SD)*	ALT (SD)	AST (SD)	GGT (SD)	ALP (SD)
rs11597390	-0.13 (-0.30, 0.03)	0.00 (-0.01, 0.01)	0.01 (0.00, 0.02)	0.00 (-0.00, 0.01)	N/A	0.06 (0.03, 0.10)	0.03 (-0.00, 0.06)	-0.08 (-0.13, -0.02)
rs2143571	-0.04 (-0.26, 0.17)	0.00 (-0.01, 0.01)	0.00 (-0.01, 0.01)	0.00 (-0.01, 0.01)	N/A	0.12 (0.07, 0.17)	0.05 (-0.00, 0.09)	0.03 (-0.04, 0.11)
rs2954021	-0.11(0.27, 0.05)	-0.01(-0.02, -0.00)	-0.01 (-0.01, 0.00)	0.00 (-0.01, 0.01)	N/A	-0.00 (-0.03, 0.03)	0.03 (-0.01, 0.06)	0.08 (0.03, 0.13)
rs6834314	0.04(-0.13, 0.22)	-0.01 (-0.02, -0.01)	-0.01 (-0.02, 0.00)	-0.00 (-0.01, 0.01)	N/A	0.04 (0.00, 0.08)	0.03 (-0.00, 0.07)	-0.03 (-0.09, 0.03)
ALT instrument	-0.07 (-0.16, 0.02)	-0.00 (-0.01, 0.00)	-0.00 (-0.00, 0.00)	0.00 (-0.00, 0.01)	N/A	0.05 (0.03, 0.07)	0.03 (0.01, 0.05)	-0.00 (-0.03, 0.03)
rs17109512	-0.27 (-0.99, 0.39)	0.00 (-0.02, 0.02)	-0.01 (-0.03, 0.01)	-0.00 (-0.03, 0.02)	-0.16 (-0.31, 0.00)	N/A	-0.12 (-0.26, 0.02)	-0.13 (-0.38, 0.13)
rs738407	-0.05 (-0.30, 0.21)	-0.00 (-0.01, 0.00)	-0.00 (-0.01, 0.01)	-0.00 (-0.01, 0.01)	0.13 (0.07, 0.19)	N/A	0.02 (-0.03, 0.07)	0.03 (-0.05, 0.12)
rs738408	-0.04 (-0.27, 0.19)	-0.01 (-0.01, 0.00)	-0.00 (-0.01, 0.01)	0.00 (-0.01, 0.01)	0.13 (0.07, 0.18)	N/A	0.01 (-0.03, 0.06)	0.02 (-0.05, 0.10)
AST instrument	-0.06 (-0.22, 0.11)	-0.00 (-0.01, 0.00)	-0.00 (-0.01, 0.00)	-0.00 (-0.01, 0.01)	0.11 (0.07, 0.15)	N/A	0.01 (-0.03, 0.04)	0.02 (-0.04, 0.07)
rs16856332	0.13 (-0.41, 0.68)	-0.01 (-0.03, 0.00)	-0.02 (-0.04, 0.00)	-0.01 (-0.03, 0.02)	0.04 (-0.09, 0.17)	-0.06 (-0.18, 0.06)	-0.11 (-0.23, -0.00)	N/A
rs174601	-0.04 (-0.20, 0.12)	0.00 (-0.00, 0.01)	0.00 (-0.00, 0.01)	0.00 (-0.00, 0.01)	-0.02 (-0.06, 0.01)	0.00 (-0.03, 0.04)	0.00 (-0.03, 0.04)	N/A
rs1780324	-0.03 (-0.18, 0.13)	-0.00 (-0.01, 0.00)	-0.00 (-0.01, 0.00)	-0.00 (-0.01, 0.00)	-0.01 (-0.05, 0.02)	-0.01 (-0.04, 0.03)	0.01 (-0.03, 0.04)	N/A
rs2236653	0.16 (-0.02, 0.33)	-0.00 (-0.01, 0.01)	-0.00 (-0.01, 0.01)	-0.00 (-0.01, 0.01)	-0.00 (-0.04, 0.04)	-0.00 (-0.04, 0.03)	0.02 (-0.01, 0.06)	N/A
