Disentangling the relationship between depression, obesity and cardiometabolic disease

Alshehri, T.

3

Metabolomics profile in depression: a pooled analysis of 230 metabolic markers in 5,283 cases with depression and 10,145 controls

Mariska Bot, Yuri Milaneschi, Tahani Alshehri, Najaf Amin, Sanzhima Garmaeva, Gerrit L.J. Onderwater, Rene Pool, Carisha S. Thesing, Lisanne S. Vijfhuizen, Nicole Vogelzangs, Ilja C.W. Arts, Ayse Demirkan, Cornelia van Duijn, Marleen van Greevenbroek, Carla J.H. van der Kallen, Sebastian Köhler, Lannie Ligthart, Arn M.J.M. van den Maagdenberg, Dennis O. Mook-Kanamori, Renée de Mutsert, Henning Tiemeier, Miranda T. Schram, Coen D.A. Stehouwer, Gisela M. Terwindt, Ko Willems van Dijk, Jingyuan Fu, Alexandra Zhernakova, Marian Beekman, P. Eline Slagboom, Dorret I. Boomsma, Brenda W.J.H. Penninx, for the BBMRI-NL Metabolomics Consortium

Biol Psychiatry, 2020. 87(5): p. 409-418

ABSTRACT

Background
Depression has been associated with metabolic alterations, which adversely impact cardiometabolic health. Here, a comprehensive set of metabolic markers, predominantly lipids, was compared between depressed and non-depressed persons.

Methods
Nine Dutch cohorts were included, comprising 10,145 controls and 5,283 persons with depression, established with diagnostic interviews or questionnaires. A proton nuclear magnetic resonance metabolomics platform provided 230 metabolite measures: 51 lipids, fatty acids and low-molecular-weight metabolites, 98 lipid composition and particle concentration measures of lipoprotein subclasses and 81 lipid and fatty acids ratios. For each metabolite measure logistic regression analyses adjusted for sex, age, smoking, fasting status and lipid-modifying medication were performed within cohort, followed by random-effects meta-analyses.

Results
Of the 51 lipids, fatty acids and low-molecular-weight metabolites, 21 were significantly related to depression (false discovery rate $q<0.05$). Higher levels of apolipoprotein B, very-low density lipoprotein cholesterol, triglycerides, diglycerides, total and mono-unsaturated fatty acids, fatty acid chain length, glycoprotein acetyls, tyrosine, and isoleucine, and lower levels of high-density lipoprotein cholesterol, acetate, and apolipoprotein A1 were associated with increased odds of depression. Analyses of lipid composition indicators confirmed a shift towards less high-density lipoprotein cholesterol and more very-low density lipoprotein cholesterol and triglycerides particles in depression. Associations appeared generally consistent across sex, age and body mass index strata, and across cohorts with depressive diagnoses versus symptoms.

Conclusions
This large-scale meta-analysis indicates a clear distinctive profile of circulating lipid metabolites associated with depression, potentially opening new prevention or treatment avenues for depression and its associated cardiometabolic comorbidity.

INTRODUCTION

Depression imposes a huge burden on individuals and society [1]. With a high annual (6%) and lifetime (19%) prevalence, depression is among the leading contributors to global disease burden [1, 2]. It has been associated with an increased risk of somatic disease, including cardiometabolic conditions such as metabolic syndrome [3], obesity [4], diabetes mellitus [5], stroke [6], and cardiovascular disease [7], as well as an increased risk of all-cause mortality [8].

Depression is correlated with metabolic alterations in peripheral bodily systems [1]. A systematic review [9] summarizing metabolomics analyses of urine, cerebrospinal fluid, and blood samples of patients with depression highlighted a set of altered metabolites implicated in energy metabolism, neuronal integrity and transmission. Meta-analyses showed that depression was associated with increased blood levels of total cholesterol [10] and triglycerides (TG) [3], and decreased low density lipoprotein (LDL) cholesterol [11], high density lipoprotein (HDL) cholesterol [3], and omega-3 polyunsaturated fatty acids [12]. However, considerable heterogeneity was noted between studies, which was partly explained by differential lipid classifications [11].

Alterations in circulating lipid concentrations may be linked to pathophysiological pathways related to depression, such as chronic activation of the hypothalamic-pituitary-adrenal (HPA) axis or chronic low-grade inflammation [1]. Glucocorticoid-induced hypercortisolemia is known to result in lipolysis, the release of fatty acids and synthesis of very-low density lipoprotein (VLDL) [13]. Similarly, activation of the pro-inflammatory response leads to a reduction in HDL cholesterol and phospholipids, and an increase in TG, caused by the compensatory production and accumulation of phospholipid-rich VLDL [14]. In addition, omega-3 fatty acids have anti-inflammatory properties, impact HPA-axis functioning, promote cell membrane fluidity, and are involved in the regulation of dopaminergic and serotonergic neurotransmission, which can be altered in depression [15]. Alterations of circulating concentrations of lipids may also represent a consequence of depression. Patients with depression are more likely to engage in unhealthy behaviors, such as sedentariness, excessive alcohol use and poor nutrition (with preference for high palatable food rich in saturated fats), which may lead to dyslipidemia [16], that can result in metabolic syndrome and cardiovascular disease.

Emerging technologies allow high-throughput profiling of lipids and other metabolites, which has led to efforts of determining metabolic signatures of various diseases [17, 18]. A few studies have applied this to depression [19, 20], but the results remain inconsistent [21, 22]; this is partly due to different

methodologies used and different metabolites (lipids, amino acids and other small molecules) analyzed [23].

This study aims to identify plasma lipids, fatty acids and low-molecular-weight metabolites associated with depression by analyzing data from nine Dutch clinical- and population-based studies, and to assess consistency of findings across studies. A strength of the study is that all metabolites were measured around the same time with the same targeted proton nuclear magnetic resonance platform that quantifies lipids, fatty acids and low-molecular-weight metabolites, including those that have been related to consequences of depression (e.g., insulin resistance [24], onset of cardiovascular events [25], and mortality [26]).

METHODS AND MATERIALS

Sample description

Eleven datasets from nine cohorts participating in the Biobanking and BioMolecular resources Research Infrastructure-The Netherlands (BBMRI-NL) were included: Cohort on Diabetes and Atherosclerosis Maastricht (CODAM) [27], The Maastricht Study [28], Erasmus Rucphen Family study (ERF) [29], Leiden University Migraine Neuro-Analysis (LUMINA) [30], Netherlands Epidemiology of Obesity study (NEO), Netherlands Study of Depression and Anxiety (NESDA), Netherlands Twin Register (NTR) [31], the Rotterdam Study (RS), and Lifelines-DEEP (LLD) [32-34]. Both CODAM and The Maastricht Study contributed two datasets stratified by diabetes mellitus status. In total, we included 5,283 persons with depression and 10,145 control subjects (see Supplement 1 for detailed cohort descriptions). All participants provided written informed consent. Studies were approved by local ethics committees.

Measurements

Depression

The presence of depression was measured either before blood sampling or up to a maximum of one month after blood sampling. Subjects were defined as cases when meeting all the criteria required for a diagnosis of major depressive disorder (MDD) in clinical structured interviews in four cohorts, or when scoring above validated clinical cut-off score in depression questionnaires in five cohorts (see Table S1 in Supplement 1 for all instruments and definitions). In the main analyses, cases included subjects with any history of depression in lifetime.

