American Journal of Epidemiology

Vol. 188, No. 11
DOI: 10.1093/aje/kwz171
Advance Access publication: July 31, 2019

Original Contribution

Metabolome-Wide Association Study of the Relationship Between Habitual Physical Activity and Plasma Metabolite Levels

Ming Ding*, Oana A. Zeleznik, Marta Guasch-Ferre, Jie Hu, Jessica Lasky-Su, I-Min Lee, Rebecca D. Jackson, Aladdin H. Shadyab, Michael J. LaMonte, Clary Clish, A. Heather Eliassen, Frank Sacks, Walter C. Willett, Frank B. Hu, Kathryn M. Rexrode, and Peter Kraft*

* Correspondence to Dr. Peter Kraft, Department of Epidemiology, Harvard T.H. Chan School of Public Health, 667 Huntington Avenue, Boston, MA 02120 (e-mail: pkraft@hsph.harvard.edu); or Dr. Ming Ding, Department of Nutrition, Harvard T.H. Chan School of Public Health, 667 Huntington Avenue, Boston, MA 02120 (e-mail: mid829@mail.harvard.edu).

Initially submitted January 23, 2019; accepted for publication July 18, 2019.

We identified plasma metabolites associated with habitual physical activity among 5,197 US participants from the Nurses' Health Study (NHS), Nurses' Health Study II (NHS II), and the Health Professionals Follow-up Study (HPFS). Physical activity was assessed every 2–4 years via self-report questionnaires. Blood was collected in the NHS in 1989–1990, in NHS II during 1996–1999, and in the HPFS during 1993–1995. Metabolic profiling was conducted by liquid chromatography–mass spectrometry. Our study included 337 known metabolites, with 256 of them classified as lipids. We corrected for multiple testing by controlling the tail probability of the proportion of false positives (TPPFP) and accounted for correlated tests using bootstrapping. Physical activity was significantly associated with 20 metabolites after correction for multiple testing (TPPFP < 0.05), and positive associations were found for most of the metabolites, including 2 amino acids (citrulline and glycine), 4 cholesteryl esters (C18:2, C18:1, C16:0, C18:3), 8 phosphocholines (PCs) (C36:4 PC-A, C34:3 PC plasmalogen, C36:3 PC plasmalogen, C34:2 PC plasmalogen, C36:2 PC) and lysophosphatidylcholines (C18:2, C20:5, C18:1), and 3 phosphatidylethanolamines (PEs) (C38:3 PE plasmalogen) and lysophosphatidylethanolamines (C18:2, C18:1). We independently replicated the 20 metabolites among 2,305 women in the Women's Health Initiative using 1993 data, and half of the metabolites were replicated. Our study may help identify biomarkers of physical activity and provide insight into biological mechanisms underlying the beneficial effect of being physically active on cardiometabolic health.

cohort studies; metabolomics; physical activity

Abbreviations: BCAA, branched-chain amino acid; BMI, body mass index; CE, cholesteryl ester; FDP, false discovery proportion; HDL, high-density lipoprotein; HPFS, Health Professionals Follow-up Study; LDL, low-density lipoprotein; LPC, lysophosphatidylcholine; LPE, lysophosphatidylethanolamine; MET, metabolic equivalent of task; NHS, Nurses' Health Study; NHS II, Nurses' Health Study II; PC, phosphocholine; PE, phosphatidylethanolamine; TPPFP, tail probability of the proportion of false positives; WHI, Women's Health Initiative.

Investigators in the field of metabolomics measure molecules that have been ingested or may be intermediates and products of metabolism (1). Metabolomics is a powerful tool for investigating disease mechanisms and discovering novel biomarkers. Because physical activity is an important component of a healthy lifestyle, being physically inactive is a leading risk factor for cardiometabolic diseases, including cardiovascular disease, diabetes mellitus, hypertension, and the metabolic syndrome (2). Identification of metabolites associated with levels of habitual physical activity has the potential to enhance our understanding of the mechanisms through which participation in regular physical activity benefits cardiometabolic health, including glycemic control, blood pressure, and subclasses of blood lipoproteins (3, 4).

Several metabolome-wide association studies have been conducted to examine associations of habitual physical activity with metabolites, and findings have varied according to coverage of metabolite classes and study sample sizes. In one prospective

study carried out among 2,106 persons, Kujala et al. (5) examined associations of physical activity with plasma subclasses of lipoproteins and several types of glucose, fatty acids, and amino acids between persistently active and inactive individuals, finding that being physically active was associated with lower isoleucine, α1-acid glycoprotein, and glucose levels. In another cross-sectional study involving 1,193 Japanese men and women, Fukai et al. (6) focused on polar metabolites, including amino acids and carbohydrates, and found that higher levels of physical activity were associated with lower concentrations of branched-chain amino acids (BCAAs) (isoleucine, leucine, and valine). In a prospective cohort study, Xiao et al. (7) measured a broader range of metabolites, including amino acids, carbohydrates, and lipids, among 277 Chinese men and women, finding that a higher amount of physical activity was associated with lower levels of BCAAs and several carbohydrate metabolites. However, none of these studies specifically focused on lipid metabolites, which are a major component of plasma metabolites.

In the present study, we examined associations of habitual physical activity with metabolites among 5,197 participants from 10 case-control studies nested within the Nurses' Health Study (NHS), Nurses' Health Study II (NHS II), and the Health Professionals Follow-up Study (HPFS), who were sampled from registered nurses and health professionals in the US population. Our candidate metabolites included 337 known metabolites, most of which were classified as lipids. Our results were independently replicated among 2,305 participants from the Women's Health Initiative (WHI).

METHODS

Study population

The NHS began in 1976, when 121,701 female registered nurses aged 30–55 years residing in 11 states were recruited to complete a baseline questionnaire. NHS II was established in 1989 and consisted of 116,429 younger female registered nurses (ages 25–42 years at baseline). The HPFS was initiated in 1986 and was composed of 51,529 male health professionals (dentists, pharmacists, veterinarians, optometrists, osteopathic physicians, and podiatrists) aged 40–75 years at baseline. The study protocols were approved by the institutional review boards of Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health (Boston, Massachusetts). Between 1989 and 1990, a total of 32,826 NHS participants provided a blood sample. With a similar protocol, blood samples were collected from 18,225 HPFS participants between 1993 and 1995 and from 29,611 NHS II participants in 1996–1999. Participants were sent a comprehensive blood collection kit, and over 90% of samples were returned within 24 hours of blood draw, after which they were centrifuged and aliquoted into cryotubes as plasma, buffy coat, and red blood cells and stored in liquid nitrogen freezers at −130°C or less (8, 9).

For the current study, we included participants who provided a blood sample and had been previously selected for 10 nested metabolomic case-control analyses of pancreatic cancer (NHS and HPFS), type 2 diabetes (NHS and NHS II), chronic stress (NHS II), breast cancer (NHS II), rheumatoid arthritis (NHS and NHS II), ovarian cancer (NHS and NHS II), Parkinson disease (NHS and HPFS), amyotrophic lateral sclerosis (NHS and HPFS), and prostate cancer (HPFS) or to participate in Mind-Body Study 1 (NHS II) (10–12). A total of 5,197 persons were included in the current analysis (2,055 from the NHS, 2,120 from NHS II, and 1,022 from the HPFS).

Launched in 1993, the WHI enrolled 161,808 women aged 50–79 years at 40 US clinical centers into one or more of 3 randomized clinical trials testing the health effects of hormone therapy, dietary modification, and/or calcium and vitamin D supplementation or into a prospective observational study. All participants provided written informed consent. Fasting blood samples were collected from all participants without cardiovascular disease at baseline. For the WHI metabolomics study, a nested case-control design was used. Participants who developed coronary heart disease after the baseline examination were selected as cases from the WHI Observational Study and the WHI Hormone Therapy Trials, and an equal number of controls were selected and matched on 5-year age, race, hysterectomy status, and 2-year enrollment window.