rs281377	0.16 (0.00, 0.33)	0.01 (0.00, 0.02)	0.01 (0.00, 0.02)	-0.01 (-0.01, 0.00)	0.00 (-0.04, 0.04)	0.00 (-0.03, 0.04)	-0.05 (-0.08, -0.01)	N/A
rs2954021	-0.11 (-0.27, 0.05)	-0.01 (-0.02, -0.00)	-0.01 (-0.01, 0.00)	0.00 (-0.01, 0.01)	0.03 (-0.00, 0.07)	-0.00 (-0.03, 0.03)	0.03 (-0.01, 0.06)	N/A
rs314253	0.15 (-0.02, 0.31)	-0.000 (-0.01, 0.01)	-0.00 (-0.01, 0.01)	-0.01 (-0.01, 0.00)	0.01 (-0.03, 0.05)	0.00 (-0.03, 0.04)	0.00 (-0.03, 0.04)	N/A
rs514708	-0.13 (-0.31, 0.05)	0.00 (-0.01, 0.01)	-0.00 (-0.01, 0.00)	-0.01 (-0.02, -0.00)	0.02 (-0.03, 0.06)	0.07 (0.03. 0.11)	0.03 (-0.01, 0.06)	N/A
rs579459	-0.12 (-0.32, 0.07)	0.00 (-0.01, 0.01)	0.00 (-0.00, 0.01)	0.00 (-0.00, 0.01)	0.00 (-0.05, 0.05)	0.05 (0.01, 0.10)	-0.05 (-0.09, -0.01)	N/A
rs6984305	-0.00 (-0.25, 0.25)	0.00 (-0.01, 0.01)	0.01 (-0.00, 0.02)	0.01 (0.00, 0.02)	0.10 (0.04, 0.16)	0.07 (0.01, 0.12)	-0.06 (-0.11, -0.01)	N/A
rs7186908	-0.09 (-0.29, 0.12)	-0.00 (-0.01, 0.01)	-0.01 (-0.02, 0.00)	-0.00 (-0.01, 0.01)	-0.01 (-0.06, 0.04)	-0.00 (-0.05, 0.04)	-0.02 (-0.07, 0.02)	N/A
rs7267979	0.11 (-0.05, 0.26)	-0.01 (-0.01, -0.00)	-0.01 (-0.02, -0.01)	-0.01 (-0.02, -0.01)	-0.03 (-0.07, 0.01)	-0.02 (-0.05, 0.02)	0.00 (-0.03, 0.03)	N/A
rs7923609	-0.01 (-0.17, 0.15)	-0.01 (-0.02, -0.00)	-0.01 (-0.02, -0.00)	-0.02 (-0.02, -0.01)	-0.00 (-0.04, 0.03)	-0.02 (-0.05, 0.01)	-0.00 (-0.03, 0.03)	N/A
rs9467160	0.19 (0.01, 0.37)	-0.01 (-0.02, -0.00)	-0.01 (-0.02, -0.00)	-0.00 (-0.01, 0.00)	-0.03 (-0.07, 0.02)	-0.04 (-0.07, 0.00)	-0.00 (-0.05, 0.02)	N/A
ALP instrument	0.02 (-0.02, 0.07)	-0.00 (-0.00, -0.00)	-0.00 (-0.00, -0.00)	-0.00 (-0.01, -0.00)	-0.00 (-0.01, 0.01)	0.00 (-0.01, 0.01)	-0.00 (-0.01, 0.00)	N/A
rs10513686	-0.02 (-0.26, 0.22)	0.01 (0.00, 0.02)	0.01 (-0.00, 0.02)	0.01 (-0.00, 0.02)	0.03 (-0.03, 0.08)	-0.01 (-0.06, 0.04)	N/A	-0.01 (-0.09, 0.07)
rs1076540	0.02 (-0.17, 0.22)	0.00 (-0.01, 0.01)	-0.00 (-0.01, 0.01)	0.01 (-0.00, 0.02)	0.03 (-0.01, 0.08)	0.02 (-0.02, 0.06)	N/A	0.01 (-0.06, 0.07)
rs10908458	0.01 (-0.16, 0.17)	0.01 (-0.00, 0.01)	-0.00 (-0.01, 0.01)	-0.01 (-0.01, 0.00)	0.01 (-0.02, 0.05)	0.01 (-0.02, 0.05)	N/A	-0.01 (-0.06, 0.04)