Metabolites

Supplement 1 shows details on blood collection (for each cohort), measurement and processing of metabolite measurements. Using targeted high-throughput

proton Nuclear Magnetic Resonance metabolomics (Nightingale Health Ltd, Helsinki, Finland), 230 metabolites or metabolite ratios were reliably quantified from ethylenediamine tetraacetate plasma samples [35]. This metabolomics platform has been used in large-scaled epidemiological studies of diabetes [24], cardiovascular disease [25], mortality [26] and alcohol intake [36]. To enhance interpretation, metabolites were classified into three clusters curated by Nightingale Health [37]: 1) lipids, fatty acids and low-molecular-weight metabolites (N=51); 2) lipid composition and particle concentration measures of lipoprotein subclasses (N=98); 3) metabolite ratios (N=81). Data were processed according to a shared protocol applied also in other studies of BBMRI-NL [38]. In each cohort, values of metabolites that could not be quantified (≤5 metabolites per cohort) were set as missing for all subjects. Furthermore, metabolites values in subjects with outlying concentrations (±5 SDs) were additionally set as missing. A value of 1 was added to all metabolite values (Supplement 1 includes extensive analyses indicating that the degree of bias potentially introduced by this transformation is likely negligible) that were subsequently natural log-transformed to approximate normality. The obtained values were scaled to standard deviations units in each cohort to enable comparison.

Statistical analyses

Per-metabolite logistic regression analyses were initially performed in each dataset. The dependent variable was depression, and independent variables were the 230 metabolite measurements. For the Netherlands Twin Register cohort, logistic regression using generalized estimating equations were conducted, accounting for family-relatedness. All models were adjusted for age, sex, fasting status, use of lipid-modifying drugs listed under ATC (Anatomical Therapeutic Chemical Classification System) code C10 and smoking (Supplement 1 for measurements). All analyses were based on available data per metabolite (pair-wise deletion). Dataset-specific estimates were combined using random-effects meta-analyses (restricted maximum-likelihood estimator) to obtain pooled odds ratios (ORs). Heterogeneity of results between datasets was quantified by I^2 [39] along with 95% confidence intervals (CI) as recommended [40, 41].

As body mass index (BMI) has been shown to be associated with depression [4] and metabolites [42], we reran the main analyses adjusting for BMI. Furthermore, to investigate whether metabolic profiles were dependent on recent presence of depression, additional analyses were conducted comparing current depressed cases (depression present ±1 month around blood sampling) and controls. We conducted sensitivity analyses in which we excluded subjects using antidepressant medication (ATC code N06A), to study the impact of depression apart from its treatment. Here, we a priori expected to find a less distinctive metabolomics profile, given that antidepressant medication prescriptions are more likely in

individuals with higher depression severity. Correlations between estimates obtained from these sensitivity analyses and estimates obtained in the main analyses were computed to measure the impact of the factors considered.

Four additional sets of stratified analysis were performed to explore whether associations between metabolites and depression were different as a function of (1) depression assessment (diagnosis vs. self-report instrument), (2) sex (men vs. women), (3) age (<50 years vs. ≥50 years) and (4) BMI (normal (18.50-24.9) vs. overweight (25.0-29.9) and vs. obesity (≥30)). A Wald-test was performed to test differences in effect sizes across these strata [43], and correlations between estimates obtained across strata were estimated.

The False Discovery Rate (FDR) method [44] was applied to address multiple testing at the meta-analysis level for 230 metabolites. Meta-analyses were conducted with the 'metafor' package (version 2.0.0) in R v3.4.2-3.4.3 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Overview of cohorts

The study population comprised 15,428 adults from 11 datasets of 9 cohorts. There were 10,145 controls, and 5,283 participants with depression. Table 1 shows the characteristics of the 11 datasets. Across the cohorts, the average age ranged from 40.4-64.8 years, the proportion of women ranged from 32% to 70%, and the median prevalence of depression was 29.5%.

Table 1. Characteristics of the study populations (N=15,428)

	CODAM DM	CODAM noDM	TMS DM	TMS noDM	ERF	LUMINA	NEO	NESDA	NTR	RS	LLD
Total N	139	416	775	723	346	231	6554	2509	1523	1188	1024
Female, n (%)	46 (33.1)	168 (40.4)	248 (32.0)	455 (62.9)	198 (57.2)	136 (58.9)	3433 (52.4)	1680 (67.0)	1072 (70.4)	755 (63.6)	596 (58.2)
Age (years), mean (SD)	61.2 (6.2)	59.0 (7.1)	62.7 (7.5)	58.8 (8.0)	48.0 (14.0)	41.2 (12.2)	55.8 (6.0)	41.8 (13.0)	40.4 (13.2)	64.8 (5.8)	44.9 (13.2)
Current smoker, n (%)	26 (18.7)	86 (20.7)	122 (15.7)	94 (13.0)	127 (36.7)	25 (10.8)	1071 (16.3)	978 (39.0)	74 (4.9)	161 (13.6)	204 (19.9)
Use of lipid-modifying medications n (%)	35 (25.2)	69 (16.6)	578 (74.6)	162 (22.4)	31 (8.95)	2 (0.9)	1024 (15.6)	177 (7.0)	77 (5.1)	257 (21.6)	45 (4.4)
Fasting, n (%)	139 (100)	416 (100)	775 (100)	723 (100)	344 (99.4)	230 (99.5)	6554 (100)	2403 (95.8)	1441 (94.6)	1113 (93.7)	1013 (98.9)
BMI (kg/m^2), mean (SD)	30.3 (4.7)	28.0 (4.1)	29.8 (4.9)	29.3 (3.6)	27.2 (4.5)	23.6 (2.4)	30.1 (4.8)	25.6 (5.0)	24.7 (4.1)	27.4 (4.3)	25.2 (4.1)
No depression, n (%)	105 (75.5)	338 (81.3)	503 (64.9)	480 (66.4)	193 (55.8)	172 (74.5)	4620 (70.5)	634 (44.8)	1353 (88.8)	737 (62.0)	1010 (98.6)
Depression, n (%)	34 (24.5)	78 (18.8)	272 (35.1)	243 (33.6)	153 (44.2)	59 (25.5)	1934 (29.5)	1875 (74.7)	170 (11.2)	451 (38.0)	14 (1.4)
Of which current depression, n (%)	34 (24.5)	78 (18.8)	46 (8.4)	24 (4.8)	25 (7.2)	14 (6.1)	1934 (29.5)	782 (55.2)	N.A.	314 (26.4)	14 (1.4)
Antidepressant use, n (%)	10 (7.2)	20 (4.8)	63 (8.1)	64 (8.9)	24 (6.9)	3 (1.3)	534 (8.1)	683 (27.2)	73 (4.8)	77 (6.5)	46 (4.5)

Abbreviations: BMI = Body mass index, N.A.=not available.

Associations of 51 lipids, fatty acids and low-molecular-weight metabolites with depression

Figure 1 shows a polar plot with ORs of meta-analyses investigating associations between depression and the 51 metabolites, after adjustment for sex, age, smoking, lipid modifying drugs, and fasting status. Of these, 21 metabolites were associated with depression at FDR q<0.05 (Table 2; Figure S1 in Supplement 1). Metabolites associated with a higher odds for depression were apolipoprotein B; remnant (non-HDL and non-LDL) cholesterol, VLDL cholesterol, and mean diameter of VLDL; the glycerides and phospholipid markers diglycerides; TG in LDL, serum TG, TG in HDL, TG in VLDL, the fatty acid measures total fatty acids, monounsaturated fatty acid, and estimated fatty acid chain length; the inflammation marker glycoprotein acetyls; and the amino acids tyrosine and isoleucine. Higher levels of metabolites that were associated with a lower odds for depression were apolipoprotein A1, cholesterol content for HDL (in particular HDL_2- and HDL_3- cholesterol), and mean diameter of HDL, and ketone body acetate.

