



OPEN Metabolome-wide association study on physical activity

Maedeh Kojouri¹, Rui Pinto^{1,2,4}, Rima Mustafa^{1,2,4}, Jian Huang¹, He Gao¹, Paul Elliott^{1,2,3}, Ioanna Tzoulaki^{1,2,3} & Abbas Dehghan^{1,2,3}✉

The underlying mechanisms linking physical activity to better health are not fully understood. Here we examined the associations between physical activity and small circulatory molecules, the metabolome, to highlight relevant biological pathways. We examined plasma metabolites associated with self-reported physical activity among 2217 participants from the Airwave Health Monitoring Study. Metabolic profiling was conducted using the mass spectrometry-based Metabolon platform (LC/GC–MS), measuring 828 known metabolites. We replicated our findings in an independent subset of the study (n = 2971) using untargeted LC–MS. Mendelian randomisation was carried out to investigate potential causal associations between physical activity, body mass index, and metabolites. Higher vigorous physical activity was associated ($P < 0.05/828 = 6.03 \times 10^{-5}$) with circulatory levels of 28 metabolites adjusted for age, sex and body mass index. The association was inverse for glutamate and diacylglycerol lipids, and direct for 3–4-hydroxyphenyllactate, phenyl lactate (PLA), alpha-hydroxy isovalerate, tiglylcarnitine, alpha-hydroxyisocaproate, 2-hydroxy-3-methylvalerate, isobutyrylcarnitine, imidazole lactate, methionine sulfone, indole lactate, plasmalogen lipids, pristanate and fumarate. In the replication panel, we found 23 untargeted LC–MS features annotated to the identified metabolites, for which we found nominal associations with the same direction of effect for three features annotated to 1-(1-enyl-palmitoyl)-2-oleoyl-GPC (P-16:0/18:1), 1-(1-enyl-palmitoyl)-2-linoleoyl-GPC (P-16:0/18:2), 1-stearoyl-2-dihomo-linolenoyl-GPC (18:0/20:3n3 or 6). Using Mendelian randomisation, we showed a potential causal relationship between body mass index and three identified metabolites. Circulatory metabolites are associated with physical activity and may play a role in mediating its health effects.

Abbreviations

LC–MS	Liquid chromatography–mass spectrometry
UPLC–MS	Ultra-performance liquid chromatography–mass spectroscopy
BMI	Body mass index
EDTA	Ethylenediaminetetraacetic acid
IPAQ-SF	International Physical Activity Questionnaire-Short Form
MET	Metabolic equivalent
HDL	High-density lipoprotein
GWAS	Genome-wide association study
IVW	Inverse variance weighted
HMD	Human Metabolome Database

Physical activity is associated with notable benefits for health and disease prevention, increasing the quality of life and decreasing morbidity and mortality¹. Although the beneficial effects of physical activity are widely accepted, the underlying pathways are not well understood. Metabolomics is the systematic profiling and identification of metabolites (small circulating molecules of our body systems) in biofluids, tissues and cells². Metabolic profiling allows us to look for metabolite differences due to the genetic profile and changes in lifestyle and environmental factors^{3,4}. Recent advances in metabolomic analytic platforms have made it possible to study metabolomic profiles in large scale epidemiologic studies⁵. Metabolites that are associated with lifestyle factors such as physical activity could, on the one hand, be further investigated as objective biomarkers of physical activity and, on the other hand, could shed light on the pathways through which physical activity affect health.

¹Department of Epidemiology and Biostatistics, School of Public Health, Faculty of Medicine, Imperial College London, London W2 1PG, UK. ²UK Dementia Research Institute, Imperial College London, London W2 1PG, UK. ³MRC Centre for Environment and Health, School of Public Health, Imperial College London, London, UK. ⁴These authors contributed equally: Rui Pinto and Rima Mustafa. ✉email: a.dehghan@imperial.ac.uk

Previous studies have suggested several metabolites to be associated with physical activity. However, their findings are not yet replicated, and it is not clear whether the associations are causal or driven by confounding or reverse causation^{6–8}. We examined the association between different intensity levels of physical activity and plasma metabolite levels of 2217 participants from the Airwave Health Monitoring Study using the Metabolon platform. We replicated our results in samples from an independent non-overlapping set of 2971 individuals of the same study using untargeted ultra-performance liquid chromatography–mass spectroscopy (UPLC–MS). Further, we performed Mendelian randomization to examine the potential causal effect of physical activity and body mass index (BMI) on metabolites⁹.