rs12145922	-0.04 (-0.20, 0.12)	-0.01 (-0.02, -0.00)	0.00 (-0.01, 0.01)	0.01 (-0.00, 0.01)	-0.02 (-0.06, 0.01)	-0.01 (-0.04, 0.02)	N/A	-0.01 (-0.06, 0.05)
rs1260326	-0.21 (-0.37, -0.05)	-0.01 (-0.02, -0.01)	-0.01 (-0.01, -0.00)	0.00 (-0.00, 0.01)	0.03 (-0.01, 0.07)	0.03 (-0.00, 0.07)	N/A	0.04 (-0.01, 0.10)
rs12968116	SNP not in UCLEB	0.00 (-0.01, 0.01)	0.00 (-0.01, 0.02)	0.01 (-0.01, 0.02)	SNP not in UCLEB	SNP not in UCLEB	N/A	SNP not in UCLEB
rs13030978	-0.07 (-0.33, 0.19)	-0.00 (-0.01, 0.01)	-0.01 (-0.01, 0.00)	-0.00 (-0.01, 0.01)	0.02 (-0.04, 0.08)	0.04 (-0.02, 0.09)	N/A	-0.09 (-0.17, -0.00)
rs1335645	SNP not in UCLEB	0.01 (-0.00, 0.02)	0.01 (-0.01, 0.02)	0.01 (-0.01, 0.02)	SNP not in UCLEB	SNP not in UCLEB	N/A	SNP not in UCLEB
rs1497406	0.08 (-0.12, 0.27)	-0.01 (-0.01, 0.00)	-0.00 (-0.01, 0.01)	0.00 (-0.01, 0.01)	0.01 (-0.03, 0.06)	0.02 (-0.02, 0.06)	N/A	-0.01 (-0.08, 0.05)
rs17145750	0.04 (-0.17, 0.25)	-0.01 (-0.02, -0.00)	-0.00 (-0.01, 0.01)	0.01 (0.00, 0.02)	-0.06 (-0.11, -0.01)	-0.04 (-0.09, 0.00)	N/A	0.01 (-0.06, 0.08)
rs2140773	-0.08 (-0.25, 0.09)	0.00 (-0.00, 0.01)	0.00 (-0.01, 0.01)	0.01 (-0.00, 0.01)	-0.00 (-0.04, 0.04)	-0.02 (-0.05, 0.02)	N/A	-0.03 (-0.08, 0.02)
rs2739330	0.09 (-0.10, 0.28)	-0.00 (-0.01, 0.01)	-0.00 (-0.01, 0.00)	0.00 (-0.01, 0.01)	-0.02 (-0.06, 0.03)	-0.00 (-0.04, 0.04)	N/A	0.04 (-0.02, 0.10)
rs339969	-0.13 (-0.31, 0.06)	0.01 (0.01, 0.02)	0.01 (-0.002, 0.02)	-0.00 (-0.01, 0.01)	0.02 (-0.02, 0.06)	0.02 (-0.02, 0.06)	N/A	-0.01 (-0.07, 0.05)
SNP/	Age (years)	BMI (SD)	Waist	Waist-hip ratio (SD)	ALT (SD)	AST (SD)	GGT (SD)	ALP (SD)

instrument			circumference (SD)					
rs4074793	-0.04 (-0.34, 0.26)	0.00 (-0.01, 0.02)	0.01 (-0.01, 0.02)	0.00 (-0.01, 0.02)	0.07 (-0.00, 0.14)	0.04 (-0.03, 0.10)	N/A	0.11 (0.01, 0.21)
rs4503880	SNP not in UCLEB	-0.00 (-0.01, 0.01)	-0.00 (-0.02, 0.01)	-0.00 (-0.02, 0.01)	SNP not in UCLEB	SNP not in UCLEB	N/A	SNP not in UCLEB
rs4547811	0.12 (-0.13, 0.37)	0.00 (-0.01, 0.01)	0.00 (-0.01, 0.01)	-0.00 (-0.01, 0.01)	0.06 (0.00, 0.12)	0.04 (-0.01, 0.09)	N/A	0.04 (-0.04, 0.13)