Table 2. Overview of the 21 lipids, fatty acids and various low-molecular-weight metabolites that are significantly related to depression in the pooled analysis at FDR q<0.05

Metabolite	Model 1			Model 2*		
	Pooled OR	p-value	FDR q-value	Pooled OR	p-value	FDR q-value
Apolipoproteins						
ApolipoproteinA1	0.90	2.71×10^{-7}	2.50×10^{-6}	0.94	0.007	0.021
ApolipoproteinB	1.08	2.40×10^{-4}	6.90×10^{-4}	1.05	0.014	0.040
Cholesterol						
Remnant cholesterol	1.07	0.003	0.006	1.05	0.014	0.038
VLDL cholesterol	1.10	1.68×10^{-4}	5.03×10^{-4}	1.07	0.001	0.002
HDL cholesterol	0.86	1.24×10^{-12}	9.47×10^{-11}	0.91	2.03×10^{-5}	2.59×10^{-4}
HDL_2 cholesterol	0.89	5.78×10^{-6}	2.79×10^{-5}	0.93	0.001	0.003
HDL_3 cholesterol	0.90	2.18×10^{-5}	8.37×10^{-5}	0.93	4.91×10^{-4}	0.002
Mean diameter of VLDL	1.13	1.30×10^{-6}	8.82×10^{-6}	1.08	2.39×10^{-4}	0.001
Mean diameter of HDL	0.91	2.10×10^{-4}	6.10×10^{-4}	0.96	0.104	0.222
Di- and triglycerides						
Diglycerides	1.09	2.56×10^{-5}	9.65×10^{-5}	1.07	0.003	0.008
Serum total TG	1.11	3.29×10^{-5}	1.15×10^{-4}	1.08	1.92×10^{-4}	0.001
VLDL TG	1.11	8.68×10^{-5}	2.77×10^{-4}	1.08	1.76×10^{-4}	0.001
LDL TG	1.05	0.015	0.032	1.04	0.101	0.218
HDL TG	1.09	0.007	0.015	1.07	0.029	0.072

Table 2. Continued.

Metabolite	Model 1			Model 2*		
	Pooled OR	p-value	FDR q-value	Pooled OR	p-value	FDR q-value
Fatty acids						
Mono Unsaturated FA	1.09	7.13×10^{-6}	3.35×10^{-5}	1.06	0.004	0.012
Total FA	1.05	0.013	0.027	1.03	0.102	0.219
Estimated FA chain length	1.10	0.020	0.043	1.08	0.060	0.140
Inflammation						
Glycoprotein acetyls	1.13	0.003	0.007	1.09	0.028	0.071
Ketone bodies						
Acetate	0.91	0.003	0.006	0.93	0.038	0.092
Amino acids						
Tyrosine	1.07	0.013	0.028	1.02	0.552	0.760
Isoleucine	1.14	8.26×10^{-6}	3.71×10^{-5}	1.08	0.001	0.004

Model 1: adjusted for sex, age, smoking, lipid modifying drugs, fasting status; Model 2: adjusted for model 1 and body mass index; Abbreviations: FDR=false discovery rate, FA=fatty acids, HDL=high-density lipoprotein, LDL=low-density lipoprotein, OR=odds ratio, TG=triglycerides, VLDL=very-low-density lipoprotein.

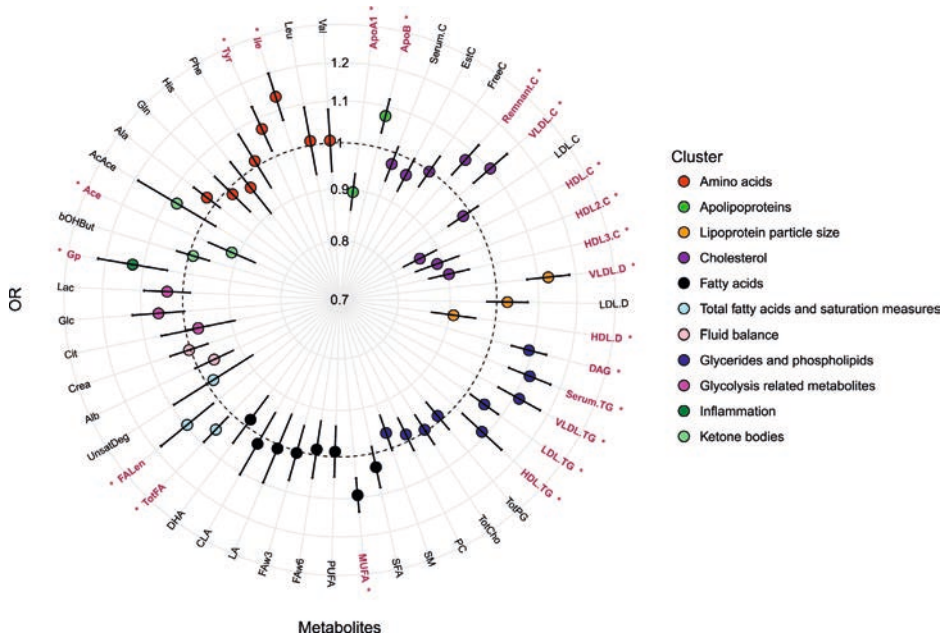

Figure 1. Polar plot illustrating pooled odds ratio and 95% confidence intervals for the association of the 51 lipids, fatty acids and various low-molecular-weight metabolites with depression

*Significant at false discovery rate q < 0.05. Dotted circle indicates an OR of 1. Density: high-density lipoprotein (HDL) subfraction 2 (HDL_2), 1.063–1.125 g/mL; HDL_3, 1.125–1.210 g/mL. AcAce, acetoacetate; Ace, acetate; Ala, alanine; Alb, albumin; ApoA1, apolipoprotein A-I; ApoB, apolipoprotein B; bOHBut, 3-hydroxybutyrate; C, cholesterol; Cit, citrate; CLA, conjugated linoleic acids; Crea, creatinine; D, mean diameter; DAG, diglycerides; DHA, docosahexaenoic acid; Est, esterified; FA, fatty acids; FALen, estimated fatty acids chain length; FAw3, ω-3 fatty acids; FAw6, ω-6 fatty acids; Glc, glucose; Gln, glutamine; Gp, glycoprotein acetyls, mainly α1-acid glycoprotein; His, histidine; IDL, intermediate-density lipoprotein; Ile, isoleucine; LA, linoleic acid (18:2); Lac, lactate; Leu, leucine; LDL, low-density lipoprotein; MUFA, monounsaturated fatty acids (16:1, 18:1); PC, phosphatidylcholine and other cholines; Phe, phenylalanine; PUFA, polyunsaturated fatty acids; Remnant, non-HDL, non-LDL cholesterol; SFA, saturated fatty acids; SM, sphingomyelins; TG, triglycerides; TotCho, total cholines; TotFA, total fatty acids; TotPG, total phosphoglycerides; Tyr, tyrosine; UnsatDeg, estimated degree of unsaturation; Val, valine; VLDL, very-low-density lipoprotein.

Heterogeneity was small (I^2<25% for 15 out of 21 metabolites) and statistically non-significant in almost all (19 out of 21) analyses. As shown in the related forest plots (Figure S1 in Supplement 1) association estimates were quite consistent across the different datasets, including those enriched for cardiometabolic risk. To confirm this, we reran the analyses after removing two datasets (CODAM subgroup with type 2 diabetes mellitus and TMS subgroup with type 2 diabetes mellitus) containing approximately 900 participants with established diabetes and elevated cardiovascular risk factors. Association estimates were highly concordant with those of the original analyses (r=0.99); all the 21 metabolites detected in the

original analyses were associated at nominal level with depression (17 at FDR q<0.05; Table S3 in Supplement 1).

