Distinct metabolic features of genetic liability to type 2 diabetes and coronary artery disease: a reverse Mendelian randomization study

Madeleine L Smith[1,2], Caroline J Bull[1,2,3], Michael V Holmes[1,2,4,5,6], George Davey Smith[1,2], Emma L Anderson[1,2], Joshua A Bell[1,2]

[1] Medical Research Council, Integrative Epidemiology Unit, University of Bristol, UK

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

[3] School of Translational Health Sciences, University of Bristol, Bristol, UK

[4] Clinical Trial Service Unit & Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

[5] MRC Population Health Research Unit at the University of Oxford, Oxford, UK

[6] National Institute for Health Research, Oxford Biomedical Research Centre, Oxford University Hospital, Oxford, UK

Corresponding author: Madeleine L Smith, Oakfield House, University of Bristol, UK, BS8 2BN, email: madeleine.smith@bristol.ac.uk

Abstract words: 355; **Main text words**: 5620; **References**: 74; **Tables & Figures**: 6; **Supplementary items**: 7 tables, 13 figures.

Keywords: Type 2 diabetes; Coronary artery disease; Genetics; Mendelian randomization; Metabolism; NMR; Epidemiology; UK Biobank

ABSTRACT

Background

Type 2 diabetes (T2D) and coronary artery disease (CAD) both have roots in disordered metabolism and known genetic determinants. The mechanisms through which their associated genetic variants lead to disease onset remain poorly understood. Here, we used large-scale metabolomics data to directly compare the metabolic features of genetic liability to T2D and to CAD.

Methods

We performed two-sample reverse Mendelian randomization (MR) to estimate effects of genetic liability to T2D and CAD on 249 circulating metabolites from targeted nuclear magnetic resonance spectroscopy in the UK Biobank (N=118,466) using 167 and 145 genetic instruments, respectively. We examined the potential for medication use to distort effect estimates by examining effects of disease liability on metformin and statin use and by conducting age-stratified metabolite analyses.

Results

Using inverse variance weighted (IVW) models, higher genetic liability to T2D was estimated to decrease high-density lipoprotein cholesterol (HDL-C) and low-density lipoprotein cholesterol (LDL-C) (e.g., HDL-C: -0.05 SD; 95% CI -0.07, -0.03, per doubling of liability), whilst increasing all triglyceride groups and branched chain amino acids (BCAAs). Estimates for CAD liability suggested an effect on reducing HDL-C as well as raising very-low density lipoprotein cholesterol (VLDL-C) and LDL-C, and LDL triglycerides. Liability to each disease was estimated to decrease apolipoprotein-A1, whilst only CAD liability was estimated with IVW to increase apolipoprotein-B (0.10 SD; 95% CI 0.03, 0.17). In pleiotropy-robust sensitivity models, T2D liability was still estimated to increase BCAAs, but several effect estimates for higher CAD liability reversed and supported decreased LDL-C and apolipoprotein-B. Moreover, higher T2D liability uniquely increased the odds of using metformin, whereas higher CAD liability uniquely

increased the odds of using statins. Estimated effects of CAD liability differed uniquely and substantially by age for non-HDL-C traits in particular, with, e.g., pleiotropy-robust models suggesting that higher CAD liability lowers LDL-C only at older ages when use of statins is common.

Conclusions

Our results support largely distinct metabolic features of genetic liability to T2D and to CAD, particularly higher BCAAs in T2D and lower LDL-C and apolipoprotein-B in CAD. Such apparently protective effects of CAD liability differ substantially by age and likely reflect mediation by statin use in adulthood.

INTRODUCTION

Globally, overall incidence rates of type 2 diabetes (T2D) and coronary artery disease (CAD) are increasing in parallel, together affecting over 500 million adults.[1,2] Both diseases have roots in disordered metabolism, but they differ in their life course development and clinical presentation, and have distinct causal architectures. T2D is characterised by insulin resistance and is identified clinically via high blood glucose, whilst CAD is characterised by atherosclerotic plaques and is often identified clinically via angina, myocardial infarction, or ischaemic cardiomyopathy. Genome-wide association studies (GWAS) have provided robust evidence that both T2D and CAD have genetic influences,[3,4] but the mechanisms through which the associated genetic variants increase disease risk remain poorly understood.

Recent advances in metabolomics technologies involve the simultaneous quantification of hundreds of molecular metabolites from a single biological sample, enabling detailed insight into the metabolic features of cardiometabolic diseases and their potential underlying mechanisms.[5] The application of metabolomics in observational epidemiological studies has helped to identify metabolites involved in the pathogenesis of T2D, or that might contribute to T2D incidence.[6–16] For example, a recent meta-analysis identified associations of incident T2D with differences in several lipid and non-lipid metabolites, including higher levels of branched chain amino acids (BCAAs), some fatty acids and acylcarnitines, amongst others.[17] For CAD, metabolomics in observational studies has helped to identify key metabolic characteristics of the disease,[18–25] to predict and diagnose CAD,[26–29] and to predict outcomes and mortality in CAD patients.[30–32] However, observational studies are limited in their ability to infer whether metabolic traits are causal, rather than just predictive for disease, as these studies are liable to bias due to residual confounding by lifestyle factors and other diseases.

Mendelian randomization (MR) is a genetic epidemiological approach that uses genetic variants reliably associated with modifiable exposures to estimate the unbiased effects of such exposures on outcomes.[33] This technique is generally implemented within an instrumental variables framework,[34] and has been previously applied in studies investigating associations between lipid levels and T2D and CAD [35–38], with the two diseases appearing to have differing lipid aetiology: higher low-density lipoprotein cholesterol (LDL-C) and triglycerides increase CAD risk, whereas higher LDL and high-density lipoprotein cholesterol (HDL-C) appear to reduce T2D risk.[37] More recently, MR has been used in 'reverse' to estimate effects of genetic liability for disease and potential biomarkers, to reveal features of the developing disease process, which include both causal and non-causal factors for disease incidence.[39] Using reverse MR, higher genetic liability to T2D diagnosed in adulthood has been estimated to decrease total lipid content in HDL particles, as well as increase BCAA and inflammatory glycoprotein acetyl levels, as early as 8 years of age.[40] Higher genetic liability to CAD has also been estimated to affect circulating metabolite levels in early life, especially higher very-low-density lipoprotein (VLDL) and total lipids in LDL [41]. These MR studies in children have the advantage of low disease prevalence (and therefore little reverse causality), and no medication use distorting effect estimates, so are useful for identifying causal factors. Studies in older adults are better for identifying predictors of disease, as there is a greater burden of prevalent disease, however increased medication use with age has the potential to distort results. Previous reverse MR studies of the impact of genetic liability to T2D or CAD on metabolic traits are limited by their small sample sizes; cohort studies of ~5,000 participants or GWAS of less than ~25,000 adults. Moreover, the metabolic profiles of genetic liability to T2D and CAD have not been directly compared within the same analytical setting, and so the ways in which these profiles differ have not yet been revealed.

Here, we compare the metabolic features of genetic liability to T2D and to CAD using a reverse MR framework, with the aim of revealing the pattern of metabolic traits which characterise the development of each disease. We used the largest, most recent summary-level GWAS data on circulating metabolites measured using targeted metabolomics (N~120,000) from the population-based UK Biobank cohort study, representing a 5-fold larger sample size than previous studies. To examine the potential for effect estimates to be distorted by medication use, we additionally examined the effects of T2D and CAD liability on the use of metformin and statins, and repeated analyses using data from age stratified GWAS of metabolites, given that medication use differs markedly by age. For comparison with metabolites, we also estimated the effects of disease liability on more distal lifestyle-related factors, including adiposity, smoking, and alcohol consumption.

METHODS

Genetic instruments for T2D liability and CAD liability

T2D liability

We identified single nucleotide polymorphisms (SNPs) that were independently associated with T2D at $P < 5 \times 10^{-8}$ from a large-scale GWAS meta-analysis combining data from 32 GWAS, including 74,124 T2D cases and 824,006 controls of European ancestry.[3] Across included studies, female sex ranged from 37 to 64% and mean (SD) age range was 23.9 (2.1) to 61.3 (2.9) years. T2D was diagnosed across included studies using a range of criteria, including use of oral diabetes medication, diagnostic fasting glucose or HbA1c levels, or self-report. The T2D GWAS was conducted with and without adjustment for BMI, and we used the 231 SNPs identified in the non-BMI-adjusted GWAS to reduce potential for collider bias.[42] For variants that were in linkage disequilibrium (LD) based on $R^2 > 0.001$, the SNP with the lowest P-value in each set of variants in LD was retained. This left 167 independent SNPs associated with T2D

(**Supplementary Table 1**) for inclusion in our reverse MR analysis. These summary level GWAS results are publicly available from the DIAGRAM consortium website, accessible at https://diagram-consortium.org/downloads.html.

CAD liability

SNPs for CAD were identified in a large-scale GWAS meta-analysis of 10 GWAS conducted among middle-aged adults, with study-specific sample and variant filters applied and adjustment for study-specific covariates. The study had a total sample of 181,522 cases among 1,165,690 study participants, who were 46% female and largely of European ancestry (>95%). Potential ancestry effects were accounted for by transethnic comparison with a GWAS from Biobank Japan. Case status was defined based on relevant hospital codes for prevalent or incident CAD. The GWAS detected 241 genetic variants independently associated with CAD at $P < 5 \times 10^{-8}$. For variants that were in LD based on $R^2 > 0.001$, the SNP with the lowest P-value in each set of variants in LD was retained, leaving 145 SNPs (**Supplementary Table 1**). These summary level GWAS results are publicly available from medrxiv[43] and will be made available on publication through the CARDIoGRAMplusC4D website (http://www.cardiogramplusc4d.org/) and the NHGRI-EBI GWAS Catalog (https://www.ebi.ac.uk/gwas/).