rs4581712	-0.12 (-0.57, 0.32)	0.00 (-0.00, 0.01)	0.00 (-0.01, 0.01)	-0.00 (-0.01, 0.01)	0.02 (-0.07, 0.11)	-0.00 (-0.10, 0.09)	N/A	0.03 (-0.11, 0.18)
rs4820599	-0.09 (-0.27, 0.08)	-0.003 (-0.01, 0.00)	-0.00 (-0.01, 0.00)	-0.00 (-0.01, 0.00)	-0.03 (-0.07, 0.01)	-0.02 (-0.06, 0.02)	N/A	-0.00 (-0.06, 0.06)
rs516246	-0.16 (-0.31, -0.00)	-0.01 (-0.01, -0.00)	-0.01 (-0.01, 0.00)	0.01 (-0.02, 0.01)	-0.00 (-0.04, 0.04)	0.00 (-0.03, 0.04)	N/A	-0.06 (-0.11, -0.01)
rs6888304	SNP not in UCLEB	-0.01 (-0.01, 0.00)	-0.01 (-0.02, 0.00)	-0.01 (-0.02, 0.00)	SNP not in UCLEB	SNP not in UCLEB	N/A	SNP not in UCLEB
rs7310409	0.00 (-0.16, 0.17)	0.00 (-0.00, 0.01)	0.00 (-0.00, 0.01)	0.01 (0.00, 0.01)	-0.01 (-0.04, 0.03)	-0.00 (-0.03, 0.03)	N/A	-0.07 (-0.13, -0.02)
rs754466	SNP not in UCLEB	-0.00 (-0.01, 0.01)	-0.00 (-0.01, 0.05)	0.00 (-0.01, 0.01)	SNP not in UCLEB	SNP not in UCLEB	N/A	SNP not in UCLEB
rs8038465	0.11 (-0.12, 0.34)	-0.00 (-0.01, 0.01)	-0.01 (-0.01, 0.00)	-0.00 (-0.01, 0.01)	0.06 (0.00, 0.11)	0.03 (-0.02, 0.08)	N/A	-0.03 (-0.11, 0.04)
rs9296736	0.09 (-0.08, 0.26)	0.01 (0.00, 0.01)	0.01 (-0.00, 0.01)	0.00 (-0.01, 0.01)	0.03 (-0.01, 0.07)	-0.01 (-0.04, 0.03)	N/A	-0.00 (-0.06, 0.05)
rs944002	-0.24 (-0.42, -0.06)	-0.00 (-0.01, 0.01)	-0.00 (-0.02, 0.01)	-0.00 (-0.01, 0.01)	-0.01 (-0.05, 0.04)	-0.01 (-0.05, 0.03)	N/A	-0.01 (-0.07, 0.05)
rs9913711	SNP not in UCLEB	-0.00 (-0.01, 0.01)	-0.01 (-0.01, 0.00)	-0.01 (-0.02, 0.00)	SNP not in UCLEB	SNP not in UCLEB	N/A	SNP not in UCLEB
GGT instrument	-0.04 (-0.08, 0.00)	-0.00 (-0.00, 0.00)	-0.00 (-0.00, 0.00)	0.00 (-0.00, 0.00)	0.00 (-0.00, 0.02)	0.00 (-0.00, 0.01)	N/A	-0.01 (-0.02, 0.00)

ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; GGT: gamma-glutamyl transferase; BMI: body mass index; N/A: not applicable; SD: standard deviation; SNPs: single nucleotide polymorphisms. GIANT: The Genetic Investigation of Anthropometric Traits. *Summary-level data publicly available from the GIANT consortium was used for these traits in addition to the data from the relevant UCLEB studies.

