

Circulating metabolic biomarkers consistently predict incident type 2 diabetes in Asian and European populations – a plasma metabolomics analysis of four ethnic groups

Jowly Yi Hoong Seah^{1*}, Yueheng Hong^{1*}, Anna Cichońska^{2*}, Charumathi Sabanayagam^{3,4}, Simon Nusinovic³, Tien Yin Wong^{3,4,5}, Ching-Yu Cheng^{3,4,5}, Pekka Jousilahti⁶, Annamari Lundqvist⁶, Markus Perola^{6,7}, Veikko Salomaa⁶, E. Shyong Tai^{1,4,5}, Peter Würtz², Rob M. van Dam^{1,8†}, Xueling Sim^{1†}

¹Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore; Singapore 117549, Singapore

²Nightingale Health Ltd, Mannerheimintie 164a, 00300 Helsinki, Finland

³Singapore Eye Research Institute, Singapore National Eye Centre, Singapore 169856, Singapore

⁴Duke-NUS Medical School, Singapore 169857, Singapore

⁵Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore 117597, Singapore

⁶Department of Public Health Solutions, Finnish Institute for Health and Welfare, P.O. Box 30, FI-00271 Helsinki, Finland

⁷Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki, 00014 Helsinki, Finland.

⁸Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA; MA 02115

*These authors contributed equally.

†These authors are co-corresponding senior authors.

Rob M. van Dam. Saw Swee Hock School of Public Health, 12 Science Drive 2, #10-01, Singapore 117549; tel: +65 6516 4980; e-mail: rob.van.dam@nus.edu.sg

Xueling Sim. Saw Swee Hock School of Public Health, 12 Science Drive 2, #10-01, Singapore 117549; tel: +65 6601 5096; e-mail: ephsx@nus.edu.sg

Number of figures (to print, not Supplementary data): 1

Number of tables (to print, not Supplementary data): 2

Supplementary data submitted: 6 tables

Running title: Metabolic biomarkers and type 2 diabetes

Abbreviations: ARIC, Atherosclerosis Risk in Communities, BCAA, branched-chain amino acids, BMI, body mass index, GlycA, glycoprotein acetyls, HbA1c, glycated hemoglobin, HDL, high-density lipoprotein, IDL, intermediate-density lipoprotein, LDL, low-density lipoprotein, MEC, Multiethnic Cohort, MUFA, monounsaturated fat, NMR, nuclear magnetic resonance, PUFA polyunsaturated fat, SEED, Singapore Epidemiology of Eye Diseases, SFA, saturated fat, T2D, type 2 diabetes, VLDL, very low-density lipoprotein

Word count: 3946 (Introduction to Discussion)

1 **Abstract**

2 **Aims/hypothesis:** We aimed to evaluate metabolic biomarkers in relation to incident
 3 type 2 diabetes (T2D) representing three major ethnic groups in Asia (Chinese,
 4 Malay, and Indian) and a European population.

5 **Methods:** We used data from male and female adult participants of multiple cohorts,
 6 including two cohorts from Singapore (n = 6,393 Asians) consisting of ethnic
 7 Chinese, Malays, and Indians, and three cohorts of European-origin participants from
 8 Finland (n = 14,558). We used nuclear magnetic resonance to quantify 154 metabolic
 9 biomarkers in plasma collected at baseline and performed multivariable logistic
 10 regression to assess associations between the metabolic measures and risk of T2D
 11 with adjustments for age, sex, BMI and either fasting glucose or glycated
 12 haemoglobin (HbA1c).

13 **Results:** Of 154 metabolic biomarkers, 59 were associated with higher risk of T2D in
 14 both Asians and Europeans ($P < 0.0003$; Bonferroni-corrected). These included
 15 branched-chain and aromatic amino acids and alanine, the inflammatory marker
 16 glycoprotein acetyls, total fatty acids, the proportion of monounsaturated fatty acids,
 17 apolipoprotein B, larger very low-density lipoprotein particle sizes, and triglycerides.
 18 A further 13 metabolic biomarkers were associated with lower T2D risk in both
 19 populations including proportion of omega-6 polyunsaturated fatty acids and larger
 20 high-density lipoprotein particle sizes. Associations were consistent within the Asian
 21 ethnic groups ($P_{\text{het}} \geq 0.05$ for all 154 metabolic biomarkers) and largely consistent
 22 with the European population ($P_{\text{het}} \geq 0.05$ for 128 of 154 metabolic biomarkers).

23 **Conclusions/interpretation:** Metabolic biomarkers across a variety of biological
 24 pathways were consistently associated with T2D risk in three Asian ethnic groups and
 25 a European population.

26

27 **Research in context (199 words, not including headings)**

28 What is already known about this subject?

29 - Previous metabolomics studies have uncovered metabolites that were consistently
30 associated with risk of type 2 diabetes (T2D)

31 - However, these studies were mostly conducted in populations of predominantly
32 European ancestry, and it remains unclear if these findings can be generalized to
33 different Asian ethnic groups

34 What is the key question?

35 - Are associations between metabolic biomarkers and T2D risk consistent between
36 major Asian ethnic groups and a European population?

37 What are the new findings?

38 - Of 154 metabolic biomarkers, 59 including branched-chain and aromatic amino
39 acids, glycoprotein acetyls, total fatty acids, the proportion of monounsaturated fatty
40 acids, apolipoprotein B, larger very low-density lipoprotein particle sizes, and
41 triglycerides were associated with higher T2D risk and a further 13 were associated
42 with lower risk in both populations including proportion of omega-6 polyunsaturated
43 fatty acids and larger high-density lipoprotein particle sizes

44 - Associations were consistent within the Asian ethnic groups ($P_{\text{het}} \geq 0.05$ for all 154
45 metabolic biomarkers) and largely consistent with the European population ($P_{\text{het}} \geq$
46 0.05 for 128 of 154 metabolic biomarkers).

47 How might this impact on clinical practice in the foreseeable future?

48 - Metabolic aberrations were similarly associated with T2D risk in major Asian ethnic
49 groups and Europeans despite environmental, lifestyle, and physiological differences.

50 These metabolic biomarkers are modifiable and may inform individualized
51 interventions for the prevention of T2D

52

53 Keywords: NMR, metabolomics, type 2 diabetes, branched-chain amino acids,
54 aromatic amino acids, glycoprotein acetyls, inflammation, fatty acids, lipoprotein,
55 HDL, LDL, triglycerides, cholesterol

56

57 **Introduction**

58 The global burden of type 2 diabetes (T2D) is high and rising (1, 2), with the
59 prevalence expected to rise to 7079 cases per 100,000 in 2030, from 6059 cases per
60 100,000 in 2017 (2). In randomized trials, pharmaceutical and lifestyle interventions
61 substantially reduced the risk of T2D (3-5) in high-risk individuals. Variation in
62 genetic, environmental and lifestyle exposures between individuals, coupled with
63 heterogeneity in T2D (6, 7) has motivated the development of ‘precision medicine’
64 approaches to prevention (8-10). Such an approach requires the identification of risk
65 factors for T2D that can inform personalized interventions. Metabolic biomarkers
66 reflect the dynamic and ongoing state of homeostasis and metabolism (11). In recent
67 years, the field of metabolomics has gained traction due to technological advances
68 (12) and has uncovered various metabolites that are consistent predictors of T2D (13).
69 As many of these metabolic biomarkers are modifiable by diet or other lifestyle
70 factors, information on metabolic risk factors may guide personalized interventions.

71
72 Previous studies on metabolic profiles and T2D risk were mostly conducted in
73 populations of predominantly European ancestry, and it remains unclear if these
74 findings can be generalized to different Asian ethnic groups. For example, individuals
75 of Asian origin tend to develop T2D at a lower body mass index (BMI) and a younger
76 age than those of European origin (7), and ethnic Indians are generally more insulin
77 resistant than ethnic Malays or Chinese (14). Although some metabolomics studies on
78 T2D incidence have been conducted in East Asian (e.g. Chinese) populations, data on
79 South East Asian (e.g. Malays) and South Asian (e.g. Indians) populations are limited.

80

81 We quantified fatty acids, lipoprotein lipids and metabolites such as amino acids
82 using nuclear magnetic resonance (NMR) metabolomic analysis in two multi-ethnic
83 cohorts in Singapore and three cohorts in Finland. NMR metabolomics allows for
84 highly reproducible identification and quantification of compounds (15, 16). We
85 subsequently performed association analyses of these metabolic measures with risk of
86 T2D in three major Asian ethnic groups (Chinese, Malay, and Indian) and the
87 European population.

88

89 **Methods**

90 *Study Population*

91 The Asian participants included in this study were from two population-based cohorts
92 in Singapore that included ethnic Indians, Malays and Chinese: the Multiethnic
93 Cohort (MEC) (17) study and the Singapore Epidemiology of Eye Diseases (SEED)
94 cohort study (18, 19). The European participants in our study were from three Finnish
95 prospective population-based cohorts (FINRISK 2002, FINRISK 2012 and Health
96 2000) (20, 21). All studies included questionnaires about socio-demographic
97 characteristics and medical history, technician-measurement of anthropometric
98 measures and blood pressure, and collection of blood samples. Details of the
99 individual cohorts are provided in the **Supplemental Text**. Informed consent was
100 obtained from all participants and ethical approvals were obtained from the respective
101 ethics review boards.