Outcome data on circulating metabolites and medication use

Summary statistics from GWAS of circulating metabolites conducted among European participants of the UK Biobank study were used. Details of the UK Biobank design, participants, genomic quality control (QC) and its strengths and limitations have been reported previously.[44–46] Briefly, 502,549 adults aged 40-69 years were recruited between 2006-2010 via 22 assessment centres across England, Wales, and Scotland. Non-fasting EDTA plasma samples from a random subset of participants (N = 118,466) were analysed for levels of 249 metabolic traits (165 concentrations and 84 derived ratios) using targeted high-throughput

proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy (Nightingale Health Ltd; biomarker quantification version 2020).[47] These metabolic traits comprised routine lipids, lipoprotein subclass profiling, fatty acid composition and various low-molecular-weight metabolites including BCAAs, ketone bodies, glycolysis-related traits and inflammatory glycoprotein acetyls. All metabolic trait measurements were normalized and standardised prior to analyses using rank-based inverse normal transformation. Metabolite GWAS were adjusted for genotype array, sex, and fasting time. We also included C-reactive protein (CRP) for comparison with glycoprotein acetyls, measured in UK Biobank via the same blood samples, analyzed by immunoturbidimetric-high sensitivity analysis on a Beckman Coulter AU5800 (N = 469,772).

Statins and metformin are commonly prescribed medications for the prevention/treatment of CAD and T2D, respectively, particularly among adults identified as having higher risk.[48] These medications can alter the levels of circulating metabolites [49,50] and thus may distort associations of genetic liability to T2D and CAD with metabolites (via mediation). We therefore additionally estimated the effects of T2D and CAD liability on the self-reported use of statins (atorvastatin and simvastatin) and metformin obtained from a sample of 462,933 UK Biobank participants using interviews, self-reporting and medical records (UK Biobank Data-Field 20003), to inform the interpretation of metabolite results. These medication outcomes also served as positive controls, with CAD expected to most strongly raise statin use, and T2D expected to most strongly raise metformin use.

For the purposes of further examining distortions by medication use, we used metabolite GWAS on the same sample which were divided into age tertiles (containing ages: 39-53 years, 53-61 years, and 61-73 years), and which used the same standardisations and covariate adjustments. Based on previous work in UK Biobank,[51] the estimated prevalence of statin medication use within these same tertiles (youngest to oldest) is 5%, 17%, and 29%, respectively. All summary-

level GWAS results are publicly available through the IEU-OpenGWAS platform, accessible at https://gwas.mrcieu.ac.uk/.

Additional outcome data on adiposity, smoking, and alcohol consumption

Summary statistics from GWAS of BMI, waist circumference, trunk fat mass, whole body fat mass, body fat percentage, smoking status, alcohol drinker status, smoking pack-years and alcohol frequency, conducted among European participants of the UK Biobank study were used. Genetic association data were previously generated for these traits using the MRC IEU UK Biobank GWAS pipeline[52], adjusting for age at baseline, sex, and genotyping array.

Standing height was measured without shoes using a Seca 202 device and weight was measured with heavy outer clothing removed, using the Tanita BC418MA body composition analyser (bioelectrical impedance). This same device was used to estimate trunk fat mass in kilograms, whole body fat mass in kilograms, and body fat percentage. Body mass index (BMI) was calculated from dividing weight in kilograms by the square of height in metres. Waist circumference (WC) was measured in centimetres at the umbilicus using a non-stretchable Wessex tape measure.

Alcohol drinking status (never, previous or current) was measured (N=336,965) via a questionnaire. Alcohol intake frequency was measured (N=462,346) via a questionnaire where they were asked “About how often do you drink alcohol?” and answered from the following options: “Daily or almost daily”, “Three or four times a week”, “Once or twice a week”, “One to three times a month”, “Special occasions only”, “Never” or “Prefer not to answer”. Smoking status (never, previous, current) was also measured via questionnaire (N=336,024). Smoking pack years was calculated for 142,387 participants that have ever smoked according to their smoking status response. The general definition of a pack year is the number of cigarettes

smoked per day, divided by twenty, multiplied by the number of years of smoking. Number of years of smoking was calculated by subtracting the age the participant started smoking from the age they stopped.

Statistical approach

To generate MR estimates of the effect of genetic liability to T2D and CAD on metabolites (with ages combined and within age tertiles), and on medication use and additional outcomes noted above, we integrated estimates of the association of SNPs with exposures (T2D and CAD separately), with estimates of the association of those same SNPs with each metabolic trait measured in UK Biobank. Where possible, when exposure SNPs were not available in the outcome GWAS, proxy SNPs were identified using 1000 genomes sample data as SNPs in LD (within 10,000kb, r^2=0.8) with the requested SNP. The number of proxy SNPs recruited differed for each outcome (maximum number of proxies for any one outcome = 8). Summary statistics were harmonised using the 'harmonise_data' function within the TwoSampleMR package.[53] Using the same package, we used inverse variance weighted (IVW) regression in the main analysis, which assumes that none of the SNPs are pleiotropic.[54] To examine and correct for potential horizontal pleiotropy, we used 3 additional sensitivity methods for each exposure-outcome pair: MR-Egger (which allows for all SNPs to be pleiotropic),[55] weighted median (which allows for up to half of the SNPs to be pleiotropic and is less influenced by outliers than other models)[54], and the weighted mode (which assumes that the most common effect is consistent with the true causal effect, i.e. not biased by horizontal pleiotropy).[56] To aid interpretation, all estimates and standard errors were multiplied by 0.693 ($\log_e 2$) and represent the normalised SD unit difference in each outcome trait (metabolites, medications, adiposity, smoking, and alcohol) per doubling of genetic liability to T2D or CAD, based on previous recommendations[57]. Statistical analyses were performed in R version 4.0.2.

RESULTS

Instrument count and strength

167 independent genetic variants were associated with T2D at $P < 5 \times 10^{-8}$, 159 of which were present in the outcome GWAS summary statistics and were used in MR analyses. For the medication use and other outcomes, between 155 and 159 SNPs were available for each and used for MR analyses. One hundred and forty-five genetic variants were associated with CAD at $P < 5 \times 10^{-8}$, 140 of which were available for each metabolic trait and thus used to instrument liability to CAD. For the medication use and other outcomes, 136-141 SNPs were available for each. The F statistics for the genetic instruments for T2D and CAD were 66.8 and 59.7, respectively, indicating instrument strength above the recommended minimum level of 10 to avoid weak instrument bias.[58]

Lipids and lipoproteins

IVW estimates for the effect of higher liability to T2D and to CAD on lipid metabolites tended to differ between the two diseases (**Figure 1**). Higher liability to T2D was estimated to decrease total, HDL-C and LDL-C (e.g., HDL-C -0.05 SD; 95% CI -0.07, -0.03), with no strong evidence of an effect on VLDL-C. IVW estimates for CAD liability on cholesterol were less precise but suggested an effect on lowering HDL-C and an effect on raising total-C, VLDL-C and LDL-C. T2D liability appeared to increase all triglyceride groups (e.g., triglycerides in VLDL 0.04 SD; 95% CI 0.01, 0.07), whilst CAD liability had little effect on triglyceride groups except for LDL triglycerides which appeared to be increased. T2D liability was estimated to increase VLDL particle size (0.04 SD; 95% CI 0.02, 0.07), and decrease LDL and HDL particle size (-0.04 SD; 95% CI -0.06, -0.02 for both), whilst CAD liability was estimated to increase LDL particle size (0.06 SD; 95% CI 0.02, 0.10), and decrease HDL particle size (-0.06 SD; 95% CI -0.11, -0.01). T2D liability appeared to decrease, whilst CAD liability appeared to increase, apolipoprotein B,

although estimates were imprecise for CAD; whilst liability to either disease appeared to decrease apolipoprotein A1 (T2D -0.03 SD; 95% CI -0.05, -0.01; CAD -0.07 SD; 95% CI -0.11, -0.03). Effect estimates across sensitivity models were imprecise for T2D and CAD with each lipid trait (e.g., MR Egger estimate for effect liability to T2D on total-C 0.00 SD; 95% CI -0.05, 0.04), but Egger P-values were largely >0.05. Of note, the effect of higher CAD liability appeared to become inverse for VLDL-C and LDL-C, and apolipoprotein B in weighted median and weighted mode models (**Supplementary Tables 2-4; Supplementary Figure 7**). Evidence of heterogeneity was strong for all lipid and lipoprotein traits, e.g., Cochran's Q P-value = $4.22e^{-188}$ for the IVW estimate for T2D liability on total-C (**Supplementary Table 2**). Where the direction of effect estimates was inconsistent across MR models, we examined scatter plots of individual SNP effects (**Supplementary Figure 5**). These supported the existence of several outlier SNPs which were biasing estimates from IVW and MR Egger models; estimates from weighted median and weighted mode models appeared not to be influenced by these outliers and were in line with the majority of SNP effects.

The magnitudes of the effect of liability to T2D or CAD on statin or metformin use were small. As expected, higher liability to T2D was most strongly estimated to increase odds of taking metformin (**Supplementary Figure 1**; OR 1.02; 95% CI 1.01, 1.02), whilst higher CAD liability was most strongly estimated to increase odds of taking simvastatin (OR 1.03; 95% CI 1.03, 1.04). There was very little evidence of an effect of CAD liability on metformin use. The effect of higher liability to T2D on metformin use was robust across weighted median and weighted mode models, but attenuated in the MR-Egger model, with an Egger intercept P-value of $6.09e^{-16}$ (**Supplementary Tables 3&4; Supplementary Figure 6**). Estimates of the effect of liability to CAD on atorvastatin and simvastatin use were consistent across all sensitivity models. Evidence of heterogeneity was strong for all medication use traits (**Supplementary Table 2**).