Methods

Study population. The Airwave Health Monitoring Study (the Airwave Study) is a longitudinal cohort study on the UK police forces. Recruitment and baseline measurements were done from 2005 to 2014. Data from the voluntary health screen program include medical, biomedical occupational and lifestyle information¹⁰. All participants were provided with written informed consent. The Airwave Study is approved by the National Health Service Multi-Site Research Ethics Committee (MREC/13/NW/0588). All methods were performed in accordance with the relevant guidelines and regulations.

Metabolic profiling. *Discovery panel.* Metabolic profiling was conducted using untargeted mass spectrometry (LC–MS) on ethylenediaminetetraacetic acid (EDTA) plasma samples by the metabolomics data supplier Metabolon, Inc. (Durham, NC, USA). The platform consolidated two separate ultra-high-performance liquid chromatography/tandem mass spectrometry (UHPLC/MS/MS2) injections and one gas-chromatography coupled with mass spectrometry injection (GC/MS) per sample. The UPLC injections have been optimised specifically for basic and acidic species¹¹. A total of 1148 metabolites were measured, comprising 828 annotated and 320 unknowns. The 828 annotated metabolites were grouped into amino acids, lipids, carbohydrates, cofactors and vitamins, energy, nucleotides, peptides and xenobiotics. The unknown metabolites were excluded from further analysis. As many metabolites had highly skewed distributions, we used natural-log transformation, and further standardised the values for the analysis. Metabolomics assays were done on 2250 individuals in the Airwave Study. Physical activity data were available for 2217 individuals.

Replication panel. To replicate the findings, we extracted the same putative metabolites from an untargeted LC–MS analysis of a sub-sample of the Airwave Study (2971) not overlapping with the discovery set. These samples were analysed at the Imperial National Phenome Centre (NPC) in the United Kingdom using three complementary assays: hydrophilic interaction liquid chromatography (HILIC) in the positive mode for small, polar molecules¹², and lipid-targeted reverse-phase chromatography for fatty acids, triglycerides and phospholipids ionised in both positive and negative modes¹³. The *m/z* values of the Metabolon sample were compared to the ones of the untargeted datasets, and all putative metabolites within 2.5e–5 Da were considered a match. Their intensities were extracted and z-scored for association analyses.

Physical activity assessment. Physical activity (PA) was assessed by a self-reported questionnaire using the International Physical Activity Questionnaire-Short Form (IPAQ-SF)¹⁴. This questionnaire collects data on walking, moderate activity, and vigorous activity. Using data on the duration and frequency of activity per week from the questionnaire, we calculated the metabolic equivalent (MET) minute per week for three physical activity levels (walking, moderate activity and vigorous activity). Applying the compendium by Ainsworth et al.¹⁵, a mean MET score was determined for each sort of activity: Walking = 3.3 METs, Moderate PA = 4.0 METs and Vigorous PA = 8.0 METs. One MET is defined as the energy consumed by a 70-kg individual in a sitting position. Total physical activity was calculated as the sum of the METs of walking, moderate activity and vigorous activity.

Covariates. Information on age, sex, education level, smoking status, and alcohol consumption was collected using a questionnaire. Other clinical measurements including height, weight, systolic and diastolic blood pressure, hip and waist girth were recorded during health screening sessions. BMI was calculated by dividing weight in kilogram by square height in meters. Blood pressure was measured using Omron HEM 705-CP digital BP. Details of each clinical measurement are described elsewhere¹⁰.

Statistical analysis. We fitted linear regression models to investigate the association between physical activity and each metabolite, adjusted for age and sex (model 1). In model 2, we further adjusted for BMI. In model 3, further adjustment was performed for smoking status (current/former/never), systolic and diastolic blood pressure (continuous), alcohol consumption (yes/no), high-density lipoprotein (HDL) and total cholesterol (continuous). Bonferroni correction was used to adjust for multiple testing ($\alpha = 0.05/\text{number of tested metabolites} = 0.05/828 = 6.03 \times 10^{-5}$). All analyses were performed using the R statistical language package (Version 3.6.1).

Pathway enrichment analysis. We performed pathway enrichment analysis to evaluate whether metabolites related to physical activity are enriched in specific metabolite pathways using MetaboAnalyst, a web-based metabolomics tool for pathway analysis and visualisation¹⁶.