Additional adjustment for BMI partially reduce the strength of the association of these 21 metabolites with depression (regression slope of the 21 beta's before versus after BMI-adjustment=0.65, whereas a beta value of 1 would indicate similar average association sizes; correlation r=0.98): of the 21 metabolites associated with depression, 16 remained significantly related to depression at p<0.05 and 13 at FDR q<0.05 (Table 2). Table S2 in Supplement 2 shows the pooled ORs and heterogeneity findings for all metabolites.

Associations of 98 detailed lipid composition and particle concentration measures of lipoprotein subclasses with depression

Figure 2 shows the ORs of the meta-analyses for the 98 lipid measures of the 14 lipoprotein subclasses, ordered from large to small particle size. Generally, there appeared to be a shift in association with depression by lipoprotein classes: VLDL lipoprotein levels were positively related to depression, intermediate-density lipoprotein (IDL) and LDL lipid levels were not consistently associated, whereas HDL lipoprotein measures were inversely related to depression. Furthermore, depression was related to higher TG levels.

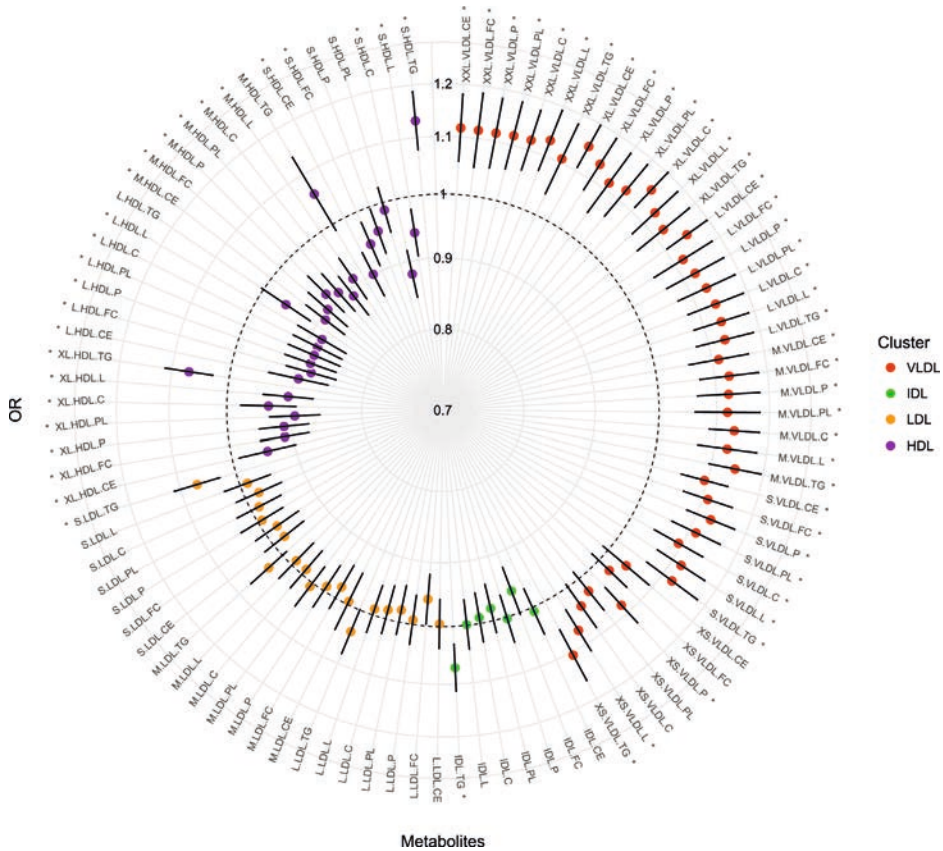

Figure 2. Pooled odds ratios (OR) and 95% confidence intervals for the association of the 98 lipid measures of lipoprotein subclasses with depression.
*Significant at false discovery rate q < 0.05. Dotted circle indicates an OR of 1. Particle sizes: extremely large (XXL) very-low-density lipoprotein (VLDL), >75 nm; very large (XL) VLDL, 64 nm; large (L) VLDL, 53.6 nm; medium (M) VLDL, 44.5 nm; small (S) VLDL, 36.8 nm; very small (XS) VLDL, 31.3 nm; intermediate-density lipoprotein (IDL), 28.6 nm; L low-density lipoprotein (LDL), 25.5 nm; M LDL, 23.0 nm; S LDL, 18.7 nm; XL high-density lipoprotein (HDL), 14.3 nm; L HDL, 12.1 nm; M HDL, 10.9 nm; S HDL, 8.7 nm. C, total cholesterol; CE, cholesterol ester; FC, free cholesterol; L, total lipids; P, particle concentration; PL, phospholipids; TC, triglycerides

Associations of 81 metabolite ratios with depression

Figure S2 in Supplement 1 shows the ORs of the meta-analyses for the 81 metabolite ratios, of which 27 were significant at FDR q<0.05. In general, TG to total lipid ratios were significantly related to an increased odds of depression. Some of the VLDL, IDL, LDL, and HDL lipid measures as percentage of total lipids were positively related to depression, whereas others were inversely related. In general, associations of the metabolite ratios with depression were less pronounced compared to those with absolute metabolite values.

Sensitivity analyses

Current depression

The original 5,283 depression cases included subjects with any lifetime history of depression. In 62% of the cases (3,265 subjects) depression was present between one month before and one month after blood draw. We repeated analyses with only these 3,265 current cases with depression (vs. 10,145 controls). Of the 51 lipids, fatty acids and low-molecular weight metabolites, 22 were associated with current depression at FDR q<0.05 (Figure S3 in Supplement 1). Notably, the strength of the associations with the 51 metabolites tended to be greater for current depression than for the original definition (regression slope of beta's for current versus broadly defined depression=1.22, r=0.95) (Table S2 in Supplement 2). Table S2 in Supplements 2 and Figure S4 and S5 in Supplement 1 show associations of the 98 lipid measures of lipoprotein subclasses, and the 81 metabolite ratios with current depression, which were largely in line with those of original analyses.

Antidepressant medication

To study whether associations were independent of concurrent antidepressant medication use, we removed 1,597 subjects across cohorts who reported use of antidepressants. The majority were depression cases (N=1,305), which was expected given that depression is the main indication for receiving antidepressant treatment. Additionally, one study (LLD) was removed because of model convergence issues. In the remaining 3,966 cases and 8,887 controls - representing a 21% decrease in effective sample size compared with the original analyses, associations with the 51 lipids, fatty acids and low-molecular-weight metabolites were generally in the same direction, but the strength of the associations was attenuated (regression slope of betas before and after exclusion of antidepressant users=0.60, r=0.88) (Figure S6 in Supplement 1). Among the 21 significantly associated metabolites in the overall sample, 8 were still associated at p<0.05, of which 2 (HDL_3- cholesterol, and acetate) at FDR q<0.05 in the antidepressant-free subsample.

Subgroups

Exploration of consistency of associations across subgroups showed that there were no significant differences (Wald-test, FDR q>0.05) in the strength of the association between metabolites and depression across subgroups with depression diagnoses vs. self-reported depression (r=0.75, Figure S7 in Supplement 1), across men vs. women (r=0.64, Figure S8 in Supplement 1), across age <50 years vs. >=50 years (r=0.84, Figure S9 in Supplement 1), and across BMI groups (normal vs. overweight r=0.68, normal vs. obese r=0.55, overweight vs. obese r=0.71, Figures S10-12 in Supplement 1).