Supplementary table 7b. Association of liver function markers-related SNPs with potentially confounding dichotomous traits meta-analysed across the relevant UCLEB studies.

	Potential confounders (Odds ratio per copy of the liver function marker raising allele (95% CI))			
SNP/ instrument	Sex	Smoking	Alcohol consumption	Social class
rs11597390	0.98 (0.91, 1.06)	0.97 (0.92, 1.02)	1.08 (0.98, 1.18)	1.06 (1.00, 1.14)
rs2143571	0.99 (0.89, 1.09)	1.01 (0.93, 1.09)	1.09 (0.96, 1.25)	1.08 (0.99, 1.18)
rs2954021	0.98 (0.91, 1.05)	1.03 (0.98, 1.09)	1.00 (0.92, 1.10)	1.00 (0.94, 1.06)
rs6834314	0.99 (0.92, 1.07)	0.99 (0.93, 1.05)	0.95 (0.86, 1.06)	1.01 (0.94, 1.08)
ALT instrument	0.98 (0.95, 1.02)	1.00 (0.97, 1.03)	1.02 (0.97, 1.08)	1.03 (1.00, 1.07)
rs17109512	1.09 (0.79, 1.52)	0.87 (0.68, 1.11)	1.37 (0.94, 1.99)	0.83 (0.63, 1.10)
rs738407	1.07 (0.95, 1.20)	1.00 (0.92, 1.10)	1.01 (0.87, 1.17)	1.02 (0.92, 1.13)
rs738408	1.01 (0.92, 1.12)	0.99 (0.91, 1.07)	1.03 (0.90, 1.17)	1.06 (0.97, 1.16)
AST instrument	1.04 (0.97, 1.12)	0.99 (0.93, 1.05)	1.04 (0.95, 1.14)	1.03 (0.96, 1.10)
rs10819937	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB
rs16856332	0.99 (0.76, 1.29)	0.95 (0.78, 1.15)	1.07 (0.78, 1.47)	1.07 (0.85, 1.33)
rs174601	1.03 (0.96, 1.11)	0.98 (0.93, 1.04)	0.95 (0.86, 1.04)	1.00 (0.94, 1.07)
rs1780324	0.97 (0.90, 1.04)	0.93 (0.88, 0.99)	0.97 (0.89, 1.07)	1.04 (0.98, 1.11)
rs2236653	0.90 (0.84, 0.98)	0.98 (0.92, 1.04)	0.95 (0.86, 1.04)	0.98 (0.91, 1.05)
rs281377	1.06 (0.99, 1.14)	0.95 (0.90, 1.00)	0.96 (0.87, 1.05)	1.04 (0.98, 1.11)
rs2954021	0.98 (0.91, 1.05)	1.03 (0.98, 1.09)	1.00 (0.92, 1.10)	1.00 (0.94, 1.06)
rs314253	0.96 (0.90, 1.04)	0.99 (0.93, 1.04)	1.04 (0.95, 1.15)	1.01 (0.94, 1.07)
rs514708	0.98 (0.90, 1.06)	0.96 (0.91, 1.02)	1.09 (0.98, 1.21)	1.03 (0.95, 1.10)
rs579459	1.01 (0.93, 1.11)	0.99 (0.92, 1.06)	1.03 (0.92, 1.15)	1.05 (0.97, 1.13)
rs6984305	1.03 (0.92, 1.16)	0.97 (0.89, 1.05)	1.04 (0.90, 1.21)	1.03 (0.93, 1.14)
rs7186908	1.05 (0.95, 1.15)	1.05 (0.98, 1.12)	1.02 (0.91, 1.15)	1.06 (0.98, 1.15)
rs7267979	1.05 (0.97, 1.12)	1.01 (0.96, 1.07)	1.01 (0.92, 1.10)	1.03 (0.96, 1.09)
rs7923609	1.00 (0.93, 1.08)	1.05 (0.99, 1.11)	1.01 (0.92, 1.11)	1.04 (0.97, 1.11)
rs9467160	0.99 (0.91, 1.07)	0.98 (0.92, 1.04)	1.01 (0.91, 1.12)	1.05 (0.98, 1.13)