102

103 For all cohorts, we excluded participants with diabetes, cardiovascular diseases or
104 cancer at baseline. From the MEC cohort of 9180 participants, we identified 694
105 incident diabetes cases using the following criteria: (a) fasting plasma glucose ≥ 7.0

106 mmol/L or HbA1c ≥ 6.5 % (47.5 mmol/mol) based on the American Diabetes
 107 Association (22) or self-reported physician-diagnosed diabetes at follow-up, or (b)
 108 record linkage with a national database of medical diagnoses as part of public
 109 hospital, polyclinic, and subsidized GP visits). Using risk-set sampling, we selected
 110 1315 controls matched for age (± 5 years), sex, date of health screening (± 2 years),
 111 and ethnicity from cohort participants who were alive, and did not have diabetes at the
 112 time the index case occurred. We used 6-year revisit data from 4384 SEED cohort
 113 participants among whom 391 developed diabetes during follow-up. Diabetes
 114 ascertainment at baseline and the revisit was based on random plasma glucose ≥ 11.1
 115 mmol/L or HbA1c ≥ 6.5 % (47.5 mmol/mol) (22), or a self-reported physician
 116 diagnosis. For the European cohorts, we used data on 14588 participants from
 117 FINRISK 2002, FINRISK 2012 and Health 2000, among whom 1055 developed
 118 diabetes during follow-up. Diabetes ascertainment at baseline was based on the
 119 National Social Insurance Institution Drug Purchase and Reimbursement Registries
 120 for purchases and reimbursements of purchases of hypoglycemic drugs, and the
 121 National Hospital Discharge Register for hospitalizations with diabetes as the
 122 diagnosis. Incident diabetes cases were identified by linking the participants to the
 123 registers described above with a unique personal identification code assigned to each
 124 Finnish citizen. Participants with prevalent or incident gestational diabetes were
 125 excluded.

126

127 *Metabolic Biomarker Quantification and Quality Control*

128 Metabolic biomarkers were quantified from plasma samples at baseline using a high-
 129 throughput ^1H -NMR metabolomics platform developed by Nightingale Health Ltd.
 130 (Helsinki, Finland; nightingalehealth.com/; [biomarker quantification version 2016](#)).

131 Details of the procedure have been described previously (23, 24). The metabolic
 132 biomarkers quantified reflected several etiologically diverse pathways, such as amino
 133 acids, fatty acids, ketone bodies, and gluconeogenesis-related measurements. In
 134 addition, lipoprotein measures including total concentration, total lipids,
 135 phospholipids, total cholesterol, cholesterol esters, free cholesterol and triglycerides
 136 within 14 lipoprotein subclasses were measured. 9 MEC and 20 SEED participants
 137 had >10% missing metabolic biomarker values and were excluded from subsequent
 138 analyses. For participants with metabolic biomarker values lower than detection level,
 139 we replaced values of 0 with a value equivalent to 0.9 multiplied by the non-zero
 140 minimum value of that measurement. Finally, we standardized the metabolic
 141 biomarker concentrations to z-scores, and all effect sizes were reported as per
 142 standard deviation increment to enable comparison of metabolic biomarker
 143 associations for measures with different units and concentration ranges. A total of 154
 144 individual metabolic biomarkers were included in our analyses.

145

146 *Statistical analyses*

147 We first calculated Pearson's correlation coefficients for all 154 metabolic
 148 biomarkers. Logistic regression was then used to assess associations between
 149 circulating metabolic biomarkers with T2D. For the matched MEC case-control study,
 150 conditional logistic regression model was used. In all studies, we adjusted for baseline
 151 age, sex, BMI, and fasting glucose (for MEC and Health 2000) or glycated
 152 hemoglobin (HbA1c, for non-fasting SEED, FINRISK 2002 and 2012 cohorts). For
 153 MEC, age² was included in the conditional logistic model as there was evidence of
 154 non-linearity in the relationship between age and T2D risk. Logistic models were
 155 performed separately by study and ethnic group to account for differences between

156 populations. Inverse variance-weighted fixed-effect meta-analysis was used to first
 157 pool ethnic-specific association statistics from the two Singapore cohorts, and then to
 158 combine association results from three ethnic groups. We also tested for heterogeneity
 159 using the Cochran Q heterogeneity. Bonferroni correction was used to account for
 160 multiple testing of 154 metabolic biomarkers and $P < 0.0003$ ($0.05/154$) was
 161 considered statistically significant.

162

163 To test the joint effects of multiple metabolic biomarkers on risk of T2D, we
 164 considered pairwise correlations between the metabolic biomarkers from MEC
 165 (Supplement Table 1). For the set of T2D associated metabolic biomarkers in the
 166 association analysis above that considered a metabolic biomarker individually across
 167 Asian and European populations, we first sorted the metabolic biomarkers by their
 168 strength of association. Starting with the most significantly associated metabolic
 169 biomarker (index-associated metabolic biomarker), we iteratively excluded other
 170 associated metabolic biomarkers that had correlations ≥ 0.80 with the index-
 171 associated metabolic biomarker (correlation-pruning). The set of remaining metabolic
 172 biomarkers thus had a Bonferroni-adjusted significant association with T2D and
 173 pairwise correlations < 0.80 . Finally, we performed stepwise regression analysis with
 174 the criterion of P -value of 0.05 in a joint model that included the selected biomarkers
 175 in the Asian and European cohorts. All statistical analyses were conducted using R
 176 3.3.2 (R Foundation for Statistical Computing, Vienna, Austria).

177

178 **Results**

179 In the nested MEC case-control study, 694 incident T2D cases and 1315 matched
 180 controls were included in this study during a mean follow-up of 6.9 y (standard

deviation SD: 2.4 y), and in the SEED cohort study, 391 incident T2D cases occurred in 4384 participants during an average of 6.2 y (SD: 1.1 y). In the European cohorts, 1055 incident T2D cases occurred in 14,558 participants during an average of 11.3 years (SD: 5.0 y). The average age was 54.0 y (SD: 9.8 y) for the Asian participants and 50.6 y (SD: 14.1 y) for the European participants, with similar representation of males and females for all populations, and Chinese, Malays and Indians for the Asian population (**Table 1**).

We observed high correlations among the branched-chain amino acids (BCAAs) leucine, isoleucine and valine, and concentrations of n-6 polyunsaturated (n-6 PUFA) fatty acids, PUFA, monounsaturated fatty acids (MUFA), saturated fatty acids and linoleic acid (18:2n-6) (all $r > 0.80$) (**Supplemental Tables 1 to 4**). In general, the lipid components (total lipids, phospholipids, cholesterol esters, free cholesterol) with the exception of triglycerides within a particular lipoprotein subclass were also highly correlated. Overall, 59 of the 154 metabolic biomarkers were consistently associated with a higher risk and 13 with a lower risk of T2D in both Asians and Europeans ($P_{meta} < 0.0003$; Bonferroni-corrected; $P_{het} < 0.05$ for 12 metabolic biomarkers) (**Table 2, Figure 1 and Supplemental Table 5**) after adjustment for age, sex, BMI and either fasting glucose or HbA1c. For the population-specific meta-analysis, we observed little heterogeneity in association across the three Asian ethnic groups ($P_{het} \geq 0.05$ for all 154 metabolic biomarkers) and across the three European cohorts ($P_{het} \geq 0.05$ for 144 of 154 metabolic biomarkers). Of the 29 metabolic biomarkers for which we did observe evidence for heterogeneity between the Asian and European populations, 23 were lipoprotein measures.

206 *Amino Acids and Inflammation*

207 Higher concentrations of BCAAs (leucine, isoleucine, and valine), aromatic amino
 208 acids (phenylalanine and tyrosine), and alanine were significantly and consistently
 209 associated with higher T2D risk in both Asians and Europeans (all $P_{\text{het}} > 0.05$). The
 210 strongest association was observed for leucine (Asians: OR 1.50 per 1 SD increment,
 211 95% CI 1.35 - 1.66; Europeans: OR 1.36, 95% CI 1.27 - 1.46). The inflammatory
 212 marker glycoprotein acetyls (GlycA) was also directly associated with risk of T2D
 213 (Asians: OR 1.41, 95% CI 1.28 - 1.55; Europeans: OR 1.34, 95% CI 1.25 - 1.43; P_{het}
 214 = 0.37). GlycA remained directly associated with T2D risk after further adjustments
 215 for CRP (data available only for MEC and the European cohorts, Asians: OR 1.94,
 216 95% CI 1.63 - 2.31; Europeans: OR 1.34, 95% CI 1.25 - 1.44).

217

218 *Fatty Acids*

219 Higher total concentrations of fatty acids in plasma were significantly associated with
 220 higher T2D risk. Among the different types of fatty acids, a higher proportion of
 221 MUFA was consistently associated with a higher T2D risk (Asians: OR 1.43, 95% CI
 222 1.30 - 1.57; Europeans: OR 1.38, 95% CI 1.27 - 1.51), and a higher proportion of n-6
 223 PUFA with a lower risk (Asians: OR 0.70, 95% CI 0.64 - 0.77; Europeans: OR 0.70,
 224 95% CI 0.65 - 0.75). The proportion of saturated fat (SFA) was directly associated
 225 with T2D risk, but this was not statistically significant after Bonferroni correction in
 226 Asians (Asians: OR 1.15, 95% CI: 1.05 - 1.25; Europeans: OR 1.21, 95% CI 1.13 -
 227 1.29). n-3 PUFAs were not significantly associated with risk of T2D in any ethnic
 228 group. There was no evidence of heterogeneity for associations of fatty acids with
 229 T2D between the Asian and European populations (all $P_{\text{het}} > 0.05$).