DISCUSSION

This meta-analysis of metabolomics and depression, is to our knowledge the largest of its kind. We analyzed data of more than 15,000 subjects from nine Dutch clinical and population-based studies in the Netherlands to identify metabolites associated with depression. Our findings showed that depression is associated with a metabolic signature towards less HDL and more VLDL and triglycerides particles. More specifically, 21 plasma lipids, fatty acids and low-molecular-weight metabolites were significantly related to depression: higher levels of apolipoprotein B, VLDL cholesterol, triglycerides, diglycerides, total and mono-unsaturated fatty acids, fatty acid chain length, glycoprotein acetyls, tyrosine, and isoleucine, and lower levels of HDL cholesterol, acetate, and apolipoprotein A1. Associations were generally consistent across sex, age and body mass index strata, and across cohorts using depression diagnoses vs. depressive symptoms. These metabolic alterations in depression could potentially explain part of the increased risk of cardiometabolic disease in individuals with depression.

Our findings that depression is related to higher VLDL, higher TG and lower VLDL are in line with previous research [3, 11, 45]. In the present study, we predominantly found differences in absolute lipid measures of the VLDL subfractions, whereas findings with lipid measures to lipid ratios in VLDL were less consistently associated with depression. This suggests that the total amount of lipids, rather than the type of lipids, is the main contributor to associations of depression with VLDL. For other metabolites, previous studies indicated more mixed findings. We did not find associations for LDL cholesterol measures, which contrasts with a previous meta-analysis that showed associations between depression and increased LDL cholesterol [11]. For measures of fatty acids, we observed that higher mono unsaturated fatty acids, total fatty acids and estimated fatty acids chain length were associated with an increased odds of depression. Most evidence for links with fatty acids in depression stems from research on omega-3 fatty acids [12], for which we did not observe a consistent, significant association with depression in the present study. The finding of proinflammatory glycoprotein acetyls being positively associated with depression is in line with the large body of evidence linking inflammation to depression [46]. The short chain fatty acid and ketone body acetate was lower in depression. It was hypothesized that a Western-style diet alters gut microbiome composition, resulting in lower acetate levels, which could subsequently induce depression [4]. Furthermore, a smaller study found lower isoleucine levels in depression [47], which contrasts our findings. Finally, a review concluded that there was no association between tyrosine and depression [48], whereas we observed higher tyrosine in depression. Discrepancies could be explained by differences in study characteristics or

variation in analytic approaches, such as selection of potentially confounding factors.

We additionally evaluated the impact of the time frame of depression assessment on the results. In secondary analyses including cases with current depression only, associations tended to become enhanced, suggesting that metabolomics alterations represent state markers reflecting current depression. Nevertheless, a similar profile of associations was found when analyzing depression cases defined in a broader timeframe. The metabolic signature identified may therefore also represent a persisting biological scar after remission of depression, or a pre-existing underlying vulnerability factor for development of depression.

The impact of antidepressant medication use on the results was explored in secondary analyses, although this observational study precludes definitive conclusions, as depression severity most likely represents the clinical indication for antidepressant treatment (confounded by indication) [49]. We reanalyzed data after excluding antidepressant users, and found that the strength of associations was attenuated. Furthermore, the reduction in effective sample size substantially impacted the power to find significant associations. Nevertheless, directions of associations were highly consistent with those obtained in the full sample. Furthermore, the literature shows that potential detrimental effects of antidepressants on dyslipidemia is evident mainly for tricyclic antidepressants (TCA) [50, 51]. Data from the NESDA cohort [51], including patients from mental health care institutions and with the highest prevalence of antidepressant users (27%, Table 1), showed that TCA antidepressant were prescribed only in 3% of the participants. As the overall prevalence of antidepressant use in other cohorts included in the present meta-analysis was lower than approximately 9%, it could be assumed that the number of TCA users may be limited. This observation, combined with the results of our sensitivity analyses, suggests that antidepressant use is unlikely to be the major driver of the associations between metabolites and depression.

Secondary analyses also indicated that results were generally attenuated when BMI was taken into account, suggesting that part of the differential metabolite levels in depression could be explained by BMI. However, interrelationships between BMI, metabolite, depression and antidepressants are particularly complex. A significant genetic correlation has been found between depression and BMI [52], indicating that they may emerge from partially shared etiological mechanisms; at the same time BMI has been shown to influence metabolite concentrations [42]. The ability to disentangle different independent effects of this complex network in observational data is limited. Nevertheless, the majority of metabolites were

associated with depression after taking into account BMI, indicating that this factor explains only a limited portion of the depression-metabolites link.

The present findings may be explained by three, non-mutually exclusive, scenarios. First, alterations of lipids may be a consequence of depression. Depressed persons are more likely to engage in unhealthy behaviors such as sedentariness, excessive alcohol use and poor nutrition (e.g., saturated fats), which may lead to dyslipidemia [16]. Second, lipid dysregulations may be part of the pathophysiological pathways implicated in depression, such as chronic HPA-axis and inflammatory activity, resulting in lipolysis, release of fatty acids, synthesis of VLDL, hypertriglyceridemia and reduction in HDL cholesterol. Third, metabolomic alterations in depression may represent epiphenomena stemming from the same root, such as a common genetic factor. A recent genome-wide association study (GWAS) of major depression involving >450,000 participants, reported a significant genetic correlation ($rg=0.14$, $p=7.8x10^{-7}$) with high TG levels, but not with LDL or HDL [53]. Furthermore, no genetic correlations emerged with metabolites of the same panel that we found to be associated with depression, although the relatively smaller sample size (~25,000) of the metabolomics GWAS may substantially limit the ability to detect correlation; the only exception was a nominally significant correlation with glycoprotein acetyls ($rg=0.15$, $p=0.03$), with the same direction of the phenotypic association we identified. Further experimental studies and genetically informed designs such as Mendelian randomization may disentangle whether depression and lipid dysregulations emerge from shared etiology, and whether depression causally determines lipid alterations or vice versa.

The present study has some limitations. Owing to limited availability or differences in assessment across datasets we cannot rule out confounding by other health-related or lifestyle factors, such as chronic cardiometabolic conditions, alcohol use or specific food intake before sample collection. Nevertheless, the associations between depression and metabolites were consistent across datasets, including those enriched for traits such as diabetes, cardiovascular risk factors and migraine. Furthermore, alcohol use may represent a mediating mechanism rather than a confounder in the metabolites-depression association, as recent evidence [54] showed that alcohol dependence is to quite some extent caused by depression. Analyses were adjusted for fasting status (>94% of subjects were fasting, Table 1), but both fasting and non-fasting samples can be reliably analyzed by the metabolomics platform used [26, 36]. We could not examine whether the associations with metabolites detected vary as a function of specific depression clinical characteristics. Strengths of the study (large samples, metabolites data generated for all studies with the same platform) have enabled the identification of the most reliable metabolic signals associated with depression. These are worth further examination in relation to clinically relevant phenotypes (e.g., age of onset,

recurrence, duration, symptom profiles) in future studies based on psychiatrically well-characterized samples.

This large-scale meta-analysis including more than 15,000 participants identified a metabolomics signature associated with depression. This biological signature is partially shared with other conditions such as diabetes, obesity and cardiovascular diseases [3, 5-7] that commonly co-occur with depression, heavily burdening public health resources. Alterations in the lipid spectrum identified in the present study may represent a substrate linking depression to cardiometabolic diseases and, therefore, a potential target for prevention and treatment of depression and its detrimental somatic sequelae.