Physical activity assessment

Physical activity was assessed by means of self-report questionnaires every 2–4 years, beginning in 1986 in the NHS and HPFS and beginning in 1989 in NHS II. In this analysis, we used the average amount of physical activity measured by the 2 questionnaires administered closest in time before and after blood draw as the main exposure. Participants were asked about the amount of time that they spent per week, on average, in each of the following physical activities: walking; jogging; running; bicycling; calisthenics, aerobics, aerobic dancing, or rowing-machine use; lap swimming; weight-lifting; playing tennis; and playing squash or racquetball. The amount of total reported physical activity was calculated as weekly energy expenditure in metabolic equivalent of task (MET)-hours per week (13). MET-hours/week for vigorous activities, defined as activities requiring ≥6 METs, were calculated from jogging, running, cycling, swimming, and playing tennis. MET-hours/week for nonvigorous activities (<6 METs) were calculated from brisk walking and stair-climbing. The reproducibility and validity of the physical activity questionnaire have been described elsewhere (14). Physical activity was assessed at baseline in the WHI using a similar self-report questionnaire that has demonstrated reproducibility and validity (15).

Metabolomics profiling

In the NHS, NHS II, HPFS, and WHI, metabolomic profiling data were obtained using liquid chromatography–mass spectrometry at the Broad Institute of MIT and Harvard (Cambridge, Massachusetts). A detailed description of the metabolite profiling methods has been previously published (10). In brief, liquid chromatography uses multiple stationary phase chemistries and provides reproducible separation of metabolites in complex mixtures. Mass spectrometry enables further resolution of metabolites based on mass:charge ratio and quantification over a wide linear dynamic range. The raw data were processed using TraceFinder 3.0 software (Thermo Fisher Scientific, Waltham, Massachusetts) to integrate chromatographic peaks of known identity, and the data were visually inspected to ensure the quality of signal integration. For the current study,

we excluded unknown metabolites or metabolites with no variance. Blood samples from the 10 studies nested within the NHS, NHS II, and HPFS that contributed samples to this analysis were processed over the course of 1 year (October 2016–August 2017). A total of 337 known metabolites were identified by combining data from the NHS, NHS II, and HPFS; 256 of these were classified as lipids, including 13 cholesteryl esters (CEs), 84 phospholipids, 81 glycerolipids, 43 sphingolipids, and 33 free fatty acids.

Covariates

Information on age, smoking status, and body weight was collected in the biennial follow-up questionnaires in the NHS, NHS II, and HPFS and at baseline in the WHI. Dietary data were collected using a 131-item food frequency questionnaire. For categorical variables, we used the questionnaire administered closest in time before blood collection, and for continuous variables, we used the average of dietary intakes derived from the 2 questionnaires administered closest in time before and after blood draw as covariates.

Statistical analysis

Within each nested case-control study, metabolite data were log-transformed and converted to z scores with a mean of 0 and a standard deviation of 1. We examined the association of physical activity with each metabolite using linear regression adjusting for age at blood draw (years; continuous), case status in the original study (yes, no), fasting status (yes, no), smoking status (current, former, or never smoker), body mass index (BMI) (defined as weight (kg)/height $(m)^2$; continuous), Alternative Healthy Eating Index score (0–100; continuous), total energy intake (kcal/day; continuous), alcohol intake (g/day; continuous), and contributing case-control study (16 categories representing each combination of endpoint and cohort). We used the tail probability of the proportion of false positives (TPPFP) to account for multiple testing. The TPPFP is defined as $\Pr(FDP > q)$, where the false discovery proportion (FDP) is the proportion of significant tests that are false-positive (tests that incorrectly reject the null value). A bootstrapping procedure was used to choose test statistic rejection regions such that $TPPFP = \Pr(FDP > q) < \alpha$ under the global null hypothesis that no metabolite is associated with physical activity. Multiple-testing-adjusted P values for individual metabolites were defined as the smallest α value such that the metabolite is significant and the overall $TPPFP < \alpha$. We set the target $TPPFP\ q = 10\%$ to balance type I and type II errors in the context of a marker discovery study. Aside from controlling the experiment-wide error, this bootstrap procedure accounts for correlation among metabolites. We used the bootstrap procedure implemented in the MTP( ) function in the "multtest" Bioconductor package in R (R Foundation for Statistical Computing, Vienna, Austria) (16) and set the number of bootstrap samples to 10,000.

We examined the direct associations of physical activity with metabolite levels and the indirect association statistically accounted for by BMI using the method proposed by VanderWeele and Vansteelandt (17) and Valeri and VanderWeele (18). Because our primary study was a pooled analysis, we conducted separate sensitivity analyses by cohort and combined the results using meta-analysis. We tested for potential effect modification by sex and cohort using the likelihood ratio test by including interaction terms for the interaction of sex and cohort with physical activity in the multivariable linear model.

We further performed enrichment analysis to examine whether physical activity–metabolite associations tended to cluster in specific classes of metabolites (11). We focused on 5 categories of metabolites: CEs, phospholipids, glycerolipids, sphingolipids, and amino acids. We obtained enrichment odds ratios comparing the number of metabolites that were significantly associated with physical activity (nominal $P < 0.05$) within a metabolite category to the rest of the metabolites. Given that metabolites within the same category were correlated with each other, we performed permutation tests to obtain the distribution of the odds ratios under the global null hypothesis that all odds ratios were equal to 1. We permuted metabolite vectors within joint strata of sex (female, male), case status in the original study (yes, no), and BMI (<25, 25–29, or ≥30). Permutation testing was repeated 10,000 times to obtain distributions of the odds ratios, and a 1-sided P value was calculated as follows: (frequency of permuted odds ratios greater than the observed odds ratio + 1)/(number of permutations + 1) (11). The analyses were performed using R, version 3.2.5, and SAS, version 9.2 (SAS Institute, Inc., Cary, North Carolina).

RESULTS

Our primary study included 5,197 participants, with 2,055 participants from the NHS, 2,120 participants from NHS II, and 1,022 participants from the HPFS. Characteristics of study participants are shown in Table 1. Mean values for physical activity, BMI, total energy intake, Alternative Healthy Eating Index score, and intakes of alcohol and fatty acids, as well as proportions of never smokers, were comparable between participants included in our study and underlying participants who provided blood specimens (see Web Table 1, available at https://academic.oup.com/aje). In the replication cohort, we included 2,305 participants from the WHI; their baseline characteristics are shown in Web Table 2.