230

231 *Lipoprotein Measures*

232 We evaluated lipoprotein particle sizes, apolipoproteins and various lipoprotein
 233 measures including total concentrations, total lipids, phospholipids, total cholesterol,
 234 cholesterol esters, free cholesterol and triglycerides for 14 lipoprotein subclasses, in
 235 relation to T2D (**Figure 1** and **Supplemental Table 5**).

236
 237 Higher levels of apolipoprotein B were associated with higher T2D risk in both
 238 Asians (OR 1.24, 95% CI 1.13 - 1.35) and Europeans (OR 1.22, 95% CI 1.13 - 1.30;
 239 $P_{\text{het}} = 0.45$). Larger VLDL particle size was among the metabolic biomarkers with the
 240 strongest direct associations with T2D risk (Asians: OR 1.44, 95% CI 1.31 - 1.59;
 241 Europeans: OR 1.46, 95% CI 1.36 - 1.57; $P_{\text{het}} = 0.83$). Larger LDL particle sizes were
 242 inversely associated with incident T2D risk in Europeans (OR 0.82, 95% CI 0.76 -
 243 0.89), but not in Asians (Asians: OR 0.98, 95% CI 0.89 - 1.07; $P_{\text{het}} = 0.01$). Larger
 244 HDL particle sizes were also more strongly inversely associated with T2D risk in
 245 Europeans (OR 0.63, 95% CI 0.58 - 0.69) as compared with Asians (OR 0.77, 95% CI
 246 0.69 - 0.86; $P_{\text{het}} = 0.01$).

247
 248 The concentration of cholesterol in very low-density lipoprotein (VLDL) particles
 249 was generally associated with higher risk of T2D in both Asians and Europeans (all
 250 $P_{\text{het}} > 0.05$), with the strongest association observed for cholesterol in large VLDL
 251 particles (Asians: OR 1.41, 95% CI 1.29 - 1.53; Europeans: OR 1.29, 95% CI 1.22 -
 252 1.37; $P_{\text{het}} = 0.10$). Because the lipid components (total lipids, phospholipids,
 253 cholesterol esters, free cholesterol, triglycerides) in VLDL particles were highly
 254 correlated (**Supplemental Tables 1 to 4**), particularly in extremely large VLDL, very

large VLDL, large VLDL and medium VLDL (Pearson's $r > 0.85$), these lipid components were similarly associated with T2D risk as VLDL cholesterol.

Total high-density lipoprotein (HDL) cholesterol was inversely associated with T2D risk in Asians (OR 0.85, 95% CI 0.76 - 0.94) and Europeans (OR 0.76, 95% CI 0.69 - 0.82; $P_{\text{het}} = 0.09$), although this did not reach Bonferroni-corrected significance in Asians. These inverse associations appeared to be driven by cholesterol in very large and large HDL particles, which were stronger in Europeans than in Asians (very large HDL: $P_{\text{het}} = 0.02$; large HDL: $P_{\text{het}} = 0.04$). Other lipid components of HDL (total lipids, phospholipids, cholesterol esters, free cholesterol) with the exception of triglycerides were similarly associated with T2D risk as the cholesterol component of HDL. There were no significant associations between intermediate-density lipoprotein (IDL) cholesterol and low-density lipoprotein (LDL) cholesterol concentrations and T2D risk.

Elevated concentrations of total triglycerides were significantly associated with a higher risk of T2D in Asians (OR 1.43, 95% CI 1.31 - 1.57) and Europeans (OR 1.32, 95% CI 1.25 - 1.40; $P_{\text{het}} = 0.14$). Results were similar when we analyzed triglycerides in subclasses of VLDL, LDL, and HDL, with the exception of triglycerides in large HDL.

Other Metabolic Biomarkers

We also evaluated ketone bodies, glycerides, phospholipids, and metabolites related to glycolysis and fluid balance (**Table 2**). Higher phosphoglyceride concentrations were associated with T2D risk but this remained significant after Bonferroni

correction only in Asians (OR 1.18, 95% CI 1.08 - 1.30) and not in Europeans (OR 1.08, 95% CI 1.01, 1.16; $P_{\text{het}} = 0.14$). Similarly, higher albumin concentrations were significantly associated with a higher risk of T2D in Asians (OR 1.23, 95% CI: 1.12 - 1.35), but not Europeans (OR 1.12, 95% CI 0.97 - 1.28; $P_{\text{het}} = 0.25$).

284

285 *Joint Effects of Metabolic Biomarkers on Risk of T2D*

Using the correlation-pruning method, 12 metabolic biomarkers in MEC were significantly associated with T2D with pairwise correlations with all other biomarkers <0.80 . These were isoleucine, valine, phenylalanine, tyrosine, alanine, GlycA, unsaturation index, proportion of n-6 PUFA, proportion of MUFA, VLDL size, small VLDL particle concentration, and very large HDL phospholipids. Stepwise regression selected a multivariable model that simultaneously included isoleucine, GlycA, unsaturation index, and very large HDL phospholipids (**Supplemental Table 6**). Higher levels of isoleucine (Asians: OR 1.25, 95% CI 1.10 - 1.42; Europeans: 1.18, 95% CI 1.06 - 1.31) and GlycA (Asians: OR 1.17, 95% CI 1.03 - 1.33; Europeans: OR 1.06, 95% CI 0.95 - 1.18) were associated with higher T2D risk, and very large HDL phospholipids with lower T2D risk (Asians: OR 0.88, 95% CI 0.77 - 0.99; Europeans: OR 0.70, 95% CI 0.62 - 0.78). There was no independent significant association between unsaturation index and T2D risk. We observed evidence of heterogeneity between the Asian and European populations only for very large HDL phospholipids ($P_{\text{het}} = 0.01$).

301

302 **Discussion**

We quantified 154 metabolic biomarkers using NMR technology in plasma samples from five prospective cohort studies representing three major ethnic groups in Asia

305 (Chinese, Malay, and Indian) and a European population. Branched-chain amino
306 acids, aromatic amino acids, alanine, the inflammatory marker GlycA, total fatty
307 acids, the proportion of MUFA, apolipoprotein B, larger VLDL particle sizes and
308 triglycerides were consistently associated with a higher T2D risk in Asians and
309 Europeans. Furthermore, the proportion of n-6 PUFA and larger HDL particle sizes
310 were consistently associated with a lower T2D risk in Asians and Europeans. Overall,
311 associations were consistent across the Asian ethnic groups, and largely consistent for
312 the Asian and European populations.

313

314 Our finding that higher concentrations of BCAAs (leucine, isoleucine and valine)
315 were associated with a higher risk of T2D is consistent with the results of previous
316 prospective studies (13, 23, 25-30). Results from Mendelian randomization studies
317 suggest that BCAA metabolites play a causal role in the pathogenesis of T2D (29). In
318 line with our results, the aromatic amino acids phenylalanine and tyrosine were
319 directly associated with T2D risk in a meta-analysis of prospective studies (13), and
320 alanine was directly associated with T2D risk in a Japanese cohort (26). Elevated
321 concentrations of aromatic amino acids and alanine have been hypothesized to induce
322 insulin resistance, possibly by inhibition of glucose transport and phosphorylation in
323 skeletal muscle (13, 31, 32).

324

325 A higher concentration of the inflammatory marker GlycA was associated with a
326 higher risk of T2D in our study, which is compatible with results from the Dutch
327 PREVEND study and the U.S. Women's Health Study (33, 34). GlycA is a composite
328 NMR measure that arises from the *N*-acetyl methyl group protons of oligosaccharide
329 moieties of acute-phase proteins (35), proteins whose concentrations change in

330 response to inflammation. As opposed to CRP, which is characterized by multi-fold
331 elevation in acute state, GlycA levels fluctuate less on short-term bases and may
332 therefore capture different aspects of the inflammatory response (36), a hypothesis
333 supported by our results and results of the PREVEND study as GlycA remains
334 directly associated with T2D risk after adjusting for CRP (33). However, in the
335 Women's Health Study, the association between GlycA and incident T2D was
336 significantly attenuated after adjusting for CRP (34). It is plausible that low-grade
337 inflammation is involved in the pathogenesis of T2D through insulin resistance and
338 beta-cell dysfunction (37, 38).

339

340 Our report of a direct association between plasma MUFA and T2D risk is consistent
341 with a previous finding in a Finnish cohort (39), but not with the Atherosclerosis Risk
342 in Communities (ARIC) study (40). Because MUFA can be synthesized endogenously
343 from SFA (41), plasma MUFAs concentrations may reflect SFA intake. We also
344 observed a direct association between plasma SFA and T2D risk although this was not
345 significant in Asians. In the ARIC study (40) and a Finnish cohort (39), a higher
346 proportion of SFA was also associated with higher T2D risk. It has been hypothesized
347 that high plasma SFAs promote insulin resistance and are lipotoxic to beta-cells (42,
348 43). Our study further showed that higher proportions of n-6 PUFAs were associated
349 with a lower risk of T2D, concordant with a previous meta-analysis of prospective
350 cohort studies (44). The association may be partly mediated by increasing insulin
351 sensitivity through increasing cellular membrane fluidity (45, 46), particularly of
352 skeletal muscle cells and hepatocytes (45, 46), and by acting as ligands for
353 peroxisome proliferator-activated receptor gamma (47, 48). These hypotheses are
354 supported by a meta-analysis of randomized feeding trials that showed improvements

355 in glucose-insulin homeostasis for dietary PUFA compared to SFA or MUFA (49). In
356 a recent Mendelian randomization study, genetic predisposition to higher levels of
357 linoleic acid was associated with a lower risk of T2D (50), supporting a causal effect
358 of linoleic acid on T2D. The lack of association between the relative fraction of n-3
359 PUFA and T2D risk is consistent with a previous meta-analysis of prospective studies
360 (51).