The estimated effects of T2D and CAD liability on several lipids, particularly non-HDL cholesterol, differed markedly by age (**Figure 2**). Of note, higher T2D liability was estimated to decrease apolipoprotein B in the oldest tertile (IVW -0.04 SD, 95% CI -0.06, -0.02), whilst this effect was null in the youngest tertile, and between the two in the intermediate tertile. Among these age-stratified results, some estimated effects of T2D liability on lipids were inconsistent across sensitivity models, including for apolipoprotein B and LDL-C (**Supplementary Figure 8**), however, there was little evidence that Egger intercepts for these estimates differed from zero (**Supplementary Table 6**). Higher CAD liability was estimated to increase total-C, VLDL-C and LDL-C in the youngest tertile (e.g., total-C IVW 0.12 SD, 95% CI 0.05, 0.20). This effect was reduced in the intermediate tertile (e.g., total-C IVW 0.08 SD, 95% CI 0.01, 0.14) and further reduced to the null in the oldest tertile (**Supplementary Table 5**). A similar attenuation with age was seen in triglycerides in LDL and apolipoprotein B. There was notable inconsistency in the direction of these effect estimates for CAD liability with lipids across sensitivity models (**Supplementary Figure 11**) and some Egger P-values were <0.05, particularly for triglyceride estimates (e.g., youngest tertile, total triglycerides, Egger P-value $2.51e^{-4}$; **Supplementary Table 6**). Based on outlier-robust weighted median and mode models, higher CAD liability was estimated to decrease LDL-C and apolipoprotein B within the oldest age tertile (statin use 29%); these effects diminished within the intermediate age tertile (statin use 17%) and were null within the youngest age tertile (statin use 5%).

Fatty acids and amino acids

Effects on fatty acid traits were largely opposite for liability to T2D and liability to CAD, but IVW estimates for CAD were often imprecise (**Figure 3**). For example, higher liability to T2D was estimated to decrease the ratio of docosahexaenoic acid to total fatty acids (-0.02 SD; 95% CI -0.03, -0.01) whereas liability to CAD was estimated to increase the same ratio (0.03 SD; 95% CI 0.00, 0.05). Liability to CAD was estimated to decrease the ratio of saturated to total fatty acids,

and increase the ratio of docosahexaenoic acid to total fatty acids, and both of these were consistent across all sensitivity models. CAD liability effects on some fatty acid ratios were stronger using the MR-Egger method (**Supplementary Table 3, Supplementary Figure 7**; e.g., ratio of monounsaturated fatty acids to total fatty acids -0.19 SD; 95% CI -0.28, -0.10), whereas estimates for T2D liability on fatty acid ratios were largely attenuated compared to IVW using MR-Egger. Higher liability to T2D was estimated to increase all amino acid traits (except for glycine), including total BCAAs (IVW 0.05 SD; 95% CI 0.04, 0.07), which was robustly positive across sensitivity models, but the MR Egger and weighted mode estimates were slightly attenuated (**Supplementary Figure 6**). There was little evidence of an effect of CAD liability on any amino acid trait, and this was consistent across all models, except for total BCAAs and valine for which the MR-Egger estimates suggested a decrease (**Supplementary Tables 2-4; Supplementary Figure 8**). Evidence of heterogeneity was strong for most fatty acid and amino acid traits, e.g., Cochran Q P = $7.55e^{-49}$ for the IVW estimate for T2D liability on total BCAAs (**Supplementary Table 2**).

The effect of liability to T2D on fatty acids and amino acids as estimated by IVW models appeared to be fairly consistent with age, with exceptions including total fatty acids, ratio of saturated to total fatty acids, and histidine (**Figure 4**). For histidine, effect estimates were consistently null in the intermediate tertile, and consistently positive in the oldest and youngest tertiles (**Supplementary Tables 5-7**). The effect of higher liability to CAD on total fatty acids decreased with age, with the youngest tertile exhibiting an increasing effect (0.06 SD, 95% CI 0.00, 0.11), the intermediate tertile a null effect and the oldest tertile a decreasing effect (-0.05 SD, 95% CI -0.09, 0.00) (**Supplementary Table 5**). The effect of higher CAD liability on both omega-3 to total fatty acids and docosahexaenoic acid to total fatty acids ratios were estimated to be null in the youngest tertile but increasing in both the intermediate and oldest tertiles, and these estimates were largely consistent across sensitivity models (**Supplementary Tables 5-7**).

Glycolysis traits and other metabolites

As expected, T2D liability was estimated to increase glucose levels (0.08 SD; 95% CI 0.06, 0.10) and this was consistent with weighted median and weighted mode estimates, but the MR-Egger estimate was attenuated (**Supplementary Tables 2-4; Supplementary Figure 6**); there was no effect of CAD liability on glucose. There was little evidence of effect of T2D liability or CAD liability on other glycolysis related metabolites, ketone bodies or fluid balance metabolites (**Figure 5**). T2D liability was estimated to slightly increase glycoprotein acetyls (0.02 SD; 95% CI 0.00, 0.04), although there were inconsistencies across different methods (**Supplementary Figure 6**), whilst CAD liability had little effect (0.01 SD; 95% CI -0.03, 0.05). Effect estimates for CRP were similar to those for glycoprotein acetyls. Evidence of heterogeneity was strong for most pre-glycaemic and other metabolic traits, e.g., Cochran Q P = $4.16e^{-112}$ for the IVW estimate for T2D liability on glucose (**Supplementary Table 2**).

Estimates of the effects of T2D and CAD liability were largely consistent across age tertiles for glycolysis traits and ketone bodies, with the exception of the effect estimate for CAD liability on lactate and pyruvate which was positive in the youngest tertile and null in the intermediate and oldest tertiles (**Figure 6**). The effect of higher liability to T2D on glycoprotein acetyls decreased with age, with an increase in the youngest tertile (0.04 SD, 95% CI 0.02, 0.06) and a null effect in the oldest tertile. However, the increasing effect in the youngest tertile was not robust across sensitivity models (**Supplementary Tables 6-7; Supplementary Figure 10**).

Adiposity, smoking, and alcohol consumption

Higher genetic liability to T2D was estimated to increase all adiposity outcomes according to IVW estimates (**Supplementary Figure 2**; e.g., BMI: 0.03 SD; 95% CI 0.01, 0.05). However, there was evidence to suggest those estimates may be biased by horizontal pleiotropy; MR-

Egger P-values were <0.05 for all adiposity traits (**Supplementary Table 3)**. There was little evidence of an effect of liability to CAD on any of the adiposity outcomes according to IVW, weighted median and weighted mode estimates, however MR-Egger estimates suggested an effect on decreased adiposity (e.g., whole body fat mass: -0.03 SD; 95% CI -0.07, 0.00). Evidence of heterogeneity was strong for all adiposity traits, e.g., Cochran's Q P-value = $2.51e^{-147}$ for the IVW estimate for CAD liability on BMI (**Supplementary Table 2**). The effects of liability to T2D and CAD on both smoking status and alcohol drinker status were null (**Supplementary Figure 3)**. Higher liability to T2D was estimated to increase alcohol intake frequency in IVW models (0.02 SD, 95% CI 0.01, 0.04), although the effect was null in sensitivity models (**Supplementary Tables 3&4**). Higher liability to CAD had no effect on alcohol intake frequency. There was no effect of liability to either disease on pack years of smoking.

Summary XY plot of T2D and CAD liability for metabolic and other outcomes

Estimates of the effects of liability to T2D and CAD on each metabolite (ages combined), adiposity, and medication use were visualized in an XY plot to compare the overall pattern of effects of liability to either disease across outcomes (**Supplementary Figure 4**). The slope of the regression line was -0.16 with an R^2 of 0.02, which suggested a lack of association between profiles, giving further evidence that genetic liability to T2D and to CAD have distinct metabolic features.

DISCUSSION

In this study, we directly compared the metabolic profiles of genetic liability to T2D and CAD, two diseases which commonly co-occur but which involve different pathophysiology and clinical presentation. We applied a reverse MR framework[39] using new summary-level GWAS data on metabolomics from UK Biobank, which enabled a sample size which was 5-times larger than

previous studies.[40,41] Our results based on pleiotropy-robust models suggest that genetic liability to T2D and to CAD have largely distinct metabolic features, including increased BCAAs in T2D and apparently decreased LDL-C and apolipoprotein B in CAD, particularly in the oldest age tertile where statin use is common. Elucidating distinct metabolic profiles of liability to each disease creates a useful opportunity for more targeted early detection and treatment of each disease. However, if there is a goal to design effective interventions that reduce T2D and CAD simultaneously, their lack of shared metabolic features could represent a challenge. Nevertheless, there is evidence that some T2D therapeutics, including sodium-glucose cotransporter 1/2 (SGLT1/2) inhibitors, reduce incidence of cardiovascular disease in patients with T2D[59,60], suggesting that it is possible to target both diseases at once via a metabolic mechanism. In addition, the lack of evidence linking CAD liability with circulating glucose in the present study suggests that the observed protective effect of SGLT1/2 inhibitors may be via an alternative mechanism to lowering serum glucose levels.

We found evidence to suggest that both LDL-C and HDL-C are reduced in response to increased liability to T2D, which is consistent with evidence from a previous reverse MR study among young people for HDL-C but not LDL-C (which was raised),[40] and a conventional 'forward' MR study that found increased HDL-C and LDL-C were estimated to decrease risk of developing T2D.[37] This is also in accordance with evidence that statin therapy (which lowers LDL-C) increases risk of T2D.[61] Our results suggest that higher liability to CAD increases LDL-C and apolipoprotein B based on IVW models which is consistent with a previous reverse MR conducted across younger samples,[41] suggesting that perturbations in LDL-C in early life due to CAD liability persist into adulthood. We did, however, see attenuation of these effects in the oldest age tertile which could be due to the increased statin use in the older sample (prevalence of statin use being 29% in the oldest tertile, versus 5% in the youngest tertile); and the direction of effect was inconsistent across sensitivity models, suggesting that these results

were vulnerable to outlier SNP effects. Results based on outlier-robust weighted median and mode models suggested that CAD liability has a small or null effect on LDL-C and apolipoprotein B, although this could reflect residual distortions by medication use.