Physical activity was significantly associated with 36 metabolites before adjustment for BMI ($TPPFP < 0.05$) (Table 2). Twenty-three of the identified metabolites showed positive associations, including 3 amino acids (glycine, citrulline, asparagine), 5 CEs, 9 phosphocholines (PCs) and lysophosphatidylcholines (LPCs), and 5 phosphatidylethanolamines (PEs) and lysophosphatidylethanolamines (LPEs). However, we also observed that physical activity was inversely associated with glutamate, 4 triglycerides, and 5 diglycerides. Given that BMI might be an important mediator between physical activity and plasma metabolites, we examined associations of physical activity with metabolites while additionally adjusting for BMI. The positive associations of physical activity with the majority of metabolites remained significant, including associations with 2 amino acids (citrulline and glycine), 4 CEs (C18:2, C18:1, C16:0, and C18:3), 8 PCs and LPCs (C36:4 PC-A, C34:3 PC plasmalogen, C36:3 PC plasmalogen, C34:2 PC plasmalogen, C36:2 PC, C18:2 LPC, C20:5 LPC, and C18:1 LPC), and 3 LPEs and PEs (C18:2 LPE, C18:1 LPE, and C38:3 PE plasmalogen). However, the glutamate, diglycerides, and triglycerides that were inversely associated with physical

Table 1. Characteristics of Participants in the Nurses' Health Study (1989–1990), Nurses' Health Study II (1996–1999), and the Health Professionals Follow-up Study (1993–1995)

Characteristic	NHS (n = 2,055)			NHS II (n = 2,120)			HPFS (n = 1,022)		
	Mean (SD)	%	No.	Mean (SD)	%	No.	Mean (SD)	%	No.
Physical activity, MET-hours/week	18.3 (19.5)			20.3 (20.6)			40.5 (38.0)		
Vigorous activity	7.1 (14.5)			12.1 (17.0)			15.0 (23.2)		
Jogging	0.3 (2.5)			1.0 (4.0)			1.4 (4.1)		
Running	0.4 (4.7)			1.3 (7.3)			2.6 (13.7)		
Cycling	2.1 (5.4)			2.0 (4.8)			3.7 (8.6)		
Swimming	1.1 (4.3)			0.7 (3.2)			0.9 (3.8)		
Tennis	1.0 (5.5)			0.6 (3.3)			3.1 (11.2)		
Nonvigorous activity	8.5 (8.3)			7.0 (6.7)			26.4 (29.9)		
Brisk walking	7.8 (8.2)			6.2 (6.6)			14.7 (16.2)		
Stair-climbing	0.7 (0.5)			0.7 (0.5)			0.4 (0.4)		
Age at blood draw, years	56.6 (6.8)			44.7 (4.5)			64.0 (8.2)		
Body mass index[a]	24.6 (5.5)			25.2 (5.0)			23.0 (6.0)		
Total energy intake, kcal/day	1,799.0 (454.0)			1,854.0 (488.0)			2,054.0 (565.0)		
AHEI score[b]	53.6 (9.7)			53.8 (10.1)			55.9 (10.6)		
Alcohol intake, g/day	7.9 (10.0)			6.0 (7.6)			14.8 (14.5)		
Fatty acid intake, g/day									
Linolenic acid	10.5 (3.8)			9.5 (3.6)			11.1 (4.4)		
α-Linolenic acid	1.1 (0.4)			1.1 (0.4)			1.1 (0.4)		
Eicosapentaenoic acid	0.1 (0.1)			0.1 (0.1)			0.1 (0.2)		
Docosahexaenoic acid	0.2 (0.1)			0.1 (0.1)			0.2 (0.2)		
Oleic acid	22.2 (7.6)			22.7 (8.3)			25.5 (10.3)		
Stearic acid	5.5 (2.0)			5.4 (2.1)			5.8 (2.6)		
Trans- fatty acid	2.9 (1.3)			3.1 (1.4)			3.3 (1.8)		
Smoking status									
Never smoker		40	823		61	1,299		42	430
Past smoker		46	946		30	628		52	530
Current smoker		14	286		9	193		6	62
Participants selected as cases in nested case-control studies		44	896		42	881		56	569
Fasting at time of blood collection		81	1,671		70	1,491		60	616

Abbreviations: AHEI, Alternative Healthy Eating Index; HPFS, Health Professionals Follow-up Study; MET, metabolic equivalent of task; NHS, Nurses' Health Study; NHS II, Nurses' Health Study II; SD, standard deviation.

[a] Body mass index is calculated as weight (kg)/height $(m)^2$.

[b] The AHEI score ranges from 0 to 100, with a higher score indicating a healthier diet.

activity were no longer significant after adjustment for BMI. In the WHI replication analysis, associations with 19 out of 36 metabolites were replicated in analyses unadjusted for BMI (nominal $P < 0.05$ and the same directionality) (Table 3). Physical activity was inversely associated with all diglycerides and triglycerides before adjustment for BMI. Physical activity also remained nominally significantly positively associated with 9 of the 19 identified metabolites in the WHI after adjustment for BMI (Table 3).

The metabolites identified after adjusting for BMI were associated with both vigorous and nonvigorous types of physical activity (Web Figure 1). Specifically, the identified metabolites were mainly associated with the most common forms of total physical activity, such as running and playing tennis for vigorous physical activity and brisk walking for nonvigorous activity. We further performed mediation analysis to examine the extent to which the associations of physical activity with the 36 metabolites identified in Table 2 might be mediated through BMI. We observed that BMI largely mediated the inverse associations of physical activity with glutamate, triglycerides, and diglycerides (Table 4). We examined whether associations between physical activity and metabolites tended to cluster in specific metabolite classes using enrichment analysis, and we found that the CE category was nominally significantly enriched for associations with physical activity (P for enrichment = 0.05) (Table 5).

Table 2. Metabolites Significantly Associated With Physical Activity in Pooled Data From Samples of Participants in the Nurses' Health Study (1989–1990), Nurses' Health Study II (1996–1999), and the Health Professionals Follow-up Study (1993–1995)

Metabolite	HMDB ID	Multivariable Model[a]		Multivariable Model Further Adjusting for BMI[b]	
		β (SE) (×10^{-3})	TPPFP *P* Value	β (SE) (×10^{-3})	TPPFP *P* Value
Cholesteryl esters					
C18:2 CE	HMDB00610	5.13 (0.80)	<0.001	4.66 (0.79)	<0.001
C18:1 CE	HMDB00918	3.45 (0.79)	<0.001	3.02 (0.79)	<0.001
C16:0 CE	HMDB00885	3.25 (0.79)	<0.001	2.84 (0.79)	0.04
C18:3 CE	HMDB10370	2.71 (0.81)	0.02	2.84 (0.81)	<0.001
C22:4 CE	HMDB06729	4.05 (1.30)	0.02	2.69 (1.29)	0.71
Phospholipids					
C18:2 LPC	HMDB10386	4.99 (0.80)	<0.001	4.46 (0.80)	<0.001
C20:5 LPC	HMDB10397	4.18 (0.81)	<0.001	3.82 (0.81)	<0.001
C18:1 LPC	HMDB02815	4.08 (0.81)	<0.001	3.61 (0.80)	<0.001
C36:4 PC-A	HMDB07983	4.68 (0.81)	<0.001	4.48 (0.81)	<0.001
C34:3 PC plasmalogen	HMDB11211	4.33 (0.80)	<0.001	3.89 (0.80)	<0.001
C36:3 PC plasmalogen	HMDB11244	3.88 (0.79)	<0.001	3.36 (0.79)	<0.001
C36:2 PC plasmalogen	HMDB11243	3.42 (0.79)	<0.001	2.90 (0.78)	0.06
C34:2 PC plasmalogen	HMDB11210	3.35 (0.79)	0.02	2.83 (0.78)	0.04
C36:2 PC	HMDB08039	2.63 (0.81)	0.10	2.83 (0.81)	0.02
C36:0 PC	HMDB08036	3.07 (0.81)	<0.001	2.91 (0.81)	0.06
C18:2 LPE	HMDB11507	3.12 (0.70)	<0.001	2.91 (0.70)	<0.001
C18:1 LPE	HMDB11506	2.91 (0.70)	<0.001	2.70 (0.71)	0.04
C38:3 PE plasmalogen	HMDB11384	3.22 (0.78)	<0.001	2.75 (0.78)	0.04
C40:6 PE	HMDB09012	−2.96 (0.78)	0.02	−2.61 (0.78)	0.16
C38:2 PE	HMDB08942	2.73 (0.81)	0.04	2.41 (0.81)	0.16
C34:0 PI	HMDB09805	3.69 (0.81)	<0.001	3.83 (0.81)	<0.001
Glycerolipids					
C34:3 DG	HMDB07132	−3.32 (0.81)	<0.001	−2.59 (0.80)	0.06
C34:2 DG	HMDB07103	−3.13 (0.80)	<0.001	−2.36 (0.79)	0.14
C32:1 DG	HMDB07099	−3.00 (0.80)	<0.001	−2.23 (0.79)	0.27
C34:1 DG	HMDB07102	−2.95 (0.80)	0.02	−2.16 (0.79)	0.29
C32:0 DG	HMDB07098	−2.80 (0.80)	0.04	−2.03 (0.79)	0.51
C50:2 TG	HMDB05377	−3.10 (0.80)	<0.001	−2.31 (0.78)	0.25
C52:2 TG	HMDB05369	−2.98 (0.80)	<0.001	−2.19 (0.79)	0.25
C50:3 TG	HMDB05433	−2.85 (0.80)	<0.001	−2.11 (0.79)	0.43
C50:1 TG	HMDB05360	−2.72 (0.80)	0.04	−1.91 (0.78)	0.70
Sphingolipids					
C22:1 SM	HMDB12104	2.40 (0.80)	0.13	2.75 (0.80)	0.04
C18 carnitine	HMDB00848	2.00 (0.56)	0.02	1.67 (0.57)	0.16
Amino acids					
Glycine	HMDB00123	2.60 (0.57)	<0.001	1.89 (0.56)	0.04
Citrulline	HMDB00904	2.16 (0.55)	<0.001	1.90 (0.55)	<0.001
Glutamate	HMDB00148	−2.19 (0.57)	<0.001	−1.34 (0.56)	0.79
Asparagine	HMDB00168	1.89 (0.57)	<0.001	1.43 (0.57)	0.46