REFERENCES

1. Otte, C., et al., *Major depressive disorder.* Nat Rev Dis Primers, 2016. **2**: p. 16065.
2. Murray, C.J., et al., *Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010.* Lancet, 2012. **380**(9859): p. 2197-223.
3. Pan, A., et al., *Bidirectional association between depression and metabolic syndrome: a systematic review and meta-analysis of epidemiological studies.* Diabetes Care, 2012. **35**(5): p. 1171-80.
4. Milaneschi, Y., et al., *Depression and obesity: evidence of shared biological mechanisms.* Mol Psychiatry, 2019. **24**(1): p. 18-33.
5. Mezuk, B., et al., *Depression and type 2 diabetes over the lifespan: a meta-analysis.* Diabetes Care, 2008. **31**(12): p. 2383-90.
6. Pan, A., et al., *Depression and risk of stroke morbidity and mortality: a meta-analysis and systematic review.* Jama, 2011. **306**(11): p. 1241-9.
7. Van der Kooy, K., et al., *Depression and the risk for cardiovascular diseases: systematic review and meta analysis.* Int J Geriatr Psychiatry, 2007. **22**(7): p. 613-26.
8. Cuijpers, P., et al., *Comprehensive meta-analysis of excess mortality in depression in the general community versus patients with specific illnesses.* Am J Psychiatry, 2014. **171**(4): p. 453-62.
9. MacDonald, K., et al., *Biomarkers for major depressive and bipolar disorders using metabolomics: A systematic review.* Am J Med Genet B Neuropsychiatr Genet, 2019. **180**(2): p. 122-137.
10. Shin, J.Y., J. Suls, and R. Martin, *Are cholesterol and depression inversely related? A meta-analysis of the association between two cardiac risk factors.* Ann Behav Med, 2008. **36**(1): p. 33-43.
11. Persons, J.E. and J.G. Fiedorowicz, *Depression and serum low-density lipoprotein: A systematic review and meta-analysis.* J Affect Disord, 2016. **206**: p. 55-67.
12. Lin, P.Y., S.Y. Huang, and K.P. Su, *A meta-analytic review of polyunsaturated fatty acid compositions in patients with depression.* Biol Psychiatry, 2010. **68**(2): p. 140-7.
13. Fardet, L. and B. Fève, *Systemic glucocorticoid therapy: a review of its metabolic and cardiovascular adverse events.* Drugs, 2014. **74**(15): p. 1731-45.
14. Esteve, E., W. Ricart, and J.M. Fernández-Real, *Dyslipidemia and inflammation: an evolutionary conserved mechanism.* Clin Nutr, 2005. **24**(1): p. 16-31.
15. Grosso, G., et al., *Omega-3 fatty acids and depression: scientific evidence and biological mechanisms.* Oxid Med Cell Longev, 2014. **2014**: p. 313570.
16. Mensink, R.P., et al., *Effects of dietary fatty acids and carbohydrates on the ratio of serum total to HDL cholesterol and on serum lipids and apolipoproteins: a meta-analysis of 60 controlled trials.* Am J Clin Nutr, 2003. **77**(5): p. 1146-55.
17. Quinones, M.P. and R. Kaddurah-Daouk, *Metabolomics tools for identifying biomarkers for neuropsychiatric diseases.* Neurobiol Dis, 2009. **35**(2): p. 165-76.

18. Gowda, G.A., et al., *Metabolomics-based methods for early disease diagnostics.* Expert Rev Mol Diagn, 2008. **8**(5): p. 617-33.

19. Shao, W.H., et al., *Combined Metabolomics and Proteomics Analysis of Major Depression in an Animal Model: Perturbed Energy Metabolism in the Chronic Mild Stressed Rat Cerebellum.* Omics, 2015. **19**(7): p. 383-92.

20. Zheng, H., et al., *Predictive diagnosis of major depression using NMR-based metabolomics and least-squares support vector machine.* Clin Chim Acta, 2017. **464**: p. 223-227.

21. Martins-de-Souza, D., *Proteomics, metabolomics, and protein interactomics in the characterization of the molecular features of major depressive disorder.* Dialogues Clin Neurosci, 2014. **16**(1): p. 63-73.

22. Guest, P.C., F.L. Guest, and D. Martins-de Souza, *Making Sense of Blood-Based Proteomics and Metabolomics in Psychiatric Research.* Int J Neuropsychopharmacol, 2016. **19**(6).

23. Gadad, B.S., et al., *Peripheral biomarkers of major depression and antidepressant treatment response: Current knowledge and future outlooks.* J Affect Disord, 2018. **233**: p. 3-14.

24. Würtz, P., et al., *Metabolic signatures of insulin resistance in 7,098 young adults.* Diabetes, 2012. **61**(6): p. 1372-80.

25. Würtz, P., et al., *Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts.* Circulation, 2015. **131**(9): p. 774-85.

26. Fischer, K., et al., *Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17,345 persons.* PLoS Med, 2014. **11**(2): p. e1001606.

27. van Greevenbroek, M.M., et al., *The cross-sectional association between insulin resistance and circulating complement C3 is partly explained by plasma alanine aminotransferase, independent of central obesity and general inflammation (the CODAM study).* Eur J Clin Invest, 2011. **41**(4): p. 372-9.

28. Schram, M.T., et al., *The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities.* Eur J Epidemiol, 2014. **29**(6): p. 439-51.

29. Sayed-Tabatabaei, F.A., et al., *Heritability of the function and structure of the arterial wall: findings of the Erasmus Rucphen Family (ERF) study.* Stroke, 2005. **36**(11): p. 2351-6.

30. van Oosterhout, W.P., et al., *Validation of the web-based LUMINA questionnaire for recruiting large cohorts of migraineurs.* Cephalalgia, 2011. **31**(13): p. 1359-67.

31. Boomsma, D.I., et al., *Netherlands Twin Register: from twins to twin families.* Twin Res Hum Genet, 2006. **9**(6): p. 849-57.

32. Scholtens, S., et al., *Cohort Profile: LifeLines, a three-generation cohort study and biobank.* Int J Epidemiol, 2015. **44**(4): p. 1172-80.

33. Tigchelaar, E.F., et al., *Gut microbiota composition associated with stool consistency.* Gut, 2016. **65**(3): p. 540-2.

34. Zhernakova, A., et al., *Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity.* Science, 2016. **352**(6285): p. 565-9.

35. Soininen, P., et al., *Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics.* Circ Cardiovasc Genet, 2015. **8**(1): p. 192-206.

36. Würtz, P., et al., *Metabolic profiling of alcohol consumption in 9778 young adults.* Int J Epidemiol, 2016. **45**(5): p. 1493-1506.

37. Würtz, P., et al., *Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies.* Am J Epidemiol, 2017. **186**(9): p. 1084-1096.

38. Onderwater, G.L.J., et al., *Large-scale plasma metabolome analysis reveals alterations in HDL metabolism in migraine.* Neurology, 2019. **92**(16): p. e1899-e1911.

39. Higgins, J.P., et al., *Measuring inconsistency in meta-analyses.* Bmj, 2003. **327**(7414): p. 557-60.

40. Viechtbauer, W., *Confidence intervals for the amount of heterogeneity in meta-analysis.* Stat Med, 2007. **26**(1): p. 37-52.

41. Ioannidis, J.P., N.A. Patsopoulos, and E. Evangelou, *Uncertainty in heterogeneity estimates in meta-analyses.* Bmj, 2007. **335**(7626): p. 914-6.

42. Würtz, P., et al., *Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change.* PLoS Med, 2014. **11**(12): p. e1001765.

43. Viechtbauer, W., *Comparing estimates of independent meta-analyses or subgroups*. 2017.

44. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* Journal of the Royal statistical society: series B (Methodological), 1995. **57**(1): p. 289-300.

45. Segoviano-Mendoza, M., et al., *Hypocholesterolemia is an independent risk factor for depression disorder and suicide attempt in Northern Mexican population.* BMC Psychiatry, 2018. **18**(1): p. 7.

46. Kiecolt-Glaser, J.K., H.M. Derry, and C.P. Fagundes, *Inflammation: depression fans the flames and feasts on the heat.* Am J Psychiatry, 2015. **172**(11): p. 1075-91.