361

362 We also identified several patterns in the lipoprotein profile that were consistently
363 associated with T2D risk across ethnic groups. Larger VLDL particles were directly
364 associated with T2D risk, similar to existing evidence from prospective studies (23,
365 52-54). VLDL is mostly comprised of triglycerides, and larger particles carry more
366 triglycerides (54). The causal relationship between plasma triglycerides and T2D is
367 still a matter of debate (55) and there was no consensus among studies on genetically
368 influenced triglyceride levels and T2D incidence (56, 57). Our finding that the
369 concentration of larger HDL particles was inversely associated with T2D risk also
370 agrees with previous prospective studies in western populations (23, 52-54). While
371 HDL lipid components (e.g. phospholipids, cholesterol esters, free cholesterol) in very
372 large and large HDL particles were associated with lower risk of T2D, they were not
373 associated with T2D risk in medium and small HDL particles. This suggests that the
374 inverse associations observed for HDL lipid components were driven by particle size
375 rather than lipid composition. Larger HDL particles have been hypothesized to be
376 more efficient in cholesterol efflux capacity than smaller HDL particles (58, 59). HDL
377 particles may also have other anti-diabetic properties such as promoting insulin
378 secretion and glucose uptake in skeletal muscle cells (60).

379

380 Our study's strengths included the use of the same highly reproducible metabolomics
381 platform in prospective cohort studies comprising Asian and European ethnic groups,
382 which allowed us to compare the results across diverse populations. However, we also
383 acknowledge certain limitations. First, a limitation of NMR technology compared to
384 MS is its lower sensitivity, and we could not quantify metabolic biomarkers that were
385 below the detection limit of NMR. Second, participants of the SEED cohorts were not
386 required to fast and this may have affected levels of certain metabolic biomarkers
387 such as triglycerides and amino acids (61, 62). However, this would have more likely
388 weakened than strengthened the observed associations. Third, different T2D case
389 ascertainment were used for the cohorts and this may have affected comparability of
390 the results. However, despite different methodologies in case ascertainment, we
391 obtained consistent results across the different cohorts. Finally, we cannot rule out
392 residual confounding due to unmeasured or imperfectly measured risk factors.

393

394 The metabolic biomarkers associated with T2D risk in our study are modifiable and
395 may thus provide targets for lifestyle and pharmacological interventions. For example,
396 replacing dietary SFA with PUFA has long been part of dietary recommendation to
397 reduce cardiovascular disease risk (63) and may also reduce T2D risk by improving
398 plasma fatty acid profiles. Similarly, dietary BCAAs and aromatic amino acids are
399 primarily from animal products, such as meat, and lower consumption may reduce
400 T2D risk partly by improving circulating amino acids (64-66). Furthermore,
401 cholesteryl ester transfer protein inhibitors are a drug class that is known to
402 preferentially increase larger HDL particles (67, 68). In a meta-analysis of four
403 randomized trials, these drugs significantly reduced T2D risk although it is unclear if
404 this effect is caused solely by changes in circulating HDL (67-69). Because these

405 metabolic biomarkers reflect modifiable pathways, metabolic profiling data could
406 inform ‘precision medicine’ approaches that prioritize specific preventive
407 interventions for individuals at high risk of T2D.

408

409 Several of our findings, including associations between GlycA and HDL particle size
410 with T2D risk are novel for Asian populations. The consistent associations for these
411 metabolic biomarkers across diverse ethnic groups suggest that differences in the
412 degree of metabolic disturbances in the evaluated pathways rather than differences in
413 the pathways involved may contribute to ethnic differences in risk of T2D. For
414 instance, South Asians have higher rates of T2D compared to European ancestry
415 populations (70). The smaller mean HDL particle sizes in South Asians as compared
416 with European ancestry populations (71) may contribute to the higher risk of T2D in
417 South Asians possibly resulting from less efficient reverse cholesterol transport and
418 dyslipidemia (58). Our observation of possibly stronger inverse associations between
419 HDL and T2D risk in Europeans than Asians, particularly for very large HDL, is
420 novel and warrants further studies.

421

422 We found consistent associations between metabolic signatures and risk of T2D in
423 prospective cohorts in Asia and Europe. Metabolic aberrations that spanned multiple
424 pathways, including amino acids, inflammation, fatty acids and lipoproteins were
425 similarly associated with T2D risk in major Asian ethnic groups and Europeans
426 despite environmental, lifestyle, and physiological differences. Our results suggest
427 that these are candidates for better prediction of T2D in different ethnic groups that
428 require further evaluation in prognostic studies. Moreover, these metabolic

429 biomarkers are modifiable and may inform individualized interventions for the
430 prevention of T2D.

431

432 **Acknowledgements**

433 All authors have no conflicts of interest.

434

435 CS, SN, TYW, C-YC, PJ, AL, MP, VS, EST, PW, RMvD and XS and conceived and
436 designed research. AC and PW conducted experiments. CS, SN, TYW, C-YC, PJ,
437 AL, MP, VS, EST, PW and RMvD contributed data. JYHS, YH, AC and XS analyzed
438 data. JYHS, RMvD and XS wrote the manuscript and had primary responsibility for
439 final content. All authors read and approved the manuscript.

440

441 The authors thank all investigators, staff members, and study participants for their
442 contributions to all the participating studies. The MEC is supported by grants from the
443 Singapore Ministry of Health, including National Medical Research Council Large
444 Collaborative Grant (MOH-000271-00), National University of Singapore and
445 National University Health System, Singapore. Metabolic profiling for MEC was
446 supported by NUHS Summit Research Program (Metabolic Diseases) and performed
447 by Nightingale Health Ltd. The SEED cohorts are supported by the National Medical
448 Research Council (NMRC), Singapore (grants 0796/2003, 1176/2008, 1149/2008,
449 STaR/0003/2008, 1249/2010, CG/SERI/2010, CIRG/1371/2013, CIRG/1417/2015,
450 CIRG/1488/2018 and OFLCG/004a/2018), and Biomedical Research Council
451 (BMRC), Singapore (08/1/35/19/550 and 09/1/35/19/616). VS was supported by the
452 Finnish Foundation for Cardiovascular Research. JYHS was supported by the

453 Singapore Ministry of Health's National Medical Research Council Large
454 Collaborative Grant (MOH-000271-00).

455

456 **Disclosures**

457 A. Cichońska and P. Wuertz are employees and shareholders of Nightingale
458 Health Ltd., a company offering NMR-based biomarker profiling.

459 B. VS has received honoraria for consulting for Novo Nordisk and Sanofi. He
460 also has ongoing research collaboration with Bayer Ltd (All unrelated to the
461 present study).

462

463 **Data availability**

464 Data described in the manuscript, code book and analytic code will be made available
465 upon request pending application and approval.

References

1. Nanditha A, Ma RCW, Ramachandran A, Snehalatha C, Chan JCN, Chia KS, Shaw JE, Zimmet PZ. Diabetes in Asia and the Pacific: Implications for the Global Epidemic. *Diabetes Care*. 2016 Mar;39:472-85.
2. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of Type 2 Diabetes - Global Burden of Disease and Forecasted Trends. *Journal of Epidemiology and Global Health*. 2020 Mar;10:107-11.
3. Uusitupa M, Khan TA, Vigiliouk E, Kahleova H, Rivelles AA, Hermansen K, Pfeiffer A, Thanopoulou A, Salas-Salvado J, Schwab U, et al. Prevention of Type 2 Diabetes by Lifestyle Changes: A Systematic Review and Meta-Analysis. *Nutrients*. 2019 Nov;11.
4. Gillies CL, Abrams KR, Lambert PC, Cooper NJ, Sutton AJ, Hsu RT, Khunti K. Pharmacological and lifestyle interventions to prevent or delay type 2 diabetes in people with impaired glucose tolerance: systematic review and meta-analysis. *British Medical Journal*. 2007 Feb;334:299-302B.
5. Sheng Z, Cao JY, Pang YC, Xu HC, Chen JW, Yuan JH, Wang R, Zhang CS, Wang LX, Dong J. Effects of Lifestyle Modification and Anti-diabetic Medicine on Prediabetes Progress: A Systematic Review and Meta-Analysis. *Frontiers in Endocrinology*. 2019 Jul;10.
6. Kodama K, Tojjar D, Yamada S, Toda K, Patel CJ, Butte AJ. Ethnic Differences in the Relationship Between Insulin Sensitivity and Insulin Response A systematic review and meta-analysis. *Diabetes Care*. 2013 Jun;36:1789-96.

7. Chan JCN, Malik V, Jia WP, Kadowaki T, Yajnik CS, Yoon KH, Hu FB. Diabetes in Asia Epidemiology, Risk Factors, and Pathophysiology. *Jama-Journal of the American Medical Association*. 2009 May;301:2129-40.
8. Philipson LH. Harnessing heterogeneity in type 2 diabetes mellitus. *Nature Reviews Endocrinology*. 2020 Feb;16:79-80.
9. Prasad RB, Groop L. Precision medicine in type 2 diabetes. *Journal of Internal Medicine*. 2019 Jan;285:40-8.
10. Palmnäs M, Brunius C, Shi L, Rostgaard-Hansen A, Torres NE, González-Domínguez R, Zamora-Ros R, Ye YL, Halkjær J, Tjønneland A, et al. Perspective: Metabotyping-A Potential Personalized Nutrition Strategy for Precision Prevention of Cardiometabolic Disease. *Adv Nutr*. 2020 05;11:524-32.
11. Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. *Nature Reviews Drug Discovery*. 2016 Jul;15:473-84.
12. Putri SP, Yamamoto S, Tsugawa H, Fukusaki E. Current metabolomics: Technological advances. *Journal of Bioscience and Bioengineering*. 2013 Jul;116:9-16.
13. Guasch-Ferre M, Hruby A, Toledo E, Clish CB, Martinez-Gonzalez MA, Salas-Salvado J, Hu FB. Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis. *Diabetes Care*. 2016 May;39:833-46.
14. Gao H, Salim A, Lee J, Tai ES, van Dam RM. Can body fat distribution, adiponectin levels and inflammation explain differences in insulin resistance between ethnic Chinese, Malays and Asian Indians? *International Journal of Obesity*. 2012 Aug;36:1086-93.