Table continues

Table 2. Continued

Metabolite	HMDB ID	Multivariable Model[a]		Multivariable Model Further Adjusting for BMI[b]	
		β (SE) (×10^{-3})	TPPFP *P* Value	β (SE) (×10^{-3})	TPPFP *P* Value
Others					
1-Methyladenosine	HMDB03331	−3.02 (0.56)	<0.001	−2.24 (0.56)	0.02
Creatine	HMDB00064	−2.01 (0.57)	0.02	−1.60 (0.57)	0.37

Abbreviations: BMI, body mass index; CE, cholesteryl ester; DG, diglyceride; HMDB, Human Metabolome Database; ID, identification number; LPC, lysophosphatidylcholines; LPE, lysophosphatidylethanolamine; PC, phosphocholine; PC-A, phosphocholine subtype A; PE, phosphatidylethanolamine; PI, phosphatidylinositol; SE, standard error; SM, sphingomyelin; TG, triglyceride; TPPFP, tail probability of the proportion of false positives.

[a] Linear regression analysis with adjustment for age (years; continuous), case-control status (case, control), case-control study nested within the 3 cohort studies (16 categories), smoking status (never, past, or current smoker), fasting status at blood collection (yes, no), alcohol intake (g/day; continuous), Alternative Healthy Eating Index score (0–100; continuous), total energy intake (kcal/day; continuous), and intakes of fatty acids (linolenic acid, α-linolenic acid, eicosapentaenoic acid, docosahexaenoic acid, oleic acid, stearic acid, and *trans*- fatty acids).

[b] BMI is calculated as weight (kg)/height (m)^2.

Nominally significant effect modification by cohort (*P* for heterogeneity < 0.05) was found for several metabolites identified in the pooled analysis (Web Tables 3 and 4). We also conducted separate analysis by cohort and combined the results using meta-analysis (Web Tables 5 and 6). There was no significant heterogeneity in associations after accounting for the number of metabolites tested. The estimates of associations between metabolites and physical activity from the meta-analysis and the pooled analysis were similar. We did not identify additional metabolites significantly associated with physical activity in any single cohort, consistent with the finding of no significant heterogeneity across cohorts. No heterogeneity was observed for identified metabolites by sex when comparing the NHS and NHS II with the HPFS (*P* for heterogeneity > 0.05). We conducted sensitivity analysis by carrying out metabolome-wide association studies restricted to controls in the 3 discovery cohorts. The directions of associations between physical activity and metabolites were consistent for 90% of all metabolites among controls and the whole population, and the TPPFP remained significant for 15 out of the 20 identified metabolites among controls (Web Table 7). In addition, physical activity was not associated with BCAAs when pooling data from the NHS, NHS II, and HPFS (valine: $\beta = -0.8 \times 10^{-3}$ (standard error, 0.6×10^{-3}), TPPFP = 1.00; isoleucine: $\beta = -0.2 \times 10^{-3}$ (standard error, 0.6×10^{-3}), TPPFP = 1.00; leucine: $\beta = -0.3 \times 10^{-3}$ (standard error, 0.6×10^{-3}), TPPFP = 1.00).

DISCUSSION

In a pooled analysis of 5,197 participants from 4 large cohort studies of men and women, habitual physical activity was significantly associated with 20 metabolites after adjustment for BMI and other factors, and most of the metabolites were CEs and phospholipids. Physical activity remained significantly associated with half of the identified metabolites in the replication cohort of postmenopausal women.

In our study, the majority of the 337 metabolites measured were lipids, including CEs, triglycerides, PCs, LPCs, PEs, LPEs, and carnitines of different chain lengths and fatty-acid desaturations, and it allowed us to identify novel lipid biomarkers associated with physical activity. In our discovery cohort, we identified 4 CEs, 8 PCs and LPCs, and 3 LPEs and PEs that were significantly associated with physical activity independently of BMI. Furthermore, most of the identified CEs and LPCs were replicable in the WHI. CEs and phospholipids are mainly located in lipoproteins in plasma, particularly high-density lipoprotein (HDL). The nascent HDL consists of cholesterol in the core and phospholipids, including PC and sphingomyelin, on the surface. HDL removes excess cholesterol from peripheral tissues and delivers it to the liver for excretion. During this reverse cholesterol transport process, lecithin:cholesterol acyl transferase on the HDL transfers fatty acids from PC to cholesterol, resulting in the formation of CE and LPC (19). Being physically active increases overall HDL level and accelerates cholesterol efflux independent of changes in body weight (20, 21), leading to lower risks of coronary heart disease and type 2 diabetes (22, 23). This may explain why we found that physical activity was positively associated with several CEs, PCs, and LPCs even after adjustment for BMI and positively associated with the CE lipid category in enrichment analysis. Although low-density lipoprotein (LDL) also contains CE and PC, whether physical activity lowers LDL level remains inconclusive (24, 25). Physical activity reduced LDL best when there was body fat loss (26), and more intense activity was required to elicit reductions in LDL than increases in HDL (27). However, in our study, we could not distinguish the CEs and PCs from HDL or LDL.

Glycerides, including triglyceride and diglyceride, are absorbed from the intestines and transported to peripheral tissues in the form of chylomicrons and very low-density lipoprotein. Habitual physical activity lowers plasma glyceride levels by increasing the activity of lipoprotein lipase, which is responsible for chylomicrons and very low-density lipoprotein glyceride hydrolysis (28). Physical activity oxidizes fatty acids in muscle and adipose tissues and lowers body weight (29). In one study involving 53 participants, Huffman et al. (29) found that 6 months of aerobic exercise training promoted mitochondrial β-oxidation. This might be the reason why the positive associations between physical activity and triglycerides and diglycerides found in our study were largely mediated by BMI.