ALP instrument	1.00 (0.98, 1.02)	0.99 (0.97, 1.01)	1.00 (0.98, 1.03)	1.03 (1.01, 1.05)
rs10513686	0.93 (0.84, 1.04)	0.98 (0.90, 1.06)	0.99 (0.86, 1.14)	0.99 (0.90, 1.09)
rs1076540	0.93 (0.86, 1.02)	1.02 (0.96, 1.09)	1.08 (0.97, 1.20)	0.96 (0.88, 1.03)
rs10908458	0.96 (0.89, 1.04)	0.99 (0.94, 1.05)	1.01 (0.92, 1.02)	1.02 (0.96, 1.09)
rs12145922	1.04 (0.97, 1.12)	0.95 (0.90, 1.01)	0.97 (0.89, 1.07)	1.00 (0.93, 1.06)
rs1260326	0.98 (0.91, 1.05)	0.98 (0.93, 1.04)	0.91(0.83, 1,00)	0.98 (0.92, 1.04)
rs12968116	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB
rs13030978	0.92 (0.82, 1.04)	1.08 (0.99, 1.19)	0.87 (0.75, 1.01)	0.94 (0.84, 1.04)
rs1335645	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB
rs1497406	1.04 (0.95, 1.14)	1.01 (0.94, 1.08)	0.96 (0.85, 1.07)	1.01 (0.94, 1.10)
rs17145750	1.00 (0.91, 1.10)	0.97 (0.90, 1.04)	0.89 (0.78, 1.01)	0.98 (0.90, 1.06)
rs2140773	1.05 (0.97, 1.13)	0.96 (0.91, 1.02)	1.01 (0.92, 1.11)	1.05 (0.98, 1.12)

rs2739330	1.08 (0.99, 1.18)	1.01 (0.94, 1.07)	0.88 (0.78, 0.98)	0.98 (0.91, 1.06)
rs339969	0.97 (0.89, 1.05)	1.03 (0.96, 1.09)	0.96 (0.86, 1.07)	1.01 (0.94, 1.09)
rs4074793	0.94 (0.82, 1.09)	1.02 (0.92, 1.13)	1.02 (0.86, 1.22)	0.90 (0.80, 1.02)
rs4503880	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB
rs4547811	1.06 (0.94, 1.19)	0.96 (0.88, 1.05)	0.97 (0.84, 1.12)	1.03 (0.93, 1.14)
rs4581712	0.91 (0.78, 1.07)	1.02 (0.89, 1.17)	1.12 (0.88, 1.44)	0.89 (0.76, 1.05)
rs4820599	0.97 (0.89, 1.05)	1.03 (0.97, 1.09)	1.00 (0.91, 1.11)	1.01 (0.94, 1.08)
rs516246	0.95 (0.88, 1.02)	1.05 (0.99, 1.11)	1.06 (0.97, 1.16)	0.96 (0.90, 1.02)
rs6888304	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB
rs7310409	0.97 (0.90, 1.04)	0.97 (0.92, 1.03)	1.04 (0.95, 1.14)	1.05 (0.98, 1.12)
rs754466	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB
rs8038465	0.97 (0.88, 1.08)	1.02 (0.94, 1.10)	1.08 (0.95, 1.24)	1.03 (0.93, 1.13)
rs9296736	1.04 (0.96, 1.13)	0.98 (0.92, 1.04)	0.95 (0.86, 1.04)	1.05 (0.98, 1.12)
rs944002	0.98 (0.90, 1.07)	1.05 (0.99, 1.12)	1.04 (0.94, 1.16)	1.00 (0.93, 1.08)
rs9913711	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB	SNP not in UCLEB
GGT instrument	0.99 (0.97, 1.01)	1.00 (0.99, 1.02)	0.99 (0.96, 1.01)	1.00 (0.98, 1.02)

ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; GGT: gamma-glutamyl transferase; SNPs: single nucleotide polymorphisms; UCLEB consortium: UCL-LSHTM-Edinburgh-Bristol consortium