15. Markley JL, Bruschweiler R, Edison AS, Eghbalnia HR, Powers R, Raftery D, Wishart DS. The future of NMR-based metabolomics. *Current Opinion in Biotechnology*. 2017 Feb;43:34-40.
16. Marshall DD, Powers R. Beyond the paradigm: Combining mass spectrometry and nuclear magnetic resonance for metabolomics. *Progress in Nuclear Magnetic Resonance Spectroscopy*. 2017 May;100:1-16.
17. Tan KHX, Tan LWL, Sim X, Tai ES, Lee JJ, Chia KS, van Dam RM. Cohort Profile: The Singapore Multi-Ethnic Cohort (MEC) study. *Int J Epidemiol*. 2018 Feb.
18. Lavanya R, Jeganathan VSE, Zheng YF, Raju P, Cheung N, Tai ES, Wang JJ, Lamoureux E, Mitchell P, Young TL, et al. Methodology of the Singapore Indian Chinese Cohort (SICC) Eye Study: Quantifying ethnic variations in the epidemiology of eye diseases in Asians. *Ophthalmic Epidemiology*. 2009 Nov-Dec;16:325-36.
19. Rosman M, Zheng YF, Wong WL, Lamoureux E, Saw SM, Tay WT, Wang JJ, Mitchell P, Tai ES, Wong TY. Singapore Malay Eye Study: rationale and methodology of 6-year follow-up study (SiMES-2). *Clinical and Experimental Ophthalmology*. 2012 Aug;40:557-68.
20. Borodulin K, Tolonen H, Jousilahti P, Jula A, Juolevi A, Koskinen S, Kuulasmaa K, Laatikainen T, Mannisto S, Peltonen M, et al. Cohort Profile: The National FINRISK Study. *International Journal of Epidemiology*. 2018 Jun;47:696-+.
21. Pirkola SP, Isometsa E, Suvisaari J, Aro H, Joukamaa M, Poikolainen K, Koskinen S, Aromaa A, Lonnqvist JK. DSM-IV mood-, anxiety- and alcohol use disorders and their comorbidity in the Finnish general population - Results from the Health 2000 Study. *Social Psychiatry and Psychiatric Epidemiology*. 2005 Jan;40:1-10.

22. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2010 Jan;33 Suppl 1:S62-9.
23. Ahola-Olli AV, Mustelin L, Kalimeri M, Kettunen J, Jokelainen J, Auvinen J, Puukka K, Havulinna AS, Lehtimäki T, Kivimäki 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 Dec;62:2298-309.
24. Soininen P, Kangas AJ, Wurtz P, Suna T, Ala-Korpela M. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and Genetics. *Circulation-Cardiovascular Genetics*. 2015 Feb;8:192-206.
25. Flores-Guerrero JL, Oste MCJ, Kieneker LM, Gruppen EG, Wolak-Dinsmore J, Otvos JD, Connelly MA, Bakker SJL, Dullaart RPF. Plasma Branched-Chain Amino Acids and Risk of Incident Type 2 Diabetes: Results from the PREVEND Prospective Cohort Study. *Journal of Clinical Medicine*. 2018 Dec;7.
26. Chen SM, Akter S, Kuwahara K, Matsushita Y, Nakagawa T, Konishi M, Honda T, Yamamoto S, Hayashi T, Noda M, et al. Serum amino acid profiles and risk of type 2 diabetes among Japanese adults in the Hitachi Health Study. *Scientific Reports*. 2019 May;9.
27. Satheesh G, Ramachandran S, Jaleel A. Metabolomics-Based Prospective Studies and Prediction of Type 2 Diabetes Mellitus Risks. *Metabolic Syndrome and Related Disorders*. 2020 Feb;18:1-9.
28. Lu YH, Wang YL, Ong CN, Subramaniam T, Choi HW, Yuan JM, Koh WP, Pan A. Metabolic signatures and risk of type 2 diabetes in a Chinese population: an untargeted metabolomics study using both LC-MS and GC-MS. *Diabetologia*. 2016 Nov;59:2349-59.

29. Lotta LA, Scott RA, Sharp SJ, Burgess S, Luan JA, Tillin T, Schmidt AF, Imamura F, Stewart ID, Perry JRB, et al. Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. *Plos Medicine*. 2016 Nov;13.
30. Tillin T, Hughes AD, Wang Q, Wurtz P, Ala-Korpela M, Sattar N, Forouhi NG, Godsland IF, Eastwood SV, McKeigue PM, et al. Diabetes risk and amino acid profiles: cross-sectional and prospective analyses of ethnicity, amino acids and diabetes in a South Asian and European cohort from the SABRE (Southall And Brent REvisited) Study. *Diabetologia*. 2015 May;58:968-79.
31. Wurtz P, Soininen P, Kangas AJ, Ronnema T, Lehtimäki T, Kahonen M, Viikari JS, Raitakari OT, Ala-Korpela M. Branched-Chain and Aromatic Amino Acids Are Predictors of Insulin Resistance in Young Adults. *Diabetes Care*. 2013 Mar;36:648-55.
32. Krebs M, Krssak M, Bernroider E, Anderwald C, Brehm A, Meyerspeer M, Nowotny P, Roth E, Waldhausl W, Roden M. Mechanism of amino acid-induced skeletal muscle insulin resistance in humans. *Diabetes*. 2002 Mar;51:599-605.
33. Connelly MA, Gruppen EG, Wolak-Dinsmore J, Matyus SP, Riphagen IJ, Shalaurova I, Bakker SJL, Otvos JD, Dullaart RPF. GlycA, a marker of acute phase glycoproteins, and the risk of incident type 2 diabetes mellitus: PREVEND study. *Clinica Chimica Acta*. 2016 Jan;452:10-7.
34. Akinkuolie AO, Pradhan AD, Buring JE, Ridker PM, Mora S. Novel protein glycan side-chain biomarker and risk of incident type 2 diabetes mellitus. *Arterioscler Thromb Vasc Biol*. 2015 Jun;35:1544-50.

35. Connelly MA, Otvos JD, Shalaurova I, Playford MP, Mehta NN. GlycA, a novel biomarker of systemic inflammation and cardiovascular disease risk. *Journal of Translational Medicine*. 2017 Oct;15.
36. Ritchie SC, Wurtz P, Nath AP, Abraham G, Havulinna AS, Fearnley LG, Sarin AP, Kangas AJ, Soininen P, Aalto K, et al. The Biomarker GlycA Is Associated with Chronic Inflammation and Predicts Long-Term Risk of Severe Infection. *Cell Systems*. 2015 Oct;1:293-301.
37. Hameed I, Masoodi SR, Mir SA, Nabi M, Ghazanfar K, Ganai BA. Type 2 diabetes mellitus: From a metabolic disorder to an inflammatory condition. *World Journal of Diabetes*. 2015 May;6:598-612.
38. Jin CC, Henao-Mejia J, Flavell RA. Innate Immune Receptors: Key Regulators of Metabolic Disease Progression. *Cell Metabolism*. 2013 Jun;17:873-82.
39. Mahendran Y, Cederberg H, Vangipurapu J, Kangas AJ, Soininen P, Kuusisto J, Uusitupa M, Ala-Korpela M, Laakso M. Glycerol and Fatty Acids in Serum Predict the Development of Hyperglycemia and Type 2 Diabetes in Finnish Men. *Diabetes Care*. 2013 Nov;36:3732-8.
40. Wang L, Folsom AR, Zheng ZJ, Pankow JS, Eckfeldt JH, Investigators AS. Plasma fatty acid composition and incidence of diabetes in middle-aged adults: the Atherosclerosis Risk in Communities (ARIC) Study. *American Journal of Clinical Nutrition*. 2003 Jul;78:91-8.
41. Nakamura MT, Nara TY. Structure, function, and dietary regulation of Delta 6, Delta 5, and Delta 9 desaturases. *Annual Review of Nutrition*. 2004;24:345-76.
42. Oh YS, Bae GD, Baek DJ, Park EY, Jun HS. Fatty Acid-Induced Lipotoxicity in Pancreatic Beta-Cells During Development of Type 2 Diabetes. *Frontiers in Endocrinology*. 2018 Jul;9.