Table 3. Replication of Metabolites Associated With Physical Activity Among 2,305 Participants in the Women's Health Initiative, 1993

Metabolite	HMDB ID	Multivariable Model[a]		Multivariable Model Further Adjusting for BMI[b]	
		β (SE) (×10^{-3})	*P* Value	β (SE) (×10^{-3})	*P* Value
Cholesteryl esters					
C18:2 CE	HMDB00610	2.52 (1.73)	0.15	1.64 (1.74)	0.35
C18:1 CE	HMDB00918	5.46 (1.72)	0.002	5.02 (1.74)	0.004
C16:0 CE	HMDB00885	4.11 (1.71)	0.02	3.82 (1.73)	0.03
C18:3 CE	HMDB10370	3.00 (1.71)	0.08	3.23 (1.72)	0.06
C22:4 CE	HMDB06729	4.24 (1.68)	0.01	2.55 (1.67)	0.13
Phospholipids					
C18:2 LPC	HMDB10386	7.41 (1.70)	<0.001	5.14 (1.68)	0.002
C20:5 LPC	HMDB10397	6.86 (1.67)	<0.001	5.49 (1.66)	0.001
C18:1 LPC	HMDB02815	6.46 (1.71)	<0.001	4.56 (1.69)	0.01
C36:4 PC-A	HMDB07983	3.55 (1.70)	0.04	2.45 (1.70)	0.15
C34:3 PC plasmalogen	HMDB11211	7.74 (1.70)	<0.001	5.77 (1.68)	<0.001
C36:3 PC plasmalogen	HMDB11244	3.22 (1.73)	0.06	1.86 (1.73)	0.28
C36:2 PC plasmalogen	HMDB11243	3.21 (1.72)	0.06	1.48 (1.71)	0.39
C34:2 PC plasmalogen	HMDB11210	3.82 (1.74)	0.03	2.52 (1.74)	0.15
C36:2 PC	HMDB08039	1.16 (1.71)	0.50	0.89 (1.72)	0.61
C36:0 PC	HMDB08036	4.48 (1.74)	0.01	3.40 (1.74)	0.05
C18:2 LPE	HMDB11507	3.14 (1.66)	0.06	2.20 (1.67)	0.19
C18:1 LPE	HMDB11506	2.89 (1.71)	0.09	1.96 (1.72)	0.25
C38:3 PE plasmalogen	HMDB11384	5.30 (1.74)	0.002	3.96 (1.74)	0.02
C40:6 PE	HMDB09012	−1.18 (1.68)	0.48	−0.35 (1.69)	0.84
C38:2 PE	HMDB08942	4.23 (1.68)	0.01	2.38 (1.66)	0.15
C34:0 PI[c]	HMDB09805				
Glycerolipids					
C34:3 DG	HMDB07132	−4.95 (1.65)	0.003	−3.15 (1.64)	0.05
C34:2 DG	HMDB07103	−5.68 (1.68)	<0.001	−3.42 (1.65)	0.04
C32:1 DG	HMDB07099	−4.77 (1.67)	0.004	−2.57 (1.64)	0.12
C34:1 DG	HMDB07102	−5.81 (1.70)	<0.001	−3.24 (1.66)	0.05
C32:0 DG	HMDB07098	−6.00 (1.70)	<0.001	−3.46 (1.66)	0.04
C50:2 TG	HMDB05377	−5.77 (1.67)	<0.001	−3.29 (1.64)	0.04
C52:2 TG	HMDB05369	−5.72 (1.69)	<0.001	−3.26 (1.65)	0.05
C50:3 TG	HMDB05433	−4.14 (1.65)	0.01	−2.26 (1.64)	0.17
C50:1 TG	HMDB05360	−5.96 (1.70)	<0.001	−3.16 (1.65)	0.06
Sphingolipids					
C22:1 SM	HMDB12104	3.57 (1.73)	0.04	3.38 (1.74)	0.05
C18 carnitine	HMDB00848	0.25 (1.66)	0.88	−0.27 (1.67)	0.87
Amino acids					
Glycine	HMDB00123	8.23 (1.71)	<0.001	6.02 (1.68)	<0.001
Citrulline	HMDB00904	2.76 (1.69)	0.10	1.54 (1.69)	0.36
Glutamate	HMDB00148	−4.14 (1.68)	0.01	−2.86 (1.68)	0.09
Asparagine	HMDB00168	6.40 (1.64)	<0.001	4.10 (1.61)	0.01

Table continues

Table 3. Continued

Metabolite	HMDB ID	Multivariable Model[a]		Multivariable Model Further Adjusting for BMI[b]	
		β (SE) (×10^{-3})	*P* Value	β (SE) (×10^{-3})	*P* Value
Others					
1-Methyladenosine	HMDB03331	−5.10 (1.72)	0.003	−3.44 (1.71)	0.04
Creatine	HMDB00064	−2.04 (1.74)	0.24	−1.79 (1.75)	0.31

Abbreviations: BMI, body mass index; CE, cholesteryl ester; DG, diglyceride; HMDB, Human Metabolome Database; ID, identification number; LPC, lysophosphatidylcholines; LPE, lysophosphatidylethanolamine; PC, phosphocholine; PC-A, phosphocholine subtype A; PE, phosphatidylethanolamine; PI, phosphatidylinositol; SE, standard error; SM, sphingomyelin; TG, triglyceride.

[a] Multivariable model adjusting for age (years; continuous), coronary heart disease case-control status (case, control), Women's Health Initiative substudy (Observational Study, Estrogen-Alone Trial, or Estrogen Plus Progestin Trial), smoking status (never, past, or current smoker), fasting status (yes, no), alcohol intake (g/day; continuous), Alternative Healthy Eating Index score (0–100; continuous), total energy intake (kcal/day; continuous), and racial/ethnic group (white, black, Hispanic, Asian, or other).

[b] BMI is calculated as weight (kg)/height (m)^2.

[c] No data were available for this metabolite in the Women's Health Initiative.

As an emerging approach, one strength of plasma metabolite profiling by liquid chromatography–mass spectrometry lies in the measurement of individual lipids with different carbon numbers and double bonds. In our study, most of the identified CEs, LPCs, PCs, and LPEs contained unsaturated fatty acid chains, particularly C18:1 and C18:2. There are 2 possible explanations for this. First, studies have shown that C18:1 and C18:2 are the most abundant fatty acids across all lipoproteins, particularly HDL and LDL (30). Second, C18:2 fatty acid (linoleic acid) is an essential fatty acid found mostly in plant oils. A plant-based diet has been found to be associated with a lower risk of cardiometabolic diseases (31), and being physically active is highly correlated with the adoption of a healthy diet (32). Therefore, the relationships with identified lipids in our study might have been confounded by dietary intakes of fatty acids; however, we controlled for a variety of dietary fatty acids, and the associations with physical activity persisted. Our study showed that physical activity was inversely associated with triglyceride with saturated (C16:0 and C18:0) and monounsaturated (C16:1, C18:1) fatty acid chains before adjustment for BMI (triglyceride: C50:3[16:0/18:3/16:0], C52:2[16:0/18:1/18:1], C50:2[16:0/16:1/18:1], C50:1[16:0/16:1/18:0]). Hydrolysis of these triglycerides releases saturated and monosaturated fatty acids. Oxidation of saturated and monounsaturated fatty acids provides more energy than oxidation of polyunsaturated fatty acids of the same length, and oxidation of saturated fatty acids in the human body decreases with increasing carbon number (33). Thus, it is possible that physical activity increases lipolysis of triglycerides with a low carbon number and double bond content and enhances β-oxidation of fatty acids released by these triglycerides within the mitochondria. One randomized trial showed that being physically inactive decreased the oxidation of saturated dietary fat (34). Another trial showed that exercise led to a faster adaptation to a high-fat diet featuring high intakes of saturated and monounsaturated fat by increasing fat oxidation (35). Similarly, previous studies also found that triglycerides with a lower carbon number and a double bond were most strongly associated with higher risks of cardiovascular disease and type 2 diabetes (36, 37), for which being physically inactive is an important risk factor.