Supplementary table 8. Fully adjusted multivariable regression and Mendelian Randomization associations / effects of liver function markers with continuous outcomes meta-analysed across UCLEB, Fenland* and GWAS studies

	Mean difference (in S.D units) for each continuous outcome per one S.D increase in each liver function marker (95% CI)					
	Glucose (mmol/L)	Insulin (μIU/ml)	HDL-C (mmol/L)	LDL-C (mmol/L)	Total cholesterol (mmol/L)	Triglycerides (mmol/L)
ALT						
Multivariable	0.08 (-0.03, 0.19)	0.19 (0.11, 0.28)	-0.01 (-0.04, 0.02)	0.03 (-0.03, 0.08)	0.06 (0, 0.11)	0.1 (0.05, 0.15)
MR-IVW	-0.01 (-0.09, 0.07)	-0.01 (-0.09, 0.06)	-0.09 (-0.35, 0.16)	0.08 (-0.33, 0.48)	0.05 (-0.39, 0.48)	0.07 (-0.45, 0.59)
MR-Egger	0.12 (-0.11, 0.34)	0.08 (-0.13, 0.29)	0.62 (0.16, 1.09)	-0.75 (-1.54, 0.04)	-0.8 (-1.7, 0.11)	-1.07 (-2.16, 0.02)
MR-Weighted median	-0.02 (-0.17, 0.12)	0 (-0.08, 0.08)	-0.09 (-0.25, 0.07)	-0.1 (-0.2, 0.01)	-0.11 (-0.21, -0.01)	-0.06 (-0.16, 0.04)
AST						
Multivariable	0 (-0.04, 0.04)	0.05 (0.01, 0.09)	0.06 (0.04, 0.09)	-0.02 (-0.05, 0.01)	0.01 (-0.02, 0.03)	0.02 (-0.04, 0.07)
MR-IVW	-0.01 (-0.07, 0.04)	0.01 (-0.04, 0.06)	-0.09 (-0.15, -0.04)	-0.07 (-0.13, -0.01)	-0.11 (-0.17, -0.04)	0.03 (-0.03, 0.08)
MR-Egger	0.01 (-0.18, 0.2)	0.08 (-0.03, 0.18)	-0.06 (-0.19, 0.07)	-0.17 (-0.31, -0.02)	-0.21 (-0.34, -0.07)	-0.03 (-0.16, 0.09)
MR-Weighted median	-0.01 (-0.06, 0.04)	0.01 (-0.04, 0.06)	-0.09 (-0.15, -0.03)	-0.07 (-0.13, -0.01)	-0.11 (-0.17, -0.04)	0.02 (-0.03, 0.08)
ALP						
Multivariable	0.01 (-0.04, 0.06)	0.02 (-0.03, 0.08)	-0.04 (-0.08, -0.01)	0.02 (-0.04, 0.07)	0.02 (-0.04, 0.07)	0.05 (0, 0.1)
MR-IVW	-0.03 (-0.08, 0.03)	-0.01 (-0.04, 0.01)	-0.06 (-0.13, 0.02)	-0.17 (-0.28, -0.06)	-0.16 (-0.28, -0.04)	0.08 (-0.07, 0.24)
MR-Egger	-0.06 (-0.22, 0.1)	0 (-0.05, 0.04)	-0.07 (-0.35, 0.21)	-0.27 (-0.46, -0.07)	-0.25 (-0.45, -0.05)	0.03 (-0.28, 0.34)
MR-Weighted median	-0.03 (-0.06, 0)	-0.01 (-0.04, 0.02)	-0.03 (-0.12, 0.06)	-0.2 (-0.34, -0.06)	-0.19 (-0.31, -0.08)	0.03 (-0.05, 0.12)
GGT						
Multivariable	0.07 (0.02, 0.11)	0.1 (0.04, 0.16)	0.04 (0.01, 0.07)	0.05 (0, 0.1)	0.11 (0.06, 0.16)	0.15 (0.1, 0.2)
MR-IVW	-0.05 (-0.12, 0.03)	-0.02 (-0.07, 0.02)	-0.04 (-0.08, 0)	-0.03 (-0.1, 0.03)	0.02 (-0.07, 0.11)	0.16 (0.02, 0.3)