Increasing HDL-C has been hypothesized as a mechanism to reduce risk of CAD, based largely on conventional non-genetic epidemiology studies.[62] However, evidence from MR studies has showed that HDL-C-raising genetic variants do not reduce risk of CAD,[38,63] and a meta-analysis showed that that HDL-C modifying treatments did not reduce cardiovascular mortality.[64] Our current study detected an effect of CAD liability on decreasing HDL-C levels, suggesting that reduced HDL-C is an early feature of CAD development, potentially explaining why lower HDL-C appears to be associated with higher CAD risk in observational settings. LDL triglycerides have been previously associated with CAD risk observationally,[65] but our overall results show little evidence of an effect of CAD liability on LDL triglycerides, suggesting that increased LDL triglyceride levels are also not an early feature of CAD. However, in our age stratified results, we did find evidence that CAD liability increases LDL triglycerides in the youngest and intermediate tertiles, suggesting that the lack of overall effect may reflect the increased statin use in the older population. Differences across age tertiles seen in other triglyceride types also support this.

Our results suggest that liability to T2D and liability to CAD confer opposite effects on fatty acid metabolism. Inverse effects of T2D liability on various fatty acid metabolites (e.g., omega-6 and polyunsaturated) support them as potential candidates for further functional studies into their role in protection against T2D. However, a systematic review of randomized controlled trials of polyunsaturated fatty acid supplementation found no strong evidence that increasing omega-3, omega-6 or total polyunsaturated fatty acids provided protection against or treated T2D.[66] We observed an effect of higher CAD liability on increasing ratios of omega-3 fatty acids and

docosahexaenoic acids to total fatty acids. This contrasts with existing evidence of a protective effect of omega-3 fatty acids on cardiovascular outcomes.[67] We saw evidence of substantial age differences in these effects, where higher liability to CAD increased these ratios in the older tertiles but not the youngest, which could be an effect of medication or improved lifestyle in response to a CAD diagnosis.

Consistent with previous studies, we found evidence that liability to T2D was robustly estimated to increase total BCAAs[17,40] whereas there was no evidence of an effect of liability to CAD on BCAAs, highlighting that increased total BCAAs are exclusively a feature of liability to T2D. Although both T2D and CAD are associated with inflammation,[68] evidence of an effect on levels of glycoprotein acetyls from this study was not robust for liability to either disease when viewing ages collectively; however, higher T2D liability did appear to modestly raise glycoprotein acetyls within the youngest age group. Together, this suggests that increased inflammation is not a major feature of either disease process. However, we only considered two inflammatory markers, omitting other potentially important ones such as interleukin-6 and interleukin-1β that may play a role in disease pathogenesis.[69,70]

UK Biobank participants are older adults and therefore many of them are likely to be taking medication, including statins and metformin for the prevention/treatment of CAD and T2D, respectively. The overall prevalence of statin use in UK Biobank is 16%, with men almost twice as likely to be taking statins as women, and use increasing markedly across age tertiles.[51] Statins and metformin are known to lower LDL-C [48,71] and will likely have direct and indirect effects on other metabolic traits.[49] As expected, our results suggest that increased liability to T2D most strongly increases the odds of metformin use whilst increased liability to CAD most strongly increases the odds of statin use, providing useful positive controls and suggesting specificity of the genetic instruments used for each disease. Medication use may therefore

modify the effect of liability to T2D or CAD on metabolic traits and lead to distortions in the form of underestimated, or even reversed, observed effects of genetic liability. Medication use may explain inconsistencies between results for T2D or CAD liability and LDL-C from this study and previous studies, particularly studies of younger people who are less likely to be taking medication.[41,65] By stratifying the cohort by age as a proxy for medication use before performing metabolite GWAS, we were likely better positioned to overcome these distortions and better able to isolate the effects of disease liability. Conditioning on age (as a proxy instead of medication use) instead of medication use itself, which is affected by disease liability, also reduced the potential for collider bias in our analyses.

It is well established that excess adiposity is a causal risk factor for higher T2D risk,[72] and there is also strong causal evidence that adiposity increases CAD risk in a conventional (forward-direction) MR study[73]. In this study, we did not find consistent evidence that liability to either T2D or CAD has an effect on raising adiposity. This likely reflects the more distal nature of adiposity as compared with circulating metabolites, and possibly highly pleiotropic variants among the outcome SNPs.[74] We also did not find consistent evidence that liability to either T2D or CAD has an effect on smoking or alcohol behaviours. Although increased smoking and drinking are risk factors for both T2D and CAD,[75,76] the lack of signal seen in our study suggests that they are not proximal features of disease liability. This provides further evidence that the reverse MR approach is useful for revealing more proximal factors like metabolites that are directly involved in the disease process and highlights the importance of combining evidence from MR study designs in different directions to reveal the full scope of factors contributing to the developing disease process.

Study limitations

A potential limitation of this study is that the samples used to generate genetic instruments for T2D and CAD included participants from UK Biobank, the same sample from which we obtained genetic association estimates for metabolic traits, medication use, and lifestyle-related outcomes. This may have led to bias in the results in the direction of the observational association, however, given the strong genetic instruments for T2D liability and CAD liability, bias from sample overlap is expected to be small.[77] Another limitation is the unrepresentative nature of the UK Biobank sample, which had an initial response rate of around 5% and therefore is vulnerable to various forms of selection bias. Replication of this study in other large cohort studies and application of different approaches will allow more robust metabolic characterisation of T2D liability and CAD liability; although UK Biobank is currently the largest such data source in existence. The metabolite GWAS with all ages collectively was not age-adjusted, leaving the potential for confounding by age within those metabolite effect estimates; although we were able to additionally examine these effects within age tertiles. Use of only summary-level data limits the capacity to fully explore effects of other factors such as medication use, sex, and ethnicity, which may affect metabolite levels. A further limitation of this study is the use of targeted NMR metabolite data, rather than mass spectrometry measures which are not yet available at large scale, but offer a broader representation of metabolites beyond the lipid subclasses. Measurement of adiposity by bio-impedance is less accurate than measurement by dual-energy x-ray absorptiometry or magnetic resonance imaging, which will be available in future in the UK Biobank cohort. The smoking pack-years GWAS was restricted to ever-smokers, which may induce sampling bias and invalidate the MR assumptions for that analysis.

Conclusions

Our results support largely distinct metabolic profiles of genetic liability to T2D and CAD using metabolite GWAS data from the largest cohort to date. Our most robust findings suggest that T2D liability increases BCAAs, and CAD liability decreases LDL-C and apolipoprotein B. Such

apparently protective effects of CAD liability differ substantially by age and likely reflect mediation by statin use in adulthood.

FUNDING

MLS is supported by the Wellcome Trust through a PhD studentship [218495/Z/19/Z]. CJB, MVH, GDS, JAB and ELA work in a unit funded by the UK MRC (MC_UU_00011/1; MC_UU_00011/4) and the University of Bristol. MVH works in a unit that receives funding from the UK MRC and is supported by a BHF Intermediate Clinical Research Fellowship (FS/18/23/33512) and the National Institute for Health Research Oxford Biomedical Research Centre. CJB is supported by Diabetes UK (17/0005587) and the World Cancer Research Fund (WCRF UK), as part of the World Cancer Research Fund International grant program (IIG_2019_2009). CJB and MLS acknowledge funding from the Wellcome Trust (202802/Z/16/Z).

CONFLICTS OF INTEREST

MVH has collaborated with Boehringer Ingelheim in research, and in adherence to the University of Oxford's Clinical Trial Service Unit & Epidemiological Studies Unit (CSTU) staff policy, did not accept personal honoraria or other payments from pharmaceutical companies. No others to declare.

DATA AVAILABILITY

All summary level GWAS results are publicly available through the IEU-OpenGWAS platform, accessible at https://gwas.mrcieu.ac.uk/.

AUTHOR CONTRIBUTIONS

JAB, ELA and CJB conceived and planned the study and supervised analyses. MLS conducted the analyses and wrote the first draft. All others critically reviewed the intellectual content of manuscript drafts and with MLS approved the final version for submission.

ACKNOWLEDGEMENTS

We thank the participants of the UK Biobank study and the genome-wide association study consortiums who made their summary statistics publicly available for this study.

REFERENCES

1 International Diabetes Federation. IDF Diabetes Atlas, 9th edn. Brussels, Belgium: 2019. https://www.diabetesatlas.org

2 Khan MA, Hashim MJ, Mustafa H, *et al.* Global Epidemiology of Ischemic Heart Disease: Results from the Global Burden of Disease Study. *Cureus* Published Online First: 2020. doi:10.7759/cureus.9349

3 Mahajan A, Taliun D, Thurner M, *et al.* Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* 2018;**50**:1505–13. doi:10.1038/s41588-018-0241-6

4 Nikpay M, Goel A, Won H-H, *et al.* A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;**47**:1121–30. doi:10.1038/ng.3396

5 Clish CB. Metabolomics: an emerging but powerful tool for precision medicine. *Mol Case Stud* 2015;**1**:a000588. doi:10.1101/mcs.a000588

6 Ottosson F, Smith E, Melander O, *et al.* Altered Asparagine and Glutamate Homeostasis Precede Coronary Artery Disease and Type 2 Diabetes. *J Clin Endocrinol Metab* 2018;**103**:3060–9. doi:10.1210/jc.2018-00546

7 Urpi-Sarda M, Almanza-Aguilera E, Llorach R, *et al.* Non-targeted metabolomic

biomarkers and metabotypes of type 2 diabetes: A cross-sectional study of PREDIMED trial participants. *Diabetes Metab* 2019;**45**:167–74. doi:10.1016/j.diabet.2018.02.006

8 Bos MM, Noordam R, Bennett K, *et al.* Metabolomics analyses in non-diabetic middle-aged individuals reveal metabolites impacting early glucose disturbances and insulin sensitivity. *Metabolomics* 2020;**16**. doi:10.1007/s11306-020-01653-7