Previous cohort studies have shown that plasma glutamate concentrations are positively associated with risk of coronary heart disease (38–40), higher BMI (41), and insulin resistance (42, 43). Consistent with these findings, our study showed that higher amounts of physical activity, a protective factor against cardiometabolic diseases, were inversely associated with plasma glutamate levels, and the inverse association became nonsignificant after adjustment for BMI. The mechanism by which physical activity may lower plasma glutamate levels might be related to metabolism of BCAAs, since glutamate is one intermediate of BCAA catabolism that is activated by exercise and high energy expenditure (44–49). Although BCAAs have been shown to be associated with higher risks of coronary heart disease (50, 51), type 2 diabetes (52, 53), and obesity (54), we did not find significant associations of physical activity with BCAA, and the reason is worth further exploration.

Our study had several strengths. First, our large sample size provided sufficient statistical power to identify metabolites associated with physical activity. Second, given the correlation among metabolites between categories and within the same category, we used a nonparametric bootstrap procedure to account for correlation among metabolites, to efficiently correct for multiple testing. Third, we replicated results in an independent data set, and physical activity was nominally significantly associated with many of the identified metabolites in the discovery study, particularly CEs and LPCs.

Our study also had several limitations. First, physical activity was measured via self-report questionnaire, which might be prone to measurement error. Second, our study was observational, and therefore cause-effect relationships could not be established. Third, the cohorts included in the discovery study were predominantly white, and all participants were medical professionals. This may limit the generalizability of our findings; additional studies in more diverse populations are needed. Fourth, participants in the discovery phase were selected from nested case-control studies of 10 diverse outcomes, with metabolites measured in the same laboratory but at different times. This may have led to ascertainment and measurement biases. However, these potential biases are unlikely to create spurious associations (i.e., inflate type I error), and our replication and

Table 4. Direct and Indirect Effects (Mediated Through Body Mass Index) of Physical Activity on Metabolite Levels in Samples of Participants From the Nurses' Health Study (1989–1990), Nurses' Health Study II (1996–1999), and the Health Professionals Follow-up Study (1993–1995)[a]

Metabolite	HMDB ID	Effect, β (SE) ($\times 10^{-3}$)			% of Effect Mediated by BMI
		Direct Effect	Indirect Effect (Mediated by BMI[b])	Total Effect	
Cholesteryl esters					
C18:2 CE	HMDB00610	4.66 (0.79)	0.52 (0.11)	5.18 (0.79)	9.98
C18:1 CE	HMDB00918	3.02 (0.79)	0.48 (0.11)	3.49 (0.79)	13.68
C16:0 CE	HMDB00885	2.84 (0.79)	0.45 (0.10)	3.29 (0.79)	13.57
C18:3 CE	HMDB10370	2.84 (0.81)	−0.14 (0.09)	2.70 (0.81)	−5.14[c]
C22:4 CE	HMDB06729	2.69 (1.29)	0.84 (0.16)	3.53 (1.28)	23.89
Phospholipids					
C18:2 LPC	HMDB10386	4.46 (0.80)	0.58 (0.12)	5.04 (0.80)	11.44
C20:5 LPC	HMDB10397	3.82 (0.81)	0.39 (0.10)	4.22 (0.81)	9.30
C18:1 LPC	HMDB02815	3.61 (0.80)	0.52 (0.11)	4.13 (0.80)	12.60
C36:4 PC-A	HMDB07983	4.48 (0.81)	0.22 (0.09)	4.70 (0.81)	4.76
C34:3 PC plasmalogen	HMDB11211	3.89 (0.80)	0.48 (0.11)	4.37 (0.80)	11.02
C36:3 PC plasmalogen	HMDB11244	3.36 (0.79)	0.57 (0.11)	3.93 (0.79)	14.54
C36:2 PC plasmalogen	HMDB11243	2.90 (0.78)	0.57 (0.11)	3.47 (0.78)	16.42
C34:2 PC plasmalogen	HMDB11210	2.83 (0.78)	0.57 (0.11)	3.40 (0.78)	16.63
C36:0 PC	HMDB08036	2.91 (0.81)	0.18 (0.09)	3.08 (0.81)	5.68
C18:2 LPE	HMDB11507	2.91 (0.70)	0.22 (0.08)	3.12 (0.70)	6.95
C18:1 LPE	HMDB11506	2.70 (0.71)	0.22 (0.08)	2.92 (0.70)	7.58
C38:3 PE plasmalogen	HMDB11384	2.75 (0.78)	0.51 (0.11)	3.26 (0.78)	15.78
C40:6 PE	HMDB09012	−2.61 (0.78)	−0.38 (0.10)	−2.99 (0.78)	12.62[c]
C38:2 PE	HMDB08942	2.41 (0.81)	0.35 (0.10)	2.76 (0.81)	12.76
C34:0 PI	HMDB09805	3.83 (0.81)	−0.14 (0.09)	3.68 (0.81)	−3.88[c]
Glycerolipids					
C34:3 DG	HMDB07132	−2.59 (0.80)	−0.79 (0.13)	−3.39 (0.80)	23.36
C34:2 DG	HMDB07103	−2.36 (0.79)	−0.84 (0.14)	−3.20 (0.80)	26.33
C32:1 DG	HMDB07099	−2.23 (0.79)	−0.84 (0.14)	−3.07 (0.79)	27.40
C34:1 DG	HMDB07102	−2.16 (0.79)	−0.87 (0.14)	−3.03 (0.79)	28.71
C32:0 DG	HMDB07098	−2.03 (0.79)	−0.84 (0.14)	−2.87 (0.79)	29.37
C50:2 TG	HMDB05377	−2.31 (0.78)	−0.86 (0.14)	−3.17 (0.79)	27.18
C52:2 TG	HMDB05369	−2.19 (0.79)	−0.86 (0.14)	−3.05 (0.80)	28.15
C50:3 TG	HMDB05433	−2.11 (0.79)	−0.81 (0.14)	−2.92 (0.80)	27.73
C50:1 TG	HMDB05360	−1.91 (0.78)	−0.89 (0.14)	−2.79 (0.79)	31.74
Sphingolipids					
C18 carnitine	HMDB00848	1.67 (0.57)	0.33 (0.07)	2.00 (0.56)	16.31
Amino acids					
Glycine	HMDB00123	1.89 (0.56)	0.71 (0.11)	2.60 (0.57)	27.33
Citrulline	HMDB00904	1.90 (0.55)	0.27 (0.07)	2.16 (0.55)	12.28
Glutamate	HMDB00148	−1.34 (0.56)	−0.84 (0.12)	−2.19 (0.57)	38.65
Asparagine	HMDB00168	1.43 (0.57)	0.46 (0.09)	1.89 (0.57)	24.25

Table continues

Table 4. Continued

Metabolite	HMDB ID	Effect, β (SE) (×10^{-3})			% of Effect Mediated by BMI
		Direct Effect	Indirect Effect (Mediated by BMI[b])	Total Effect	
Others					
1-Methyladenosine	HMDB03331	−2.24 (0.56)	−0.77 (0.12)	−3.02 (0.56)	25.69
Creatine	HMDB00064	−1.60 (0.57)	−0.41 (0.08)	−2.01 (0.57)	20.31

Abbreviations: BMI, body mass index; CE, cholesteryl ester; DG, diglyceride; HMDB, Human Metabolome Database; ID, identification number; LPC, lysophosphatidylcholines; LPE, lysophosphatidylethanolamine; PC, phosphocholine; PE, phosphatidylethanolamine; PI, phosphatidylinositol; SE, standard error; TG, triglyceride.