47. Baranyi, A., et al., *Branched-Chain Amino Acids as New Biomarkers of Major Depression - A Novel Neurobiology of Mood Disorder.* PLoS One, 2016. **11**(8): p. e0160542.

48. Parker, G. and H. Brotchie, *Mood effects of the amino acids tryptophan and tyrosine: 'Food for Thought' III.* Acta Psychiatr Scand, 2011. **124**(6): p. 417-26.

49. Kyriacou, D.N. and R.J. Lewis, *Confounding by Indication in Clinical Research.* Jama, 2016. **316**(17): p. 1818-1819.

50. McIntyre, R.S., et al., *The association between conventional antidepressants and the metabolic syndrome: a review of the evidence and clinical implications.* CNS Drugs, 2010. **24**(9): p. 741-53.

51. van Reedt Dortland, A.K., et al., *Metabolic syndrome abnormalities are associated with severity of anxiety and depression and with tricyclic antidepressant use.* Acta Psychiatr Scand, 2010. **122**(1): p. 30-9.

52. Milaneschi, Y., et al., *Genetic Association of Major Depression With Atypical Features and Obesity-Related Immunometabolic Dysregulations.* JAMA Psychiatry, 2017. **74**(12): p. 1214-1225.

53. Wray, N.R., et al., *Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression.* Nat Genet, 2018. **50**(5): p. 668-681.

54. Polimanti, R., et al., *Evidence of causal effect of major depression on alcohol dependence: findings from the psychiatric genomics consortium.* Psychol Med, 2019. **49**(7): p. 1218-1226.

SUPPLEMENTARY MATERIAL

Full version of supplementary materials can be found through the following link:
https://ars.els-cdn.com/content/image/1-s2.0-S0006322319316282-mmc1.pdf
https://ars.els-cdn.com/content/image/1-s2.0-S0006322319316282-mmc2.xlsx
https://ars.els-cdn.com/content/image/1-s2.0-S0006322319316282-mmc3.xlsx

Information about BBMRI-NL consortium can be found through the following link:
https://www.bbmri.nl/bbmri-metabolomics-consortium

Classification of depressed cases and controls

Controls were those with a negative diagnostic interview for lifetime depression, or had a score on the depression questionnaires below established cut-off scores (i.e., CES-D<16, HADS-D<8 and/or IDS-SR30<14). If multiple self-reports of depressive symptoms before blood sampling were available, controls needed to score below the established cut-offs during all these assessments. When diagnostic data on other psychiatric disorders were available (e.g., anxiety disorders), persons with other psychiatric disorders were excluded from the controls.

Metabolomics assessment

A total of 230 metabolites or metabolite ratios were reliably quantified from Ethylenediaminetetraacetic acid (EDTA) plasma samples using targeted high-throughput proton Nuclear Magnetic Resonance (^{1}H-NMR) metabolomics (Nightingale Health Ltd, Helsinki, Finland) [19]. This platform provides simultaneous quantification of routine lipids, lipoprotein subclass profiling with lipid concentrations within 14 subclasses, fatty acid composition, and various low-molecular-weight metabolites including amino acids, ketone bodies and gluconeogenesis-related metabolites in molar concentration units. This metabolomics platform has been extensively used and described in numerous studies (see https://nightingalehealth.com/publications for an overview), including large-scaled epidemiological studies in the field of type 2 diabetes [20], cardiovascular disease [21], mortality [22], and lifestyle factors such as alcohol intake [23]. Details of the experimentation and applications of the ^{1}H-NMR metabolomics platform have been extensively described previously [19, 24, 25].

The entire process from sample handling to data processing is highly standardized and fully automated. Samples were prepared irrespective of depression status, because depression cases and controls entered each study at random order (i.e. unrelated to depression status), and the laboratory analyzing the samples was unaware of depression cases vs. control status when preparing the samples. Automated liquid handlers mixed 260 µL buffer (75 mM Na_2HPO_4 in 80%/20% H_2O/D_2O, pH 7.4; 4.64 mM sodium 3-(trimethylsilyl)propionate-2,2,3,3-d_4, and

6.15 mM sodium azide) with the plasma in 1:1 ratio and moved the prepared samples to 96-format racks of NMR tubes, which were subsequently moved to the robotic sample changer, cooled to refrigerator temperature. Each rack contained 2 quality control samples: 1 serum mimic and a mixture of 2 low-molecular-weight metabolites. For the native plasma samples, the lipoprotein (80k data points after 4 dummy scans using 8 transients, 90° pulse) and low-molecular-weight metabolites (64k data points, using 24 (or 16) transients acquired after 4 steady state scans, T2-relaxation-filtered pulse sequence) data were automatically collected at 310.1K either with the 500 MHz or the 600 MHz Bruker AVANCE IIIHD NMR spectrometer, with a relaxation delay of 3.0 seconds [19, 25].

The NMR spectra are converted to absolute concentrations via Bayesian modeling performed via advanced proprietary software and integrates quality control checks. Several of the metabolic biomarkers have already been 'validated' with other techniques (i.e. routine clinical chemistry assays, gas chromatography, an enzymatic method, and/or mass spectrometry) [21, 24, 26-28]. Furthermore, genetic studies [29-31] performed on the same metabolomics platform showed that the labels applied to the metabolites are coherent and linked with biologically relevant and plausible genes.

The 14 lipoprotein subclass sizes were defined as follows: extremely large VLDL with particle diameters from 75 nm upwards and a possible contribution of chylomicrons, five VLDL subclasses, IDL, three LDL subclasses and four HDL subclasses. The following components of the lipoprotein subclasses were quantified: phospholipids (PL), TG, cholesterol (C), free cholesterol (FC), and cholesteryl esters (CE). The mean size for VLDL, LDL and HDL particles was calculated by weighting the corresponding subclass diameters with their particle concentrations.

NMR spectroscopy provides highly consistent biomarker quantification. This is due to the inherently reproducible nature of the technology; the samples never come into contact with the radiofrequency detector in the NMR spectrometer. Biomarker quantification directly from plasma, without any sample extraction procedures, further contributes to the high reproducibility [24]. Representative coefficients of variations (CVs) for the metabolic biomarkers are published as Supplementary Data 3 in Kettunen et al. [30] with the CVs determined for 9,600 samples. Values ranged between 0.3 and 19.5 (mean 4.5%), and most values are comparable to routinely used assays in clinical chemistry.

Covariates

To be largely in line with previous metabolomics meta-analytic studies, [23], we adjusted analyses for the following potentially confounding variables: age (in

years), sex, fasting status (yes/no), use of lipid modifying medication (yes/no), and current smoking (yes/no). The lipid modifying drugs were defined according to the related Anatomical Therapeutic Chemical Classification System (ATC) code C10 (Lipid modifying agents) in order to capture all the medications falling under this category, including the use of single agents (C10A - Lipid modifying agents, plain: C10AA HMG CoA reductase inhibitors; C10AB Fibrates; C10AC Bile acid sequestrants; C10AD Nicotinic acid and derivatives; C10AX Other lipid modifying agents) and all their potential combinations (C10B - Lipid modifying agents, combination: C10BA HMG CoA reductase inhibitors in combination with other lipid modifying agents; C10BX HMG CoA reductase inhibitors, other combinations). The antidepressant medications selected for the sensitivity analyses included all classes listed under the ATC code N06A (N06AA Non-selective monoamine reuptake inhibitors, N06AB Selective serotonin reuptake inhibitors, N06AF Monoamine oxidase inhibitors, non-selective, N06AG Monoamine oxidase A inhibitors, N06AX Other antidepressants). Given the bidirectional relationship between depression and obesity and their shared biological processes (including genes, endocrine and immuno-inflammatory mechanisms) [32], the role of obesity was explored in greater detail in sensitivity analysis (see Statistical analyses). Body mass index (BMI) was calculated as measured weight (kg)/length $(m)^2$, and divided into normal weight (BMI=18.50-24.99), overweight (BMI=25.00-29.99) and obesity (BMI≥30).