9 Guasch-Ferré M, Hruby A, Toledo E, *et al.* Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis. *Diabetes Care* 2016;:833–46. doi:10.2337/dc15-2251

10 Ahola-Olli A V., Mustelin L, Kalimeri M, *et al.* Circulating metabolites and the risk of type 2 diabetes: A prospective study of 11,896 young adults from four Finnish cohorts. *Diabetologia* 2019;:2298–309. doi:10.1101/513648

11 Al-Sulaiti H, Diboun I, Agha M V., *et al.* Metabolic signature of obesity-associated insulin resistance and type 2 diabetes. *J Transl Med* 2019;**17**:1–11. doi:10.1186/s12967-019-2096-8

12 Concepcion J, Chen K, Saito R, *et al.* Identification of pathognomonic purine synthesis biomarkers by metabolomic profiling of adolescents with obesity and type 2 diabetes. *PLoS One* 2020;**15**:e0234970. doi:10.1371/journal.pone.0234970

13 Gu X, Al Dubayee M, Alshahrani A, *et al.* Distinctive Metabolomics Patterns Associated With Insulin Resistance and Type 2 Diabetes Mellitus. *Front Mol Biosci* 2020;**7**. doi:10.3389/fmolb.2020.609806

14 Liu J, Semiz S, van der Lee SJ, *et al.* Metabolomics based markers predict type 2 diabetes in a 14-year follow-up study. *Metabolomics* 2017;**13**:104. doi:10.1007/s11306-017-1239-2

15 Roberts LD, Koulman A, Griffin JL. Towards metabolic biomarkers of insulin resistance and type 2 diabetes: progress from the metabolome. *Lancet Diabetes Endocrinol* 2014;**2**:65–75. doi:10.1016/S2213-8587(13)70143-8

16 Suhre K, Meisinger C, Döring A, *et al.* Metabolic Footprint of Diabetes: A Multiplatform Metabolomics Study in an Epidemiological Setting. *PLoS One* 2010;**5**:e13953. doi:10.1371/journal.pone.0013953

17 Sun Y, Gao H-Y, Fan Z-Y, *et al.* Metabolomics Signatures in Type 2 Diabetes: A Systematic Review and Integrative Analysis. *J Clin Endocrinol Metab* 2020;**105**:1000–8. doi:10.1210/clinem/dgz240

18 Li Y, Zhang D, He Y, *et al.* Investigation of novel metabolites potentially involved in the pathogenesis of coronary heart disease using a UHPLC-QTOF/MS-based metabolomics approach. *Sci Rep* 2017;**7**. doi:10.1038/s41598-017-15737-3

19 Smith E, Ottosson F, Hellstrand S, *et al.* Ergothioneine is associated with reduced mortality and decreased risk of cardiovascular disease. *Heart* 2020;**106**:691–7. doi:10.1136/heartjnl-2019-315485

20 Tzoulaki I, Castagné R, Boulangé CL, *et al.* Serum metabolic signatures of coronary and carotid atherosclerosis and subsequent cardiovascular disease. *Eur Heart J* 2019;**40**:2883–96. doi:10.1093/eurheartj/ehz235

21 Amin AM, Mostafa H, Arif NH, *et al.* Metabolomics profiling and pathway analysis of human plasma and urine reveal further insights into the multifactorial nature of coronary artery disease. *Clin Chim Acta* 2019;**493**:112–22. doi:10.1016/j.cca.2019.02.030

22 Li R, Li F, Feng Q, *et al.* An LC-MS based untargeted metabolomics study identified novel biomarkers for coronary heart disease. *Mol Biosyst* 2016;**12**:3425–34. doi:10.1039/c6mb00339g

23 Basak T, Varshney S, Hamid Z, *et al.* Identification of metabolic markers in coronary artery disease using an untargeted LC-MS based metabolomic approach. *J Proteomics* 2015;**127**:169–77. doi:10.1016/j.jprot.2015.03.011

24 Fan Y, Li Y, Chen Y, *et al.* Comprehensive Metabolomic Characterization of Coronary Artery Diseases. *J Am Coll Cardiol* 2016;**68**:1281–93. doi:10.1016/j.jacc.2016.06.044

25 Vernon ST, Tang O, Kim T, *et al.* Metabolic Signatures in Coronary Artery Disease: Results from the BioHEART-CT Study. *Cells* 2021;**10**:980. doi:10.3390/cells10050980

26 Wang Z, Zhu C, Nambi V, *et al.* Metabolomic Pattern Predicts Incident Coronary Heart Disease. *Arterioscler Thromb Vasc Biol* 2019;**39**:1475–82. doi:10.1161/atvbaha.118.312236

27 Gao X, Ke C, Liu H, *et al.* Large-scale Metabolomic Analysis Reveals Potential Biomarkers for Early Stage Coronary Atherosclerosis. *Sci Rep* 2017;**7**. doi:10.1038/s41598-017-12254-1

28 Xu X, Gao B, Guan Q, *et al.* Metabolomic profile for the early detection of coronary artery disease by using UPLC-QTOF/MS. *J Pharm Biomed Anal* 2016;**129**:34–42. doi:10.1016/j.jpba.2016.06.040

29 Wurtz P, Raiko JR, Magnussen CG, *et al.* High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis. *Eur Heart J* 2012;**33**:2307–16. doi:10.1093/eurheartj/ehs020

30 Mehta A, Liu C, Nayak A, *et al.* Untargeted high-resolution plasma metabolomic profiling predicts outcomes in patients with coronary artery disease. *PLoS One* 2020;**15**:e0237579. doi:10.1371/journal.pone.0237579

31 Shah SH, Sun J-L, Stevens RD, *et al.* Baseline metabolomic profiles predict cardiovascular events in patients at risk for coronary artery disease. *Am Heart J* 2012;**163**:844-850.e1. doi:10.1016/j.ahj.2012.02.005

32 Karagiannidis E, Sofidis G, Papazoglou AS, *et al.* Correlation of the severity of coronary artery disease with patients' metabolic profile- rationale, design and baseline patient characteristics of the CorLipid trial. *BMC Cardiovasc Disord* 2021;**21**. doi:10.1186/s12872-021-01865-2

33 Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*

2003;**32**:1–22. doi:10.1093/ije/dyg070

34 Sanderson E, Glymour MM, Holmes M V, *et al.* Mendelian randomization. *Nat Rev Methods Prim* 2022;**2**. doi:10.1038/s43586-021-00092-5

35 Ference BA, Yoo W, Alesh I, *et al.* Effect of Long-Term Exposure to Lower Low-Density Lipoprotein Cholesterol Beginning Early in Life on the Risk of Coronary Heart Disease. *J Am Coll Cardiol* 2012;**60**:2631–9. doi:10.1016/j.jacc.2012.09.017

36 Holmes M V, Asselbergs FW, Palmer TM, *et al.* Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J* 2015;**36**:539–50. doi:10.1093/eurheartj/eht571

37 White J, Swerdlow DI, Preiss D, *et al.* Association of Lipid Fractions With Risks for Coronary Artery Disease and Diabetes. *JAMA Cardiol* 2016;**1**:692. doi:10.1001/jamacardio.2016.1884

38 Richardson TG, Sanderson E, Palmer TM, *et al.* Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLOS Med* 2020;**17**:e1003062. doi:10.1371/journal.pmed.1003062

39 Holmes M V, Davey Smith G. Can Mendelian Randomization Shift into Reverse Gear? *Clin Chem* 2019;**65**:363–6. doi:10.1373/clinchem.2018.296806

40 Bell JA, Bull CJ, Gunter MJ, *et al.* Early Metabolic Features of Genetic Liability to Type 2 Diabetes: Cohort Study With Repeated Metabolomics Across Early Life. *Diabetes Care* 2020;**43**:1537–45. doi:10.2337/dc19-2348

41 Battram T, Hoskins L, Hughes DA, *et al.* Coronary artery disease, genetic risk and the metabolome in young individuals. *Wellcome Open Res* 2019;**3**:114. doi:10.12688/wellcomeopenres.14788.2

42 Hartwig FP, Tilling K, Davey Smith G, *et al.* Bias in two-sample Mendelian randomization when using heritable covariable-adjusted summary associations. *Int J Epidemiol* Published Online First: 2021. doi:10.1093/ije/dyaa266

43 Aragam KG, Jiang T, Goel A, *et al.* Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *medRxiv* 2021;:2021.05.24.21257377. doi:10.1101/2021.05.24.21257377

44 Littlejohns TJ, Sudlow C, Allen NE, *et al.* UK Biobank: opportunities for cardiovascular research. *Eur Heart J* 2019;**40**:1158–66. doi:10.1093/eurheartj/ehx254

45 Haworth S, Mitchell R, Corbin L, *et al.* Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nat Commun* 2019;**10**:333. doi:10.1038/s41467-018-08219-1

46 Sudlow C, Gallacher J, Allen N, *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;**12**:e1001779. doi:10.1371/journal.pmed.1001779

47 Würtz P, Kangas AJ, Soininen P, *et al.* Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies. *Am J Epidemiol* 2017;**186**:1084–96. doi:10.1093/aje/kwx016

48 Ference BA, Ginsberg HN, Graham I, *et al.* Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *Eur Heart J* 2017;**38**:2459–72. doi:10.1093/eurheartj/ehx144

49 Würtz P, Wang Q, Soininen P, *et al.* Metabolomic Profiling of Statin Use and Genetic Inhibition of HMG-CoA Reductase. *J Am Coll Cardiol* 2016;**67**:1200–10. doi:10.1016/j.jacc.2015.12.060

50 Swerdlow DI, Preiss D, Kuchenbaecker KB, *et al.* HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. *Lancet* 2015;**385**:351–61. doi:10.1016/s0140-6736(14)61183-1

51 Bell JA, Richardson TG, Wang Q, *et al.* Dominant role of abdominal adiposity in circulating lipoprotein, lipid, and metabolite levels in UK Biobank: Mendelian

randomization study. *medRxiv* 2021;:2021.05.29.21258044. doi:10.1101/2021.05.29.21258044

52 Mitchell R, Elsworth B, Raistrick C, *et al.* MRC IEU UK Biobank GWAS pipeline version 2. *data.bris* Published Online First: 2019. doi:https://doi.org/10.5523/bris.pnoat8cxo0u52p6ynfaekeigi

53 Hemani G, Zheng J, Elsworth B, *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 2018;**7**. doi:10.7554/elife.34408

54 Bowden J, Davey Smith G, Haycock PC, *et al.* Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* 2016;**under revi**:Available at: https://www.academia.edu/15479132/Co.