[a] Linear regression analysis with adjustment for age (years; continuous), case-control status (case, control), case-control study nested within the 3 cohort studies (16 categories), smoking status (never, past, or current smoker), fasting status at blood collection (yes, no), alcohol intake (g/day; continuous), Alternative Healthy Eating Index score (0–100; continuous), total energy intake (kcal/day; continuous), and intakes of fatty acids (linolenic acid, α-linolenic acid, eicosapentaenoic acid, docosahexaenoic acid, oleic acid, stearic acid, and *trans*- fatty acids).

[b] BMI is calculated as weight (kg)/height (m)^2.

[c] The mediation effect was not significant.

sensitivity analyses (restricting to controls in the original studies, performing analyses stratified by cohort) suggested that the observed associations were unlikely to be solely due to bias. Finally, the metabolites measured in our study did not include glucose; therefore, it is impossible to compare our study with previous studies showing that physical activity was associated with lower levels of carbohydrate metabolites. Furthermore, although our study identified positive associations for CEs and phospholipids that can be explained in the context of lipoproteins, our study did not measure metabolites contained in lipoproteins, which is an important area for future research.

In summary, in this analysis of 337 metabolites, most of which were lipids, in 5,197 men and women, we identified 20 metabolites that were significantly associated with physical activity independently of BMI. These findings may help identify biomarkers of physical activity and provide insight into biological mechanisms underlying the beneficial effect of being physically active on cardiometabolic health.

Table 5. Associations of Physical Activity With Metabolite Groups, Calculated Using Enrichment Analysis, in Samples of Participants From the Nurses' Health Study (1989–1990), Nurses' Health Study II (1996–1999), and the Health Professionals Follow-up Study (1993–1995)

Metabolite Group	Multivariable Model[a]		Multivariable Model Further Adjusting for BMI[b,c]	
	Enrichment OR	*P* for Enrichment[d]	Enrichment OR	*P* for Enrichment[d]
Cholesteryl esters	2.98	0.14	6.19	0.05
Phospholipids	1.69	0.25	2.46	0.14
Glycerolipids	1.40	0.26	0.79	0.44
Sphingolipids	0.50	0.59	0.86	0.43
Amino acids	1.36	0.35	0.80	0.57

Abbreviations: BMI, body mass index; OR, odds ratio.

[a] Permutation was performed within joint strata of sex (female, male) and case status in the original study (yes, no). The enrichment OR was calculated from the 2 × 2 table of metabolites (in/out of group vs. significant/nonsignificant). Determination of whether individual metabolites were significantly associated with physical activity (nominal $P < 0.05$) was obtained using linear regression analysis adjusting for age (years; continuous), case-control status (case, control), case-control study nested within the 3 cohort studies (16 categories), smoking status (never, past, or current smoker), fasting status at blood collection (yes, no), alcohol intake (g/day; continuous), Alternative Healthy Eating Index score (0–100; continuous), total energy intake (kcal/day; continuous), BMI (continuous), and intakes of fatty acids (linolenic acid, α-linolenic acid, eicosapentaenoic acid, docosahexaenoic acid, oleic acid, stearic acid, and *trans*- fatty acids).

[b] BMI is calculated as weight (kg)/height (m)^2.

[c] Permutation was performed within joint strata of sex (female, male), case status in the original study (yes, no), and BMI (<25, 25–29, ≥30).

[d] The *P* value for enrichment was based on 10,000 permutations to simulate the distribution of the enrichment OR under the global null hypothesis that no metabolite is associated with physical activity.

ACKNOWLEDGMENTS

Author affiliations: Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts (Ming Ding, Marta Guasch-Ferre, Frank Sacks, Walter C. Willett, Frank B. Hu); Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts (Oana A. Zeleznik, Marta Guasch-Ferre, Jessica Lasky-Su, A. Heather Eliassen, Walter C. Willett, Frank B. Hu); Division of Preventive Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts (Jie Hu, I-Min Lee, Kathryn M. Rexrode); Division of Women's Health, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts (Jie Hu, Kathryn M. Rexrode); Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts (I-Min Lee, A. Heather Eliassen, Walter C. Willett, Frank B. Hu, Peter Kraft); Department of Endocrinology, Diabetes and Metabolism, School of Medicine, Ohio State University, Columbus, Ohio (Rebecca D. Jackson); Department of Family Medicine and Public Health, School of Medicine, University of California, San Diego, La Jolla, California (Aladdin H. Shadyab); Department of Epidemiology and Environmental Health, School of Public Health and Health Professions, State University of New York at Buffalo, Buffalo, New York (Michael J. LaMonte); and Broad Institute of MIT and Harvard, Cambridge, Massachusetts (Clary Clish).

This study was supported by grants UM1 CA186107, P01 CA87969, R01 CA49449, R01 HL034594, R01 HL088521, UM1 CA176726, R01 CA67262, UM1 CA167552, and R01 HL35464 from the National Institutes of Health. Metabolomic analysis in the Women's Health Initiative (WHI) was funded by the National Heart, Lung, and Blood Institute through contract HHSN268201300008C. The WHI program is funded by the National Heart, Lung, and Blood Institute through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C.

Conflict of interest: none declared.

REFERENCES

1. Jones DP, Park Y, Ziegler TR. Nutritional metabolomics: progress in addressing complexity in diet and health. *Annu Rev Nutr*. 2012;32:183–202.
2. Sattelmair J, Pertman J, Ding EL, et al. Dose response between physical activity and risk of coronary heart disease: a meta-analysis. *Circulation*. 2011;124(7):789–795.
3. Stampfer MJ, Hu FB, Manson JE, et al. Primary prevention of coronary heart disease in women through diet and lifestyle. *N Engl J Med*. 2000;343(1):16–22.
4. Neufer PD, Bamman MM, Muoio DM, et al. Understanding the cellular and molecular mechanisms of physical activity-induced health benefits. *Cell Metab*. 2015;22(1):4–11.
5. Kujala UM, Mäkinen VP, Heinonen I, et al. Long-term leisure-time physical activity and serum metabolome. *Circulation*. 2013;127(3):340–348.
6. Fukai K, Harada S, Iida M, et al. Metabolic profiling of total physical activity and sedentary behavior in community-dwelling men. *PLoS One*. 2016;11(10):e0164877.
7. Xiao Q, Moore SC, Keadle SK, et al. Objectively measured physical activity and plasma metabolomics in the Shanghai Physical Activity Study. *Int J Epidemiol*. 2016;45(5):1433–1444.
8. Hankinson SE, Willett WC, Manson JE, et al. Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. *J Natl Cancer Inst*. 1998;90(17):1292–1299.
9. Chu NF, Spiegelman D, Yu J, et al. Plasma leptin concentrations and four-year weight gain among US men. *Int J Obes Relat Metab Disord*. 2001;25(3):346–353.
10. Mayers JR, Wu C, Clish CB, et al. Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. *Nat Med*. 2014;20(10):1193–1198.
11. Jiang X, Zeleznik OA, Lindstrom S, et al. Metabolites associated with the risk of incident venous thromboembolism: a metabolomic analysis. *J Am Heart Assoc*. 2018;7(22):e010317.
12. Huang T, Zeleznik OA, Poole EM, et al. Habitual sleep quality, plasma metabolites and risk of coronary heart disease in post-menopausal women [published online ahead of print October 26, 2018]. *Int J Epidemiol*. (doi:10.1093/ije/dyy234).
13. Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of Physical Activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc*. 1993;25(1):71–80.
14. Wolf AM, Hunter DJ, Colditz GA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. *Int J Epidemiol*. 1994;23(5):991–999.
15. Pettee Gabriel K, McClain JJ, Lee CD, et al. Evaluation of physical activity measures used in middle-aged women. *Med Sci Sports Exerc*. 2009;41(7):1403–1412.
16. Dudoit S, van der Laan MJ. *Multiple Testing Procedures With Applications to Genomics*. New York, NY: Springer Publishing Company; 2008.
17. Vanderweele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. *Am J Epidemiol*. 2010;172(12):1339–1348.
18. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods*. 2013;18(2):137–150.
19. Fielding CJ, Shore VG, Fielding PE. Lecithin: cholesterol acyltransferase: effects of substrate composition upon enzyme activity. *Biochim Biophys Acta*. 1972;270(4):513–518.
20. Kodama S, Tanaka S, Saito K, et al. Effect of aerobic exercise training on serum levels of high-density lipoprotein cholesterol: a meta-analysis. *Arch Intern Med*. 2007;167(10):999–1008.
21. Kokkinos PF, Holland JC, Narayan P, et al. Miles run per week and high-density lipoprotein cholesterol levels in healthy, middle-aged men. A dose-response relationship. *Arch Intern Med*. 1995;155(4):415–420.
22. Hu FB, Leitzmann MF, Stampfer MJ, et al. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch Intern Med*. 2001;161(12):1542–1548.
23. Li TY, Rana JS, Manson JE, et al. Obesity as compared with physical activity in predicting risk of coronary heart disease in women. *Circulation*. 2006;113(4):499–506.
24. Kraus WE, Houmard JA, Duscha BD, et al. Effects of the amount and intensity of exercise on plasma lipoproteins. *N Engl J Med*. 2002;347(19):1483–1492.