MR-Egger	0.07 (-0.11, 0.24)	0.1 (-0.01, 0.22)	-0.07 (-0.17, 0.03)	-0.11 (-0.38, 0.16)	-0.17 (-0.37, 0.02)	-0.14 (-0.44, 0.17)
MR-Weighted median	0 (-0.11, 0.11)	-0.01 (-0.05, 0.03)	-0.03 (-0.07, 0.01)	-0.06 (-0.15, 0.03)	-0.04 (-0.11, 0.04)	0 (-0.04, 0.04)

*Fenland study was only included in the instrumental variables analyses of the associations between the exposures ALT, ALP, GGT and lipid outcomes. IVW: inverse-variance weighted method. ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; GGT: gamma-glutamyl transferase; UCLEB consortium: UCL-LSHTM-Edinburgh-Bristol consortium; GWAS: Genome-wide association study; MR: Mendelian randomization; IVW: Inverse variance weighted estimator; S.D: standard deviation; HDL-C: high density lipoprotein-cholesterol; LDL-C: low density lipoprotein-cholesterol

Supplementary table 9. MR-Egger intercept values (pooled across UCLEB, Fenland and GWAS studies)

Liver function marker	Mean difference (S.D) of each continuous outcome per one S.D increase in the liver function marker (95% CI)					
	Glucose (mmol/L)	Insulin (μIU/ml)	HDL-C (mmol/L)	LDL-C (mmol/L)	Total cholesterol (mmol/L)	Triglycerides (mmol/L)
ALT						
MR-Egger intercept	-0.01 (-0.02, 0.01)	0 (-0.01, 0.01)	-0.04 (-0.06, -0.01)	0.05 (0.02, 0.09)	0.06 (0.02, 0.1)	0.06 (0.01, 0.12)
AST						
MR-Egger intercept	0.03 (-0.01, 0.07)	0.01 (-0.02, 0.04)	0 (-0.02, 0.03)	0.01 (-0.05, 0.07)	0.01 (-0.06, 0.08)	0.01 (-0.02, 0.03)
ALP						
MR-Egger intercept	0.01 (-0.01, 0.02)	0 (-0.01, 0.01)	0 (-0.03, 0.02)	0.01 (-0.01, 0.02)	0.01 (-0.01, 0.02)	0.01 (-0.01, 0.02)
GGT						
MR-Egger intercept	-0.01 (-0.03, 0)	-0.01 (-0.02, 0)	0 (-0.01, 0.01)	0.01 (-0.02, 0.03)	0.01 (0, 0.03)	0.02 (0, 0.04)

ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; GGT: gamma-glutamyl transferase; HDL-C: high density lipoprotein-cholesterol; LDL-C: low density lipoprotein-cholesterol; UCLEB consortium: UCL-LSHTM-Edinburgh-Bristol consortium; GWAS: Genome-wide association study; MR: Mendelian randomization; IVW: Inverse variance weighted estimator; S.D: standard deviation.

References

1. Wadsworth M, Kuh D, Richards M, Hardy R. Cohort Profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development). Int J Epidemiol. 2006;35(1):49-54.
2. Rantakallio P. The longitudinal study of the northern Finland birth cohort of 1966. Paediatr Perinat Epidemiol. 1988;2(1):59-88.