43. Sobczak AIS, Blindauer CA, Stewart AJ. Changes in Plasma Free Fatty Acids Associated with Type-2 Diabetes. *Nutrients*. 2019 Sep;11.
44. Wu JHY, Marklund M, Imamura F, Tintle N, Korat AVA, de Goede J, Zhou X, Yang WS, Otto MCD, Kroger J, et al. Omega-6 fatty acid biomarkers and incident type 2 diabetes: pooled analysis of individual-level data for 39 740 adults from 20 prospective cohort studies. *Lancet Diabetes & Endocrinology*. 2017 Dec;5:965-74.
45. Kroger J, Jacobs S, Jansen E, Fritsche A, Boeing H, Schulze MB. Erythrocyte membrane fatty acid fluidity and risk of type 2 diabetes in the EPIC-Potsdam study. *Diabetologia*. 2015 Feb;58:282-9.
46. Weijers RNM. Lipid Composition of Cell Membranes and Its Relevance in Type 2 Diabetes Mellitus. *Current Diabetes Reviews*. 2012;8:390-400.
47. Lai MC, Teng TH, Yang C. The Natural PPAR Agonist Linoleic Acid Stimulated Insulin Release in the Rat Pancreas. *Journal of Veterinary Medical Science*. 2013 Nov;75:1449-54.
48. Echeverria F, Ortiz M, Valenzuela R, Videla LA. Long-chain polyunsaturated fatty acids regulation of PPARs, signaling: Relationship to tissue development and aging. *Prostaglandins Leukotrienes and Essential Fatty Acids*. 2016 Nov;114:28-34.
49. Imamura F, Micha R, Wu JHY, Otto MCD, Otite FO, Abioye AI, Mozaffarian D. Effects of Saturated Fat, Polyunsaturated Fat, Monounsaturated Fat, and Carbohydrate on Glucose-Insulin Homeostasis: A Systematic Review and Meta-analysis of Randomised Controlled Feeding Trials. *Plos Medicine*. 2016 Jul;13.
50. Yuan S, Larsson SC. Association of genetic variants related to plasma fatty acids with type 2 diabetes mellitus and glycaemic traits: a Mendelian randomisation study. *Diabetologia*. 2020 Jan;63:116-23.

51. Wu JHY, Micha R, Imamura F, Pan A, Biggs ML, Ajaz O, Djousse L, Hu FB, Mozaffarian D. Omega-3 fatty acids and incident type 2 diabetes: a systematic review and meta-analysis. *British Journal of Nutrition*. 2012 Jun;107:S214-S27.
52. Mackey RH, Mora S, Bertoni AG, Wassel CL, Carnethon MR, Sibley CT, Goff DC. Lipoprotein Particles and Incident Type 2 Diabetes in the Multi-Ethnic Study of Atherosclerosis. *Diabetes Care*. 2015 Apr;38:628-36.
53. Hodge AM, Jenkins AJ, English DR, O'Dea K, Giles GG. NMR-determined lipoprotein subclass profile predicts type 2 diabetes. *Diabetes Research and Clinical Practice*. 2009 Jan;83:132-9.
54. Mora S, Otvos JD, Rosenson RS, Pradhan A, Buring JE, Ridker PM. Lipoprotein Particle Size and Concentration by Nuclear Magnetic Resonance and Incident Type 2 Diabetes in Women. *Diabetes*. 2010 May;59:1153-60.
55. Hölzl B, Paulweber B, Sandhofer F, Patsch JR. Hypertriglyceridemia and insulin resistance. *J Intern Med*. 1998 Jan;243:79-82.
56. De Silva NM, Freathy RM, Palmer TM, Donnelly LA, Luan J, Gaunt T, Langenberg C, Weedon MN, Shields B, Knight BA, et al. Mendelian randomization studies do not support a role for raised circulating triglyceride levels influencing type 2 diabetes, glucose levels, or insulin resistance. *Diabetes*. 2011 Mar;60:1008-18.
57. Qi QB, Liang LM, Doria A, Hu FB, Qi L. Genetic Predisposition to Dyslipidemia and Type 2 Diabetes Risk in Two Prospective Cohorts. *Diabetes*. 2012 Mar;61:745-52.
58. Khera AV, Demler OV, Adelman SJ, Collins HL, Glynn RJ, Ridker PM, Rader DJ, Mora S. Cholesterol Efflux Capacity, High-Density Lipoprotein Particle Number, and Incident Cardiovascular Events: An Analysis From the JUPITER Trial

(Justification for the Use of Statins in Prevention: An Intervention Trial Evaluating Rosuvastatin). *Circulation*. 2017 Jun;135:2494-504.

59. Rysz J, Gluba-Brzozka A, Rysz-Gorzynska M, Franczyk B. The Role and Function of HDL in Patients with Chronic Kidney Disease and the Risk of Cardiovascular Disease. *International Journal of Molecular Sciences*. 2020 Jan;21.
60. Drew BG, Duffy SJ, Formosa MF, Natoli AK, Henstridge DC, Penfold SA, Thomas WG, Mukhamedova N, de Courten B, Forbes JM, et al. High-Density Lipoprotein Modulates Glucose Metabolism in Patients With Type 2 Diabetes Mellitus. *Circulation*. 2009 Apr;119:2103-U134.
61. Jackson KG, Poppitt SD, Minihane AM. Postprandial lipemia and cardiovascular disease risk: Interrelationships between dietary, physiological and genetic determinants. *Atherosclerosis*. 2012 Jan;220:22-33.
62. Ottosson F, Ericson U, Almgren P, Nilsson J, Magnusson M, Fernandez C, Melander O. POSTPRANDIAL LEVELS OF BRANCH CHAINED AND AROMATIC AMINO ACIDS ASSOCIATED WITH FASTING GLYCAEMIA. *Journal of Hypertension*. 2016 Sep;34:E145-E.
63. Sacks FM, Lichtenstein AH, Wu JHY, Appel LJ, Creager MA, Kris-Etherton PM, Miller M, Rimm EB, Rudel LL, Robinson JG, et al. Dietary Fats and Cardiovascular Disease A Presidential Advisory From the American Heart Association. *Circulation*. 2017 Jul;136:E1-+.
64. Isanejad M, LaCroix AZ, Thomson CA, Tinker L, Larson JC, Qi QB, Qi LH, Cooper-DeHoff RM, Phillips LS, Prentice RL, et al. Branched-chain amino acid, meat intake and risk of type 2 diabetes in the Women's Health Initiative. *British Journal of Nutrition*. 2017 Jun;117:1523-30.

65. Elshorbagy A, Jerneren F, Basta M, Basta C, Turner C, Khaled M, Refsum H. Amino acid changes during transition to a vegan diet supplemented with fish in healthy humans. *European Journal of Nutrition*. 2017 Aug;56:1953-62.
66. Gunther SH, Khoo CM, Sim X, Tai ES, van Dam RM. Diet, Physical Activity and Adiposity as Determinants of Circulating Amino Acid Levels in a Multiethnic Asian Population. *Nutrients*. 2020 Aug;12.
67. Krauss RM, Wojnooski K, Orr J, Geaney JC, Pinto CA, Liu Y, Wagner JA, Luk JM, Johnson-Levonas AO, Anderson MS, et al. Changes in lipoprotein subfraction concentration and composition in healthy individuals treated with the CETP inhibitor anacetrapib. *Journal of Lipid Research*. 2012 Mar;53:540-7.
68. Kontush A. HDL particle number and size as predictors of cardiovascular disease. *Frontiers in Pharmacology*. 2015 Oct;6.
69. Masson W, Lobo M, Siniawski D, Huerin M, Molinero G, Valero R, Nogueira JP. Therapy with cholesteryl ester transfer protein (CETP) inhibitors and diabetes risk. *Diabetes & Metabolism*. 2018 Dec;44:508-13.
70. Narayan KMV, Kanaya AM. Why are South Asians prone to type 2 diabetes? A hypothesis based on underexplored pathways. *Diabetologia*. 2020 Jun;63:1103-9.
71. Bilen O, Kamal A, Virani SS. Lipoprotein abnormalities in South Asians and its association with cardiovascular disease: Current state and future directions. *World Journal of Cardiology*. 2016 Mar;8:247-57.

Legends for figures

Figure 1. Associations between baseline circulating lipoprotein measures and risk of incident type 2 diabetes. Values are ORs (95% CIs) per 1 SD increment in metabolic biomarkers levels. The results for Asian participants (n=6393) were meta-analysed using data from Chinese (n=2676), Malay (n=1872) and Indian (n=1845) participants of the MEC and the SEED cohorts which consist of the Singapore Chinese Eye Study, Singapore Malay Eye Study and the Singapore Indian Eye Study. Results for the European participants (n=14558) were meta-analysed from 3 prospective Finnish cohorts (FINRISK 2002, FINRISK 2012 and Health 2000). We adjusted or matched for age, age² (only for MEC), sex, BMI and fasting glucose (MEC and Health 2000) or HbA1c (SEED cohorts, FINRISK 2002 and FINRISK 2012). We accounted for multiple testing using Bonferroni correction and *P*-values < 0.0003 were considered statistically significant. Abbreviations: ApoA1, apolipoprotein A1, ApoB, apolipoprotein B, HbA1c, glycated hemoglobin A1c, HDL, high-density lipoprotein, IDL, intermediate-density lipoprotein, LDL, low-density lipoprotein, MEC, Multiethnic Cohort, SEED, Singapore Epidemiology of Eye Diseases, VLDL, very low-density lipoprotein

Table 1. Baseline characteristics of participants of the cohorts included in the study¹