55 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**:512–25. doi:10.1093/ije/dyv080

56 Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* Published Online First: 2017. doi:10.1093/ije/dyx102

57 Burgess S, Labrecque JA. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur J Epidemiol* 2018;**33**:947–52. doi:10.1007/s10654-018-0424-6

58 Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011;**40**:755–64. doi:10.1093/ije/dyr036

59 Udell JA, Yuan Z, Rush T, *et al.* Cardiovascular Outcomes and Risks After Initiation of a Sodium Glucose Cotransporter 2 Inhibitor. *Circulation* 2018;**137**:1450–9. doi:10.1161/circulationaha.117.031227

60 Wiviott SD, Raz I, Bonaca MP, *et al.* Dapagliflozin and Cardiovascular Outcomes in Type

2 Diabetes. *N Engl J Med* 2019;**380**:347–57. doi:10.1056/nejmoa1812389

61 Crandall JP, Mather K, Rajpathak SN, *et al.* Statin use and risk of developing diabetes: results from the Diabetes Prevention Program. *BMJ Open Diabetes Res Care* 2017;**5**:e000438. doi:10.1136/bmjdrc-2017-000438

62 Linsel-Nitschke P, Tall AR. HDL as a target in the treatment of atherosclerotic cardiovascular disease. *Nat Rev Drug Discov* 2005;**4**:193–205. doi:10.1038/nrd1658

63 Voight BF, Peloso GM, Orho-Melander M, *et al.* Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet* 2012;**380**:572–80. doi:10.1016/S0140-6736(12)60312-2

64 Riaz H, Khan SU, Rahman H, *et al.* Effects of high-density lipoprotein targeting treatments on cardiovascular outcomes: A systematic review and meta-analysis. *Eur J Prev Cardiol* 2019;**26**:533–43. doi:10.1177/2047487318816495

65 März W, Scharnagl H, Winkler K, *et al.* Low-density lipoprotein triglycerides associated with low-grade systemic inflammation, adhesion molecules, and angiographic coronary artery disease: The Ludwigshafen Risk and Cardiovascular Health Study. *Circulation* 2004;**110**:3068–74. doi:10.1161/01.CIR.0000146898.06923.80

66 Brown TJ, Brainard J, Song F, *et al.* Omega-3, omega-6, and total dietary polyunsaturated fat for prevention and treatment of type 2 diabetes mellitus: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2019;:l4697. doi:10.1136/bmj.l4697

67 Khan SU, Lone AN, Khan MS, *et al.* Effect of omega-3 fatty acids on cardiovascular outcomes: A systematic review and meta-analysis. *EClinicalMedicine* 2021;**38**:100997. doi:10.1016/j.eclinm.2021.100997

68 Tsalamandris S, Antonopoulos AS, Oikonomou E, *et al.* The Role of Inflammation in Diabetes: Current Concepts and Future Perspectives. *Eur Cardiol* 2019;**14**:50–9. doi:10.15420/ecr.2018.33.1

69 The Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6R MR) Consortium. The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. *Lancet* 2012;**379**:1214–24. doi:10.1016/s0140-6736(12)60110-x

70 Ridker PM, Everett BM, Thuren T, *et al.* Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N Engl J Med* 2017;**377**:1119–31. doi:10.1056/nejmoa1707914

71 Solymár M, Ivic I, Pótó L, *et al.* Metformin induces significant reduction of body weight, total cholesterol and LDL levels in the elderly – A meta-analysis. *PLoS One* 2018;**13**:e0207947. doi:10.1371/journal.pone.0207947

72 Eckel RH, Kahn SE, Ferrannini E, *et al.* Obesity and Type 2 Diabetes: What Can Be Unified and What Needs to Be Individualized? *Diabetes Care* 2011;**34**:1424–30. doi:10.2337/dc11-0447

73 Lv W-Q, Zhang X, Fan K, *et al.* Genetically driven adiposity traits increase the risk of coronary artery disease independent of blood pressure, dyslipidaemia, glycaemic traits. *Eur J Hum Genet* 2018;**26**:1547–53. doi:10.1038/s41431-018-0180-9

74 Corbin LJ, Richmond RC, Wade KH, *et al.* BMI as a Modifiable Risk Factor for Type 2 Diabetes: Refining and Understanding Causal Estimates Using Mendelian Randomization. *Diabetes* 2016;**65**:3002–7. doi:10.2337/db16-0418

75 Chi Y, Wang X, Jia J, *et al.* Smoking Status and Type 2 Diabetes, and Cardiovascular Disease: A Comprehensive Analysis of Shared Genetic Etiology and Causal Relationship . Front. Endocrinol. . 2022;**13**.https://www.frontiersin.org/article/10.3389/fendo.2022.809445

76 Larsson SC, Burgess S, Mason AM, *et al.* Alcohol Consumption and Cardiovascular Disease. *Circ Genomic Precis Med* 2020;**13**. doi:10.1161/circgen.119.002814

77 Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol* 2016;**40**:597–608. doi:10.1002/gepi.21998

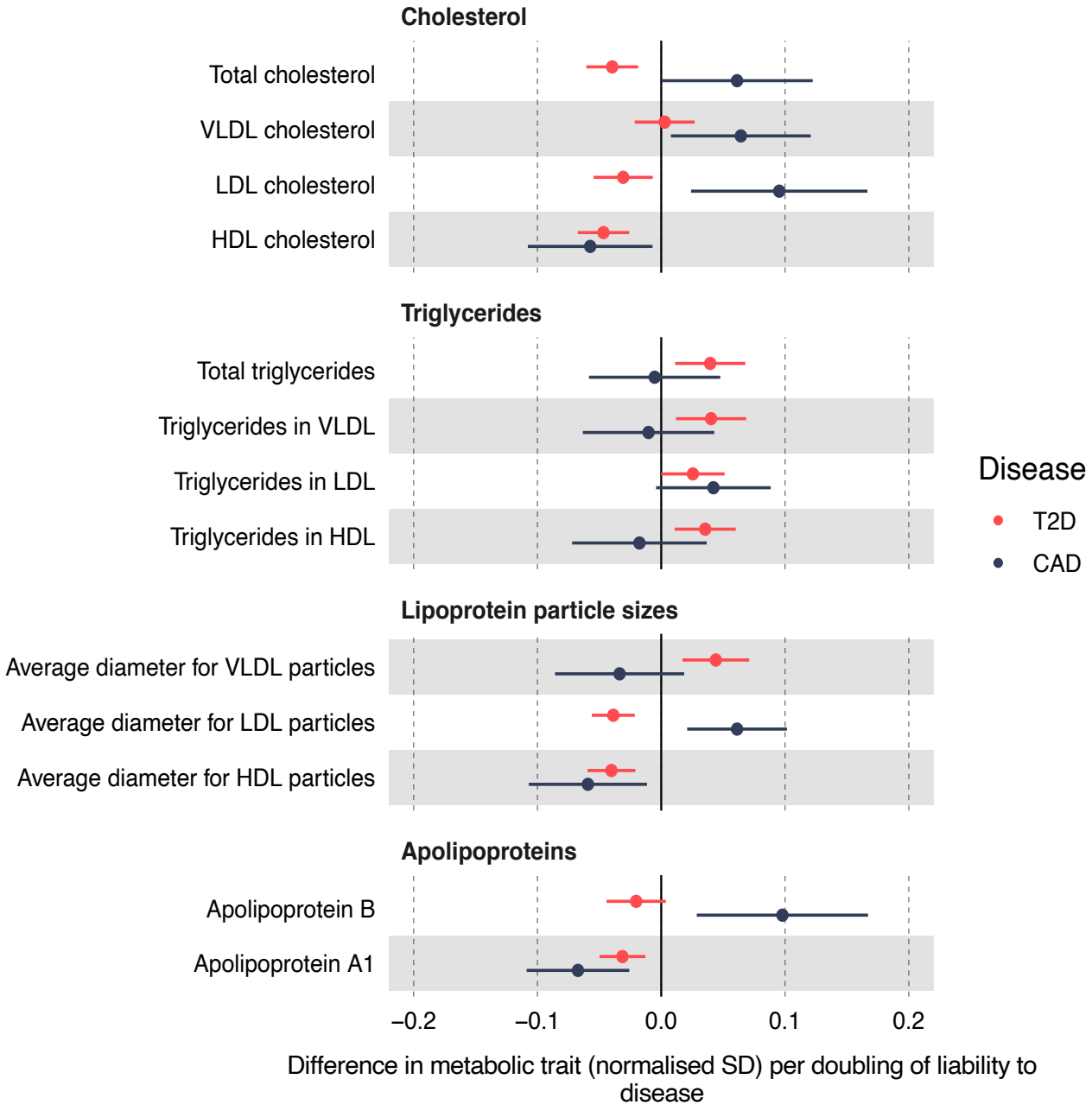

Figure 1: Effect of T2D and CAD liability on lipids and lipoproteins. Effect estimates are normalised SD unit differences in metabolite per doubling of liability to T2D or CAD based on IVW models.