Mendelian randomisation analysis. Mendelian randomisation uses genetic variants as instrumental variables to investigate causality. Since genetic variants are randomly inherited, they are not associated with

confounding factors⁹. We applied two-sample Mendelian randomisation using the TwoSampleMR package in R¹⁷. We used 9 independent genetic instruments for physical activity from a genome-wide association study (GWAS) on physical activity¹⁸. The association of the genetic instruments with metabolites were extracted using GWAS on metabolites levels from 1942 participants of the Airwave Study. GWAS was performed using the high-dimensional association analyses (HASE) framework with adjustment for age, sex, and the first ten principal components¹⁹.

We selected 503 independent SNPs as genetic instruments for BMI using data from a recent GWAS on BMI, based on p -value $< 5 \times 10^{-8}$ and LD cut-off of $R^2 < 0.001$ ²⁰. Further, we applied Mendelian randomisation analysis of BMI on identified metabolites using the Airwave dataset and another separate set of analyses using a dataset from Shin et al.²¹ to investigate the possible causal relation between BMI and identified metabolites.

We used the inverse variance weighted (IVW) method to estimate the causal effect. We further applied robust methods including weighted median and MR-Egger to rule out potential pleiotropic effects^{22–24}. Bonferroni correction was used for multiple testing ($\alpha = 0.05/\text{number of tested metabolites} = 0.05/36 = 0.001$).

Ethics approval and consent to participate. The Airwave Study is approved by the National Health Service Multi-Site Research Ethics Committee (MREC/13/NW/0588).

Results

The baseline characteristics of the participants of the Airwave Study are shown in Table 1.

Vigorous physical activity was associated with 36 plasma metabolites in model 1, adjusted for age and sex (Fig. 1, Supplementary Table 1). The association was direct for 24 metabolites (amino acids, plasmalogen lipids, fatty acid lipid, and tricarboxylic acid cycle metabolite) and inverse for 12 metabolites [diacylglycerol lipids, monoacylglycerol lipid, phospholipid metabolism product, and amino acid (glutamate)]. We examined the correlation between these metabolites using Pearson correlation, with many of the metabolites associated with vigorous physical activity correlated with each other (Fig. 2).

With further adjustment for BMI in model 2, the association remained meaningful for 28 metabolites (Fig. 1, Supplementary Table 2). The monoacylglycerol lipid 1-linolenoylglycerol and the phospholipid metabolism product 1-stearoyl-2-dihomo-linolenoyl-GPC were no longer significant after this adjustment. In model 3, the effect estimate did not change after adjustment.

Moderate physical activity was inversely associated with one medium-chain length fatty acid (heptanoate) ($\beta = -0.03$, p -value = 2.5×10^{-5}), which was minimally affected by adjustment for BMI in model 2 ($\beta = -0.03$, p -value = 2.28×10^{-5}) and was not affected by adjustment in model 3.

Total physical activity was positively associated with the amino acid 3-(4-hydroxyphenyl) lactate ($\beta = 0.002$, p -value = 3.62×10^{-6}) and inversely associated with the medium-chain length fatty acid heptanoate ($\beta = -0.01$, p -value = 1.78×10^{-5}) (Supplementary Table 1). The associations remained unchanged after further adjustment in model 2 and model 3 (Supplementary Table 2).

Characteristic	Mean (SD)/number (%)
Age, years	41.2 (8.3)
Body mass index, kg/m ²	26.8 (4.0)
Systolic blood pressure, mmHg	125.6 (13.9)
Diastolic blood pressure, mmHg	77.2 (9.3)
Waist girth, cm	89.1 (11.5)
Waist to hip ratio	0.86 (0.07)
Total cholesterol, mmol/l	5.2 (0.95)
High-density lipoprotein (HDL), mmol/l	1.43 (0.36)
Walking, MET-hours/week	6.6 (1.6)
Moderate physical activity, MET-hours/week	3.9 (3.2)
Vigorous physical activity, MET-hours/week	4.8 (3.5)
Total physical activity MET-hours/week	9.8 (10.5)
Sedentary time min/day	2087 (1098.1)
White ethnicity	2003 (90%)
Education A levels/higher or equivalent (NVQ3)	687 (30%)
GSCE/O-level/CSE	608 (27%)
Postgraduate qualifications	146 (6%)
Bachelor's degree or equivalent (NVQ4)	547 (24%)
Left school before taking O levels/GCSEs	70 (3%)
Vocational qualifications (NVQ1 + 2)	159 (7%)
Alcohol consumption positive	1996 (90%)

Table 1. Baseline characteristics of participants in the Airwave Study.