Characteristic	MEC			SEED ²			European cohorts		
	Chinese	Malay	Indian	Chinese	Malay	Indian	FINRISK 2002	FINRISK 2012	Health 2000
No. cases / No. controls	256/499	216/411	222/405	96/1825	132/1113	163/1055	350/3904	78/4475	627/6179
Follow-up duration (y)	7.4 ± 2.5	6.7 ± 2.5	6.6 ± 2.3	5.8 ± 0.9	7.1 ± 0.9	6.0 ± 0.9	13.8 ± 0.1	3.8 ± 0.1	15.1 ± 0.1
Age at interview (y)	53.6 ± 9.5	49.1 ± 9.9	47.8 ± 10.4	57 ± 8.6	54.8 ± 10	54.4 ± 8.4	48.5 ± 13.7	48.8 ± 13.8	53.2 ± 14.2
Male [n (%)]	356 (47.1)	234 (37.3)	257 (41)	889 (46.3)	563 (45.2)	568 (46.6)	1794 (45.9)	2088 (46.7)	2748 (44.5)
Education level ³									
Low	243 (32.2)	208 (33.2)	199 (31.8)	905 (47.1)	773 (62.2)	588 (48.4)	-	-	-
Intermediate	415 (55)	404 (64.5)	356 (56.9)	818 (42.6)	455 (36.6)	478 (39.3)	-	-	-
High	97 (12.8)	14 (2.2)	71 (11.3)	198 (10.3)	15 (1.2)	149 (12.3)	-	-	-
Body mass index (kg/m ²)	23.8 ± 3.7	26.9 ± 5	26.7 ± 4.9	23.4 ± 3.5	26 ± 4.7	25.7 ± 4.2	26.6 ± 4.5	26.6 ± 4.6	26.8 ± 4.6
Fasting glucose (mmol/L)	4.9 ± 0.6	5 ± 0.6	5.1 ± 0.6	-	-	-	-	-	5.4 ± 0.8
HbA1c (%) ⁴	5.6 ± 0.4	5.6 ± 0.4	5.8 ± 0.4	5.8 ± 0.3	5.7 ± 0.4	5.7 ± 0.4	5.4 ± 0.4	5.2 ± 0.3	-
Total cholesterol (mmol/L)	5.3 ± 0.9	5.6 ± 1	5.3 ± 0.9	5.6 ± 1	5.7 ± 1.1	5.5 ± 1	5.6 ± 1.1	5.4 ± 1.0	6.0 ± 1.1
HDL cholesterol (mmol/L)	1.4 ± 0.3	1.3 ± 0.3	1.1 ± 0.3	1.4 ± 0.4	1.4 ± 0.3	1.1 ± 0.3	1.5 ± 0.4	1.5 ± 0.4	1.4 ± 0.4
LDL cholesterol (mmol/L)	3.4 ± 0.8	3.6 ± 0.9	3.5 ± 0.8	3.4 ± 0.9	3.6 ± 1	3.6 ± 0.8	3.4 ± 0.9	3.3 ± 0.9	3.8 ± 1.1

Triglycerides (mmol/L)	1.2 (0.9,1.7)	1.2 (0.9,1.7)	1.3 (0.9,1.8)	1.4 (1.0,2.1)	1.3 (0.7,2.1)	1.6 (1.1,2.3)	1.1 (0.8, 1.6)	1.1 (0.8, 1.6)	1.3 (1.0, 1.8)
DBP (mmHg)	78.7 ± 11.3	76.4 ± 11.7	74.5 ± 11.5	77.6 ± 9.9	79.2 ± 11.0	77.9 ± 10.1	78.7 ± 11.4	81.3 ± 11.1	82.0 ± 11.0
SBP (mmHg)	134.2 ± 20.6	131.9 ± 20	125.3 ± 21.4	133.4 ± 18.4	141 ± 22.3	130.9 ± 18.7	135.0 ± 20.2	132.7 ± 18.6	134.2 ± 20.7

¹ Mean ± SD; median (Q1, Q3)

² SEED cohorts consist of a collection of three cohorts: Singapore Chinese Eye Study, Singapore Malay Eye Study and the Singapore Indian Eye Study. Participants of the SEED cohorts were not required to fast. LDL cholesterol was measured in the Singapore Chinese Eye Study and Singapore Indian Eye Study, and calculated in the Singapore Malay Eye Study.

³ We classified education level as follows: low – none, Primary or equivalent, intermediate – Secondary, diploma, Junior College, high school or equivalent, and high – college degree and above. The data for European cohorts could not be accessed.

⁴ Only a subset of participants in MEC (N = 598, 531, 516, for Chinese, Malay, and Indian) had HbA1c measurements. To convert HbA1c from % to mmol/mol, first multiply HbA1c (%) by 10.93 and then subtract 23.50.

Abbreviations: DBP, diastolic blood pressure, HbA1c, glycated hemoglobin A1c, HDL, high-density lipoprotein, LDL, low-density lipoprotein, MEC, Multiethnic Cohort, SEED, SBP, systolic blood pressure, Singapore Epidemiology of Eye Diseases, Q1, quantile 1, Q3, quantile 3

	Chinese	Malay	Indian	Meta-analysis of Asians			Meta-analysis of Europeans			P-het
Metabolite	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	P-value	P-het	OR (95% CI)	P-value	P-het	(Asians and Europeans)
Branched-chain amino acids										
Leucine	1.68 (1.41 - 2.00)	1.44 (1.19 - 1.73)	1.38 (1.16 - 1.64)	1.50 (1.35 - 1.66)	<0.0003	0.26	1.36 (1.27 - 1.46)	<0.0003	0.55	0.14
Isoleucine	1.68 (1.42 - 1.98)	1.43 (1.19 - 1.72)	1.33 (1.12 - 1.58)	1.48 (1.34 - 1.63)	<0.0003	0.15	1.36 (1.27 - 1.45)	<0.0003	0.76	0.17
Valine	1.54 (1.30 - 1.83)	1.35 (1.14 - 1.61)	1.36 (1.14 - 1.61)	1.42 (1.28 - 1.57)	<0.0003	0.48	1.25 (1.16 - 1.35)	<0.0003	0.39	0.05
Aromatic amino acids										
Phenylalanine	1.27 (1.07 - 1.51)	1.29 (1.04 - 1.59)	1.24 (1.04 - 1.47)	1.26 (1.14 - 1.40)	<0.0003	0.96	1.22 (1.14 - 1.31)	<0.0003	0.75	0.63
Tyrosine	1.34 (1.13 - 1.59)	1.21 (1.04 - 1.42)	1.22 (1.05 - 1.41)	1.25 (1.14 - 1.37)	<0.0003	0.64	1.21 (1.13 - 1.29)	<0.0003	0.79	0.56
Other amino acids										
Histidine	1.08 (0.91 - 1.27)	1.10 (0.91 - 1.32)	0.94 (0.80 - 1.11)	1.03 (0.93 - 1.14)	0.55	0.40	1.08 (1.00 - 1.16)	0.06	0.83	0.51
Alanine	1.35 (1.14 - 1.61)	1.15 (0.97 - 1.38)	1.27 (1.09 - 1.47)	1.26 (1.14 - 1.38)	<0.0003	0.45	1.25 (1.14 - 1.39)	<0.0003	0.19	0.96
Inflammation										
Glycoprotein acetyls	1.62 (1.38 - 1.90)	1.29 (1.05 - 1.58)	1.32 (1.15 - 1.52)	1.41 (1.28 - 1.55)	<0.0003	0.11	1.34 (1.25 - 1.43)	<0.0003	0.95	0.37
Fatty acids										

Total fatty acids	1.41 (1.22 - 1.64)	1.22 (1.03 - 1.45)	1.29 (1.12 - 1.49)	1.31 (1.20 - 1.44)	<0.0003	0.45	1.21 (1.13 - 1.29)	<0.0003	0.93	0.14
SFA (%)	1.16 (1.00 - 1.34)	1.22 (1.03 - 1.46)	1.09 (0.95 - 1.26)	1.15 (1.05 - 1.25)	0.002	0.62	1.21 (1.13 - 1.29)	<0.0003	0.56	0.36
MUFA (%)	1.52 (1.30 - 1.78)	1.48 (1.22 - 1.81)	1.32 (1.14 - 1.53)	1.43 (1.30 - 1.57)	<0.0003	0.42	1.38 (1.27 - 1.51)	<0.0003	0.28	0.63
PUFA (%)	0.66 (0.56 - 0.78)	0.67 (0.55 - 0.81)	0.78 (0.67 - 0.90)	0.71 (0.64 - 0.78)	<0.0003	0.30	0.72 (0.67 - 0.77)	<0.0003	0.48	0.78
n-6 PUFA (%)	0.64 (0.55 - 0.75)	0.68 (0.56 - 0.81)	0.76 (0.66 - 0.88)	0.70 (0.64 - 0.77)	<0.0003	0.27	0.70 (0.65 - 0.75)	<0.0003	0.52	0.96
Linoleic acid (%)	0.76 (0.65 - 0.88)	0.77 (0.65 - 0.92)	0.79 (0.68 - 0.92)	0.77 (0.71 - 0.85)	<0.0003	0.92	0.72 (0.67 - 0.77)	<0.0003	0.41	0.21
n-3 PUFA (%)	0.96 (0.83 - 1.10)	0.93 (0.78 - 1.10)	1.05 (0.90 - 1.23)	0.98 (0.89 - 1.07)	0.60	0.56	0.97 (0.91 - 1.04)	0.44	0.55	0.94
DHA (%)	0.95 (0.83 - 1.09)	0.97 (0.82 - 1.15)	1.03 (0.88 - 1.20)	0.98 (0.90 - 1.07)	0.67	0.78	0.96 (0.90 - 1.03)	0.30	0.46	0.75
Other metabolic biomarkers										
Sphingomyelins	1.01 (0.87 - 1.18)	0.92 (0.79 - 1.08)	1.07 (0.91 - 1.25)	1.00 (0.91 - 1.10)	0.98	0.44	0.90 (0.84 - 0.97)	0.007	0.70	0.08
Phosphoglycerides	1.24 (1.06 - 1.45)	1.11 (0.94 - 1.31)	1.19 (1.02 - 1.39)	1.18 (1.08 - 1.30)	<0.0003	0.60	1.08 (1.01 - 1.16)	0.03	0.67	0.14
Phosphatidylcholines	1.21 (1.04 - 1.42)	1.08 (0.92 - 1.28)	1.17 (1.01 - 1.36)	1.16 (1.06 - 1.27)	0.001	0.60	1.05 (0.98 - 1.13)	0.19	0.64	0.09
Cholines	1.15 (0.98 - 1.34)	1.03 (0.87 - 1.21)	1.14 (0.98 - 1.32)	1.11 (1.01 - 1.21)	0.03	0.58	1.02 (0.95 - 1.10)	0.63	0.71	0.16
Triglyceride by phosphoglyceride ratio	1.60 (1.37 - 1.86)	1.52 (1.24 - 1.85)	1.27 (1.11 - 1.45)	1.42 (1.30 - 1.55)	<0.0003	0.06	1.33 (1.26 - 1.42)	<0.0003	0.46	0.25
Citrate	1.15 (0.96 - 1.37)	1.05 (0.90 - 1.23)	1.38 (1.13 - 1.67)	1.16 (1.05 - 1.29)	0.003	0.11	0.91 (0.80 - 1.03)	0.13	0.11	0.00
Lactate	1.16 (0.98 - 1.37)	1.10 (0.91 - 1.34)	1.11 (0.95 - 1.29)	1.12 (1.02 - 1.24)	0.02	0.91	1.15 (0.98 - 1.35)	0.09	0.01	0.82
beta-hydroxybutyrate	1.22 (1.04 - 1.43)	1.10 (0.94 - 1.29)	1.28 (1.04 - 1.56)	1.19 (1.08 - 1.31)	0.001	0.48	0.99 (0.91 - 1.07)	0.80	0.79	0.01