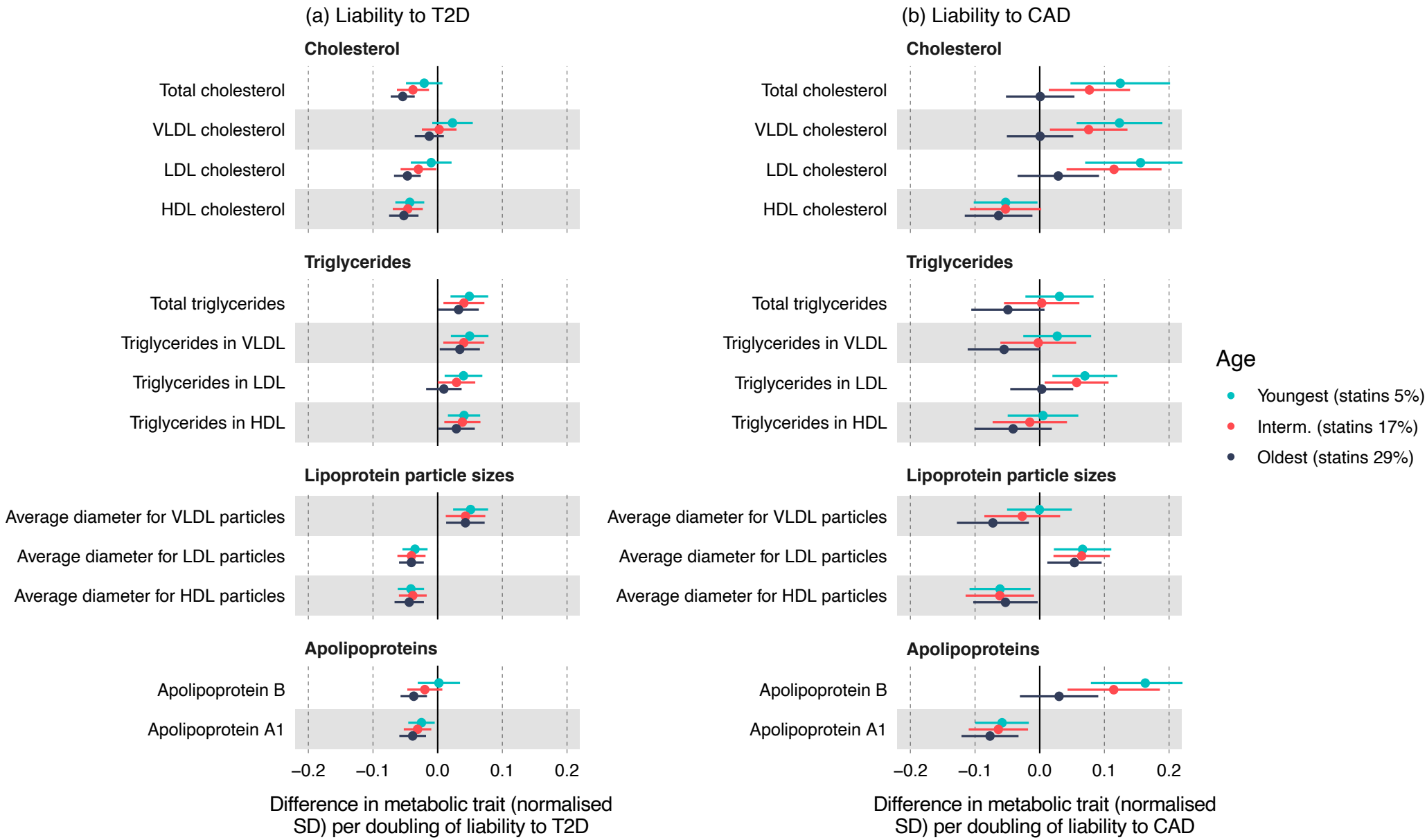

Figure 2. Effect of T2D (a) and CAD (a) liability on lipids and lipoproteins in age tertiles. Effect estimates are normalised SD unit differences in metabolite per doubling of liability to disease based on IVW models.

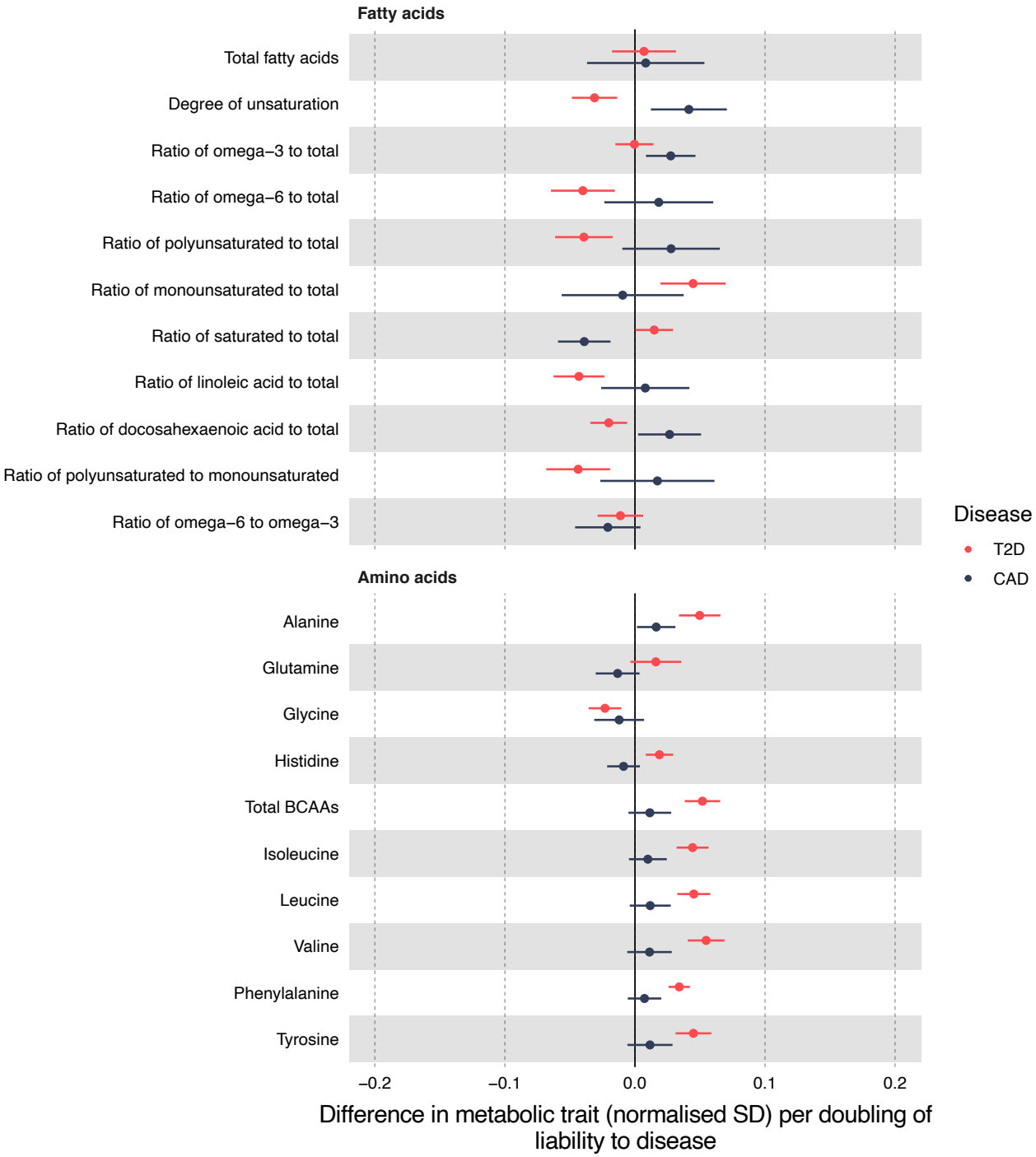

Figure 3: Effect of T2D and CAD liability on fatty acids and amino acids. Effect estimates are normalised SD unit differences in metabolite per doubling of liability to T2D or CAD based on IVW models.