Assessment of potential bias due to metabolites data transformation

According to the standardized protocol of data processing applied in the present study a constant of 1 was added to the metabolite values before log-transformation. This common practice, adopted also in several other studies also from the same BBMRI-NL Metabolomics Consortium [33], aims to achieve normalization of the distribution also for metabolites with initial values equaling zero. Nevertheless, it is important to acknowledge that this transformation may have had introduced some bias due to the high variability in the normal range of different metabolite. In the present analyses we aimed to estimate the potential degree of bias introduced by comparing the results of the metabolites-depression associations obtained applying three different transformation before log-transformation: A) adding a constant of 1; B) adding the value of the 10^{th} percentile of the distribution (excluding 0 values) of each metabolite, a value therefore within the normal range of the original metabolite; C) excluding all 0 values, a more conservative approach.

Analyses were performed in the NESDA sample (N=2,509), the most representative dataset for the trait under study, which involves subjects well phenotyped in psychiatric terms including healthy controls and depressed patients from various settings and developmental stages of psychopathology. Furthermore, analyses

focused on the 51 metabolites classified in the cluster of "lipids, fatty acids and various low-molecular-weight metabolites".

Ridge plots in Figure S13 shows the distribution (per SD increase) of the (log) values of the metabolites after the three different transformation. The three sets of values were used in logistic regression analyses estimating the association between metabolites and lifetime depression, adjusting for sex, age, smoking, lipid modifying drugs and fasting status. Results were highly similar across the three transformations. In Figure S14 the estimates obtained used the original transformation A were plotted against estimates obtained with transformation B (panel 1), and against those obtained with transformation C (panel 2). In both instances the correlation between association effect sizes equaled 1 as the estimates were substantially identical across transformation (coefficient from regressing estimates of transformation A on those from transformation B = 1.02, se=0.01; coefficient from regressing estimates of transformation A on those from transformation C = 1.00, se=0.02). Overall, these results suggests that the degree of bias potentially introduced by the transformation applied in original analyses is minimal and negligible.

REFERENCES

1. van Greevenbroek, M.M., et al., *The cross-sectional association between insulin resistance and circulating complement C3 is partly explained by plasma alanine aminotransferase, independent of central obesity and general inflammation (the CODAM study).* Eur J Clin Invest, 2011. **41**(4): p. 372-9.
2. van Dam, R.M., et al., *Parental history of diabetes modifies the association between abdominal adiposity and hyperglycemia.* Diabetes Care, 2001. **24**(8): p. 1454-9.
3. Radloff, L.S., *The CES-D Scale:A Self-Report Depression Scale for Research in the General Population.* Applied Psychological Measurement, 1977. **1**(3): p. 385-401.
4. Lewinsohn, P.M., et al., *Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults.* Psychol Aging, 1997. **12**(2): p. 277-87.
5. Schram, M.T., et al., *The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities.* Eur J Epidemiol, 2014. **29**(6): p. 439-51.
6. Sheehan, D.V., et al., *The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10.* J Clin Psychiatry, 1998. **59 Suppl 20**: p. 22-33;quiz 34-57.
7. Sayed-Tabatabaei, F.A., et al., *Heritability of the function and structure of the arterial wall: findings of the Erasmus Rucphen Family (ERF) study.* Stroke, 2005. **36**(11): p. 2351-6.
8. Zigmond, A.S. and R.P. Snaith, *The hospital anxiety and depression scale.* Acta Psychiatr Scand, 1983. **67**(6): p. 361-70.
9. van Oosterhout, W.P., et al., *Validation of the web-based LUMINA questionnaire for recruiting large cohorts of migraineurs.* Cephalalgia, 2011. **31**(13): p. 1359-67.
10. Rush, A.J., et al., *The Inventory for Depressive Symptomatology (IDS): preliminary findings.* Psychiatry Res, 1986. **18**(1): p. 65-87.
11. Rush, A.J., et al., *The Inventory of Depressive Symptomatology (IDS): psychometric properties.* Psychol Med, 1996. **26**(3): p. 477-86.
12. Penninx, B.W., et al., *The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods.* Int J Methods Psychiatr Res, 2008. **17**(3): p. 121-40.
13. Wittchen, H.U., *Reliability and validity studies of the WHO--Composite International Diagnostic Interview (CIDI): a critical review.* J Psychiatr Res, 1994. **28**(1): p. 57-84.
14. Boomsma, D.I., et al., *Netherlands Twin Register: from twins to twin families.* Twin Res Hum Genet, 2006. **9**(6): p. 849-57.
15. Hofman, A., et al., *The Rotterdam Study: 2016 objectives and design update.* Eur J Epidemiol, 2015. **30**(8): p. 661-708.
16. Scholtens, S., et al., *Cohort Profile: LifeLines, a three-generation cohort study and biobank.* Int J Epidemiol, 2015. **44**(4): p. 1172-80.
17. Tigchelaar, E.F., et al., *Gut microbiota composition associated with stool consistency.* Gut, 2016. **65**(3): p. 540-2.

18. Zhernakova, A., et al., *Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity.* Science, 2016. **352**(6285): p. 565-9.

19. Soininen, P., et al., *Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics.* Circ Cardiovasc Genet, 2015. **8**(1): p. 192-206.

20. Würtz, P., et al., *Metabolic signatures of insulin resistance in 7,098 young adults.* Diabetes, 2012. **61**(6): p. 1372-80.

21. Würtz, P., et al., *Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts.* Circulation, 2015. **131**(9): p. 774-85.

22. Fischer, K., et al., *Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17,345 persons.* PLoS Med, 2014. **11**(2): p. e1001606.

23. Würtz, P., et al., *Metabolic profiling of alcohol consumption in 9778 young adults.* Int J Epidemiol, 2016. **45**(5): p. 1493-1506.

24. Würtz, P., et al., *Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies.* Am J Epidemiol, 2017. **186**(9): p. 1084-1096.

25. Soininen, P., et al., *High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism.* Analyst, 2009. **134**(9): p. 1781-5.

26. Tynkkynen, J., et al., *Association of branched-chain amino acids and other circulating metabolites with risk of incident dementia and Alzheimer's disease: A prospective study in eight cohorts.* Alzheimers Dement, 2018. **14**(6): p. 723-733.

27. Holmes, M.V., et al., *Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke.* J Am Coll Cardiol, 2018. **71**(6): p. 620-632.

28. Ritchie, S.C., et al., *The Biomarker GlycA Is Associated with Chronic Inflammation and Predicts Long-Term Risk of Severe Infection.* Cell Syst, 2015. **1**(4): p. 293-301.

29. Kettunen, J., et al., *Genome-wide association study identifies multiple loci influencing human serum metabolite levels.* Nat Genet, 2012. **44**(3): p. 269-76.

30. Kettunen, J., et al., *Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA.* Nat Commun, 2016. **7**: p. 11122.

31. Tukiainen, T., et al., *Detailed metabolic and genetic characterization reveals new associations for 30 known lipid loci.* Hum Mol Genet, 2012. **21**(6): p. 1444-55.

32. Milaneschi, Y., et al., *Depression and obesity: evidence of shared biological mechanisms.* Mol Psychiatry, 2019. **24**(1): p. 18-33.

33. Onderwater, G.L.J., et al., *Large-scale plasma metabolome analysis reveals alterations in HDL metabolism in migraine.* Neurology, 2019. **92**(16): p. e1899-e1911.