Acetate	1.03 (0.84 - 1.27)	0.74 (0.50 - 1.10)	0.94 (0.81 - 1.09)	0.95 (0.85 - 1.06)	0.36	0.34	0.96 (0.88 - 1.05)	0.43	0.40	0.81
Acetoacetate	1.05 (0.91 - 1.22)	0.89 (0.74 - 1.07)	0.86 (0.73 - 1.02)	0.94 (0.86 - 1.04)	0.24	0.16	1.04 (0.96 - 1.12)	0.35	0.69	0.13
Albumin	1.28 (1.09 - 1.50)	1.26 (1.06 - 1.50)	1.17 (1.01 - 1.35)	1.23 (1.12 - 1.35)	<0.0003	0.68	1.12 (0.97 - 1.28)	0.13	0.10	0.25
Creatinine	0.96 (0.75 - 1.23)	0.71 (0.56 - 0.92)	0.96 (0.80 - 1.15)	0.89 (0.78 - 1.01)	0.07	0.14	0.97 (0.84 - 1.13)	0.72	0.06	0.36

[†]Other than fatty acids and lipoprotein measures, results for all metabolic biomarkers are shown here. For individual fatty acids, we only showed fatty acids expressed as a percentage of total fatty acids to avoid confounding by total fatty acids. All results for lipoprotein measures are displayed in Figure 1. Estimates are odds ratios (95% confidence intervals) per SD increase. The results for Chinese (n=2676), Malay (n=1872) and Indian (n=1845) participants were meta-analysed from the MEC and the SEED cohorts which consist of the Singapore Chinese Eye Study, Singapore Malay Eye Study and the Singapore Indian Eye Study. Results for the European participants (n=14558) were meta-analysed from 3 prospective Finnish cohorts (FINRISK 2002, FINRISK 2012 and Health 2000). We adjusted or matched for age, age2 (only for MEC), sex, BMI and fasting glucose (MEC and Health 2000) or HbA1c (SEED cohorts, FINRISK 2002, FINRISK 2012). We accounted for multiple testing using Bonferroni correction and P-values < 0.0003 were considered statistically significant. P-het denotes Cochran Q heterogeneity.

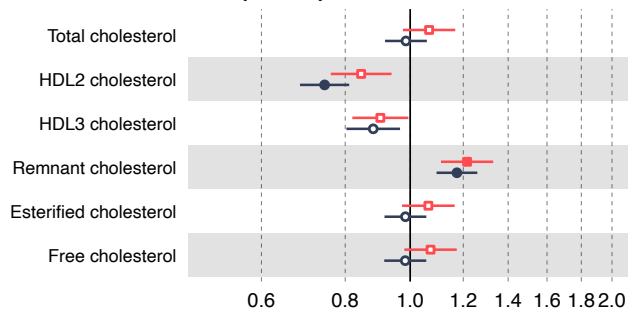
Abbreviations: HbA1c, glycated hemoglobin A1c, MEC, Multiethnic Cohort, MUFA, monounsaturated fatty acids, n-3, omega-3, n-6 - omega-6, PUFA, polyunsaturated fatty acids, SEED, Singapore Epidemiology of Eye Diseases, SFA, saturated fatty acids

Asians Finns

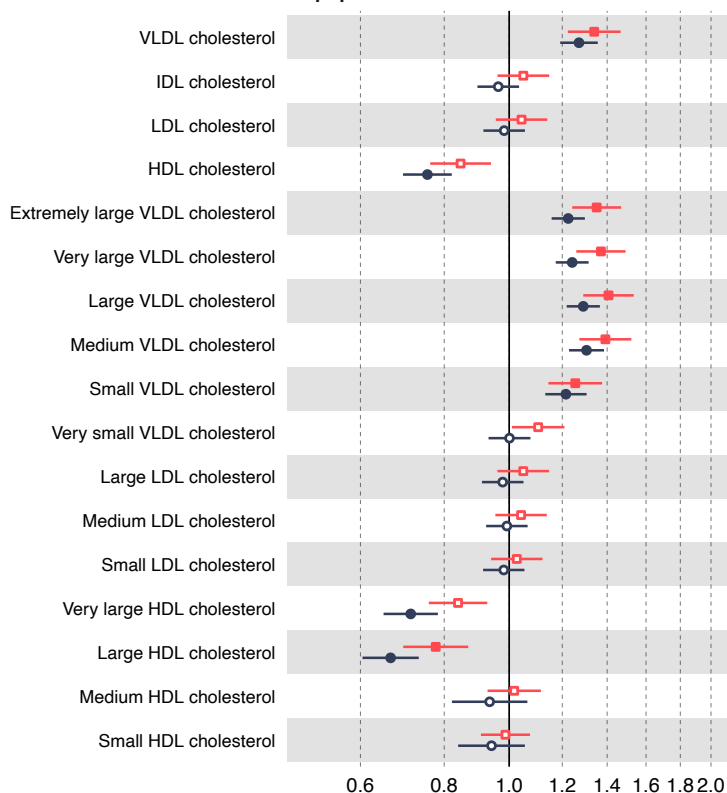
p-value < 0.0003
 p-value ≥ 0.0003

Asians Finns

Composite lipid measures

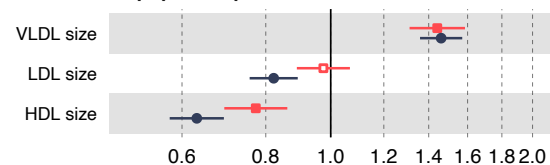


Lipoprotein cholesterol

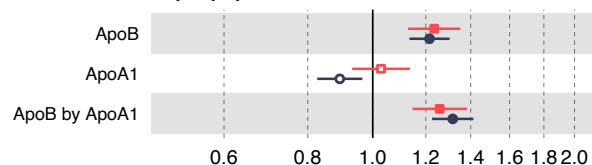


Odds ratio for incident type 2 diabetes (95% CI) per 1-SD increment in metabolite concentration

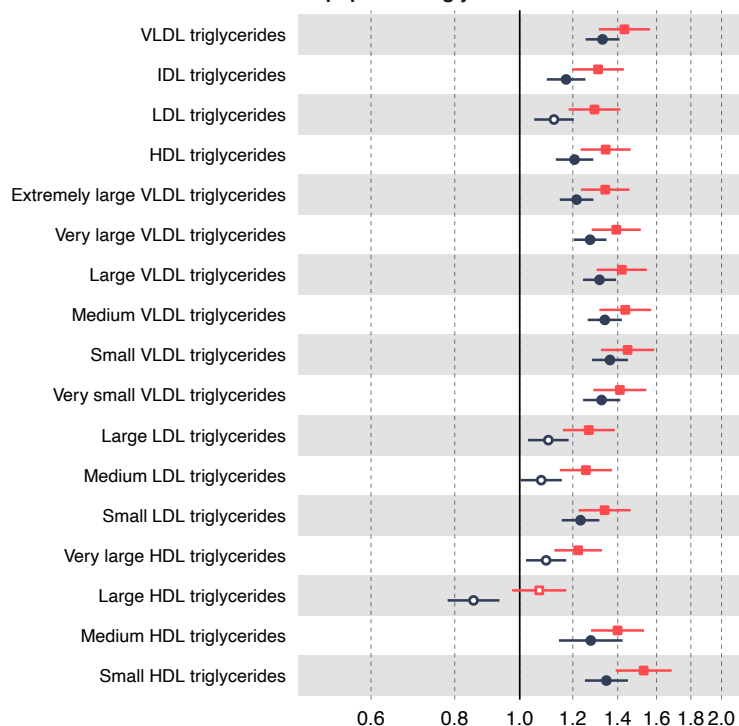
Lipoprotein particle size



Apolipoproteins



Lipoprotein triglycerides



Odds ratio for incident type 2 diabetes (95% CI) per 1-SD increment in metabolite concentration