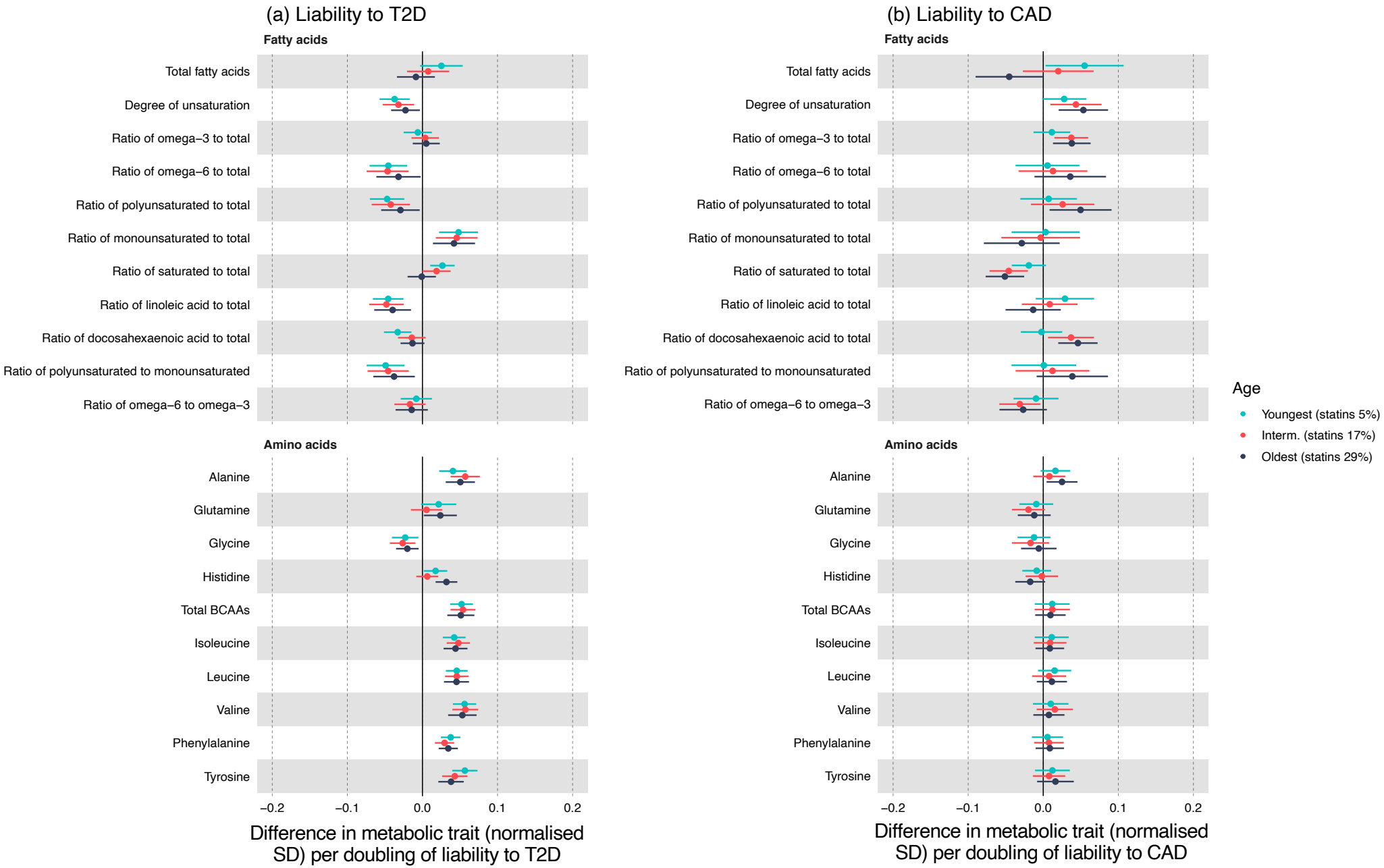

Figure 4. Effect of T2D (a) and CAD (b) liability on fatty acids and amino acids in age tertiles. Effect estimates are normalised SD unit differences in metabolite per doubling of liability to disease based on IVW models.

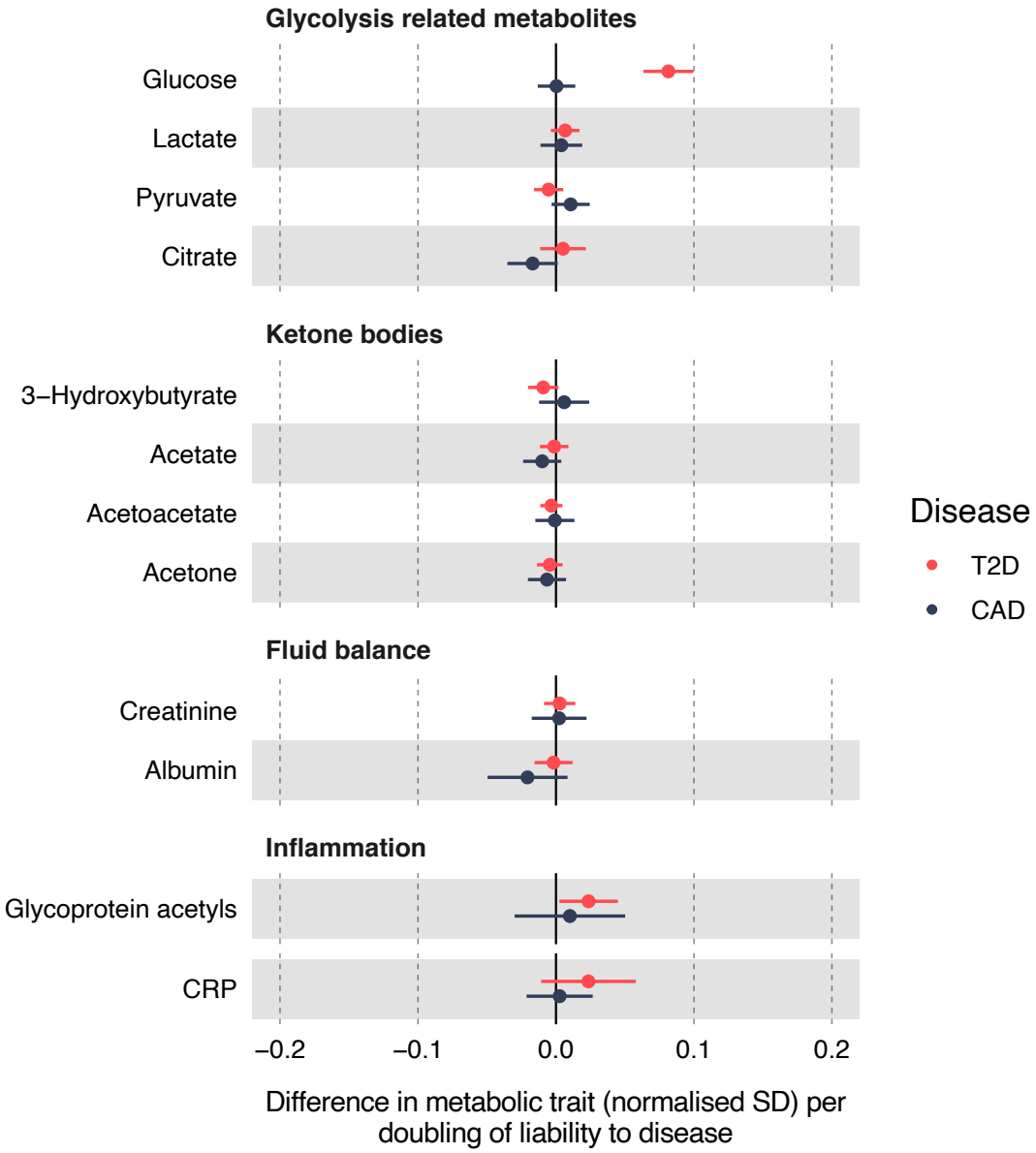

Figure 5: Effect of T2D and CAD liability on glycolysis related metabolites, ketone bodies, fluid balance metabolites and inflammation metabolites. Effect estimates are normalised SD unit differences in metabolite per doubling of liability to T2D or CAD based on IVW models.

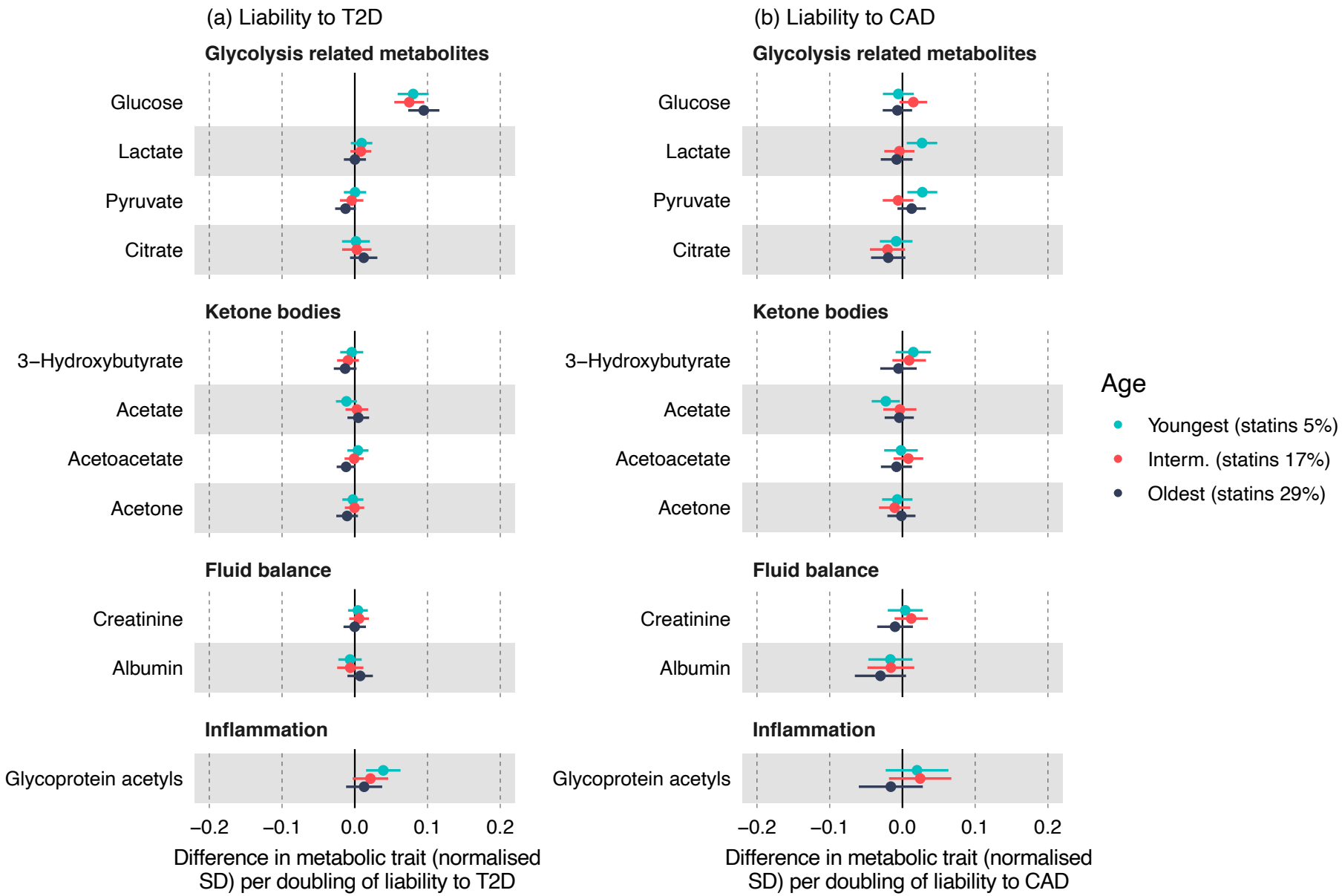

Figure 6. Effect of T2D (a) and CAD (b) liability on glycolysis related metabolites, ketone bodies, fluid balance metabolites and glycoprotein acetyls in age tertiles. Effect estimates are normalised SD unit differences in metabolite per doubling of liability to disease based on IVW models.