25. Slentz CA, Houmard JA, Johnson JL, et al. Inactivity, exercise training and detraining, and plasma lipoproteins. STRRIDE: a randomized, controlled study of exercise intensity and amount. *J Appl Physiol (1985)*. 2007;103(2):432–442.
26. Wang Y, Xu D. Effects of aerobic exercise on lipids and lipoproteins. *Lipids Health Dis*. 2017;16(1):Article 132.
27. Mann S, Beedie C, Jimenez A. Differential effects of aerobic exercise, resistance training and combined exercise modalities on cholesterol and the lipid profile: review, synthesis and recommendations. *Sports Med*. 2014;44(2):211–221.
28. Huttunen JK. Physical activity and plasma lipids and lipoproteins. *Ann Clin Res*. 1982;14(suppl 34):124–129.
29. Huffman KM, Slentz CA, Bateman LA, et al. Exercise-induced changes in metabolic intermediates, hormones, and inflammatory markers associated with improvements in insulin sensitivity. *Diabetes Care*. 2011;34(1):174–176.
30. Christinat N, Masoodi M. Comprehensive lipoprotein characterization using lipidomics analysis of human plasma. *J Proteome Res*. 2017;16(8):2947–2953.
31. Satija A, Bhupathiraju SN, Spiegelman D, et al. Healthful and unhealthful plant-based diets and the risk of coronary heart disease in US adults. *J Am Coll Cardiol*. 2017;70(4):411–422.
32. Veronese N, Li Y, Manson JE, et al. Combined associations of body weight and lifestyle factors with all cause and cause specific mortality in men and women: prospective cohort study. *BMJ*. 2016;355:i5855.
33. DeLany JP, Windhauser MM, Champagne CM, et al. Differential oxidation of individual dietary fatty acids in humans. *Am J Clin Nutr*. 2000;72(4):905–911.
34. Bergouignan A, Schoeller DA, Normand S, et al. Effect of physical inactivity on the oxidation of saturated and monounsaturated dietary fatty acids: results of a randomized trial. *PLoS Clin Trials*. 2006;1(5):e27.
35. Cooper JA, Watras AC, Shriver T, et al. Influence of dietary fatty acid composition and exercise on changes in fat oxidation from a high-fat diet. *J Appl Physiol (1985)*. 2010;109(4):1011–1018.
36. Rhee EP, Cheng S, Larson MG, et al. Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. *J Clin Invest*. 2011;121(4):1402–1411.
37. Stegemann C, Pechlaner R, Willeit P, et al. Lipidomics profiling and risk of cardiovascular disease in the prospective population-based Bruneck Study. *Circulation*. 2014;129(18):1821–1831.
38. Zheng Y, Hu FB, Ruiz-Canela M, et al. Metabolites of glutamate metabolism are associated with incident cardiovascular events in the PREDIMED PREvencion con DIeta MEDiterranea (PREDIMED) Trial. *J Am Heart Assoc*. 2016;5(9):e003755.
39. Menni C, Fauman E, Erte I, et al. Biomarkers for type 2 diabetes and impaired fasting glucose using a nontargeted metabolomics approach. *Diabetes*. 2013;62(12):4270–4276.
40. Zhang S, Zeng X, Ren M, et al. Novel metabolic and physiological functions of branched chain amino acids: a review. *J Anim Sci Biotechnol*. 2017;8:Article 10.
41. Opara EC, Petro A, Tevrizian A, et al. L-glutamine supplementation of a high fat diet reduces body weight and attenuates hyperglycemia and hyperinsulinemia in C57BL/6J mice. *J Nutr*. 1996;126(1):273–279.
42. Greenfield JR, Farooqi IS, Keogh JM, et al. Oral glutamine increases circulating glucagon-like peptide 1, glucagon, and insulin concentrations in lean, obese, and type 2 diabetic subjects. *Am J Clin Nutr*. 2009;89(1):106–113.
43. Iwashita S, Williams P, Jabbour K, et al. Impact of glutamine supplementation on glucose homeostasis during and after exercise. *J Appl Physiol (1985)*. 2005;99(5):1858–1865.
44. Batch BC, Hyland K, Svetkey LP. Branch chain amino acids: biomarkers of health and disease. *Curr Opin Clin Nutr Metab Care*. 2014;17(1):86–89.
45. Shimomura Y, Murakami T, Nakai N, et al. Exercise promotes BCAA catabolism: effects of BCAA supplementation on skeletal muscle during exercise. *J Nutr*. 2004;134(6 suppl):1583S–1587S.
46. van Hall G, MacLean DA, Saltin B, et al. Mechanisms of activation of muscle branched-chain alpha-keto acid dehydrogenase during exercise in man. *J Physiol*. 1996;494(3):899–905.
47. Harris RA, Joshi M, Jeoung NH, et al. Overview of the molecular and biochemical basis of branched-chain amino acid catabolism. *J Nutr*. 2005;135(6 suppl):1527S–1530S.
48. Krug S, Kastenmüller G, Stückler F, et al. The dynamic range of the human metabolome revealed by challenges. *FASEB J*. 2012;26(6):2607–2619.
49. Rheaume C, Arsenault BJ, Dumas MP, et al. Contributions of cardiorespiratory fitness and visceral adiposity to six-year changes in cardiometabolic risk markers in apparently healthy men and women. *J Clin Endocrinol Metab*. 2011;96(5):1462–1468.
50. Magnusson M, Lewis GD, Ericson U, et al. A diabetes-predictive amino acid score and future cardiovascular disease. *Eur Heart J*. 2013;34(26):1982–1989.
51. Shah SH, Bain JR, Muehlbauer MJ, et al. Association of a peripheral blood metabolic profile with coronary artery disease and risk of subsequent cardiovascular events. *Circ Cardiovasc Genet*. 2010;3(2):207–214.
52. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. *Nat Med*. 2011;17(4):448–453.
53. Guasch-Ferré M, Hruby A, Toledo E, et al. Metabolomics in prediabetes and diabetes: a systematic review and meta-analysis. *Diabetes Care*. 2016;39(5):833–846.
54. Newgard CB, An J, Bain JR, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab*. 2009;9(4):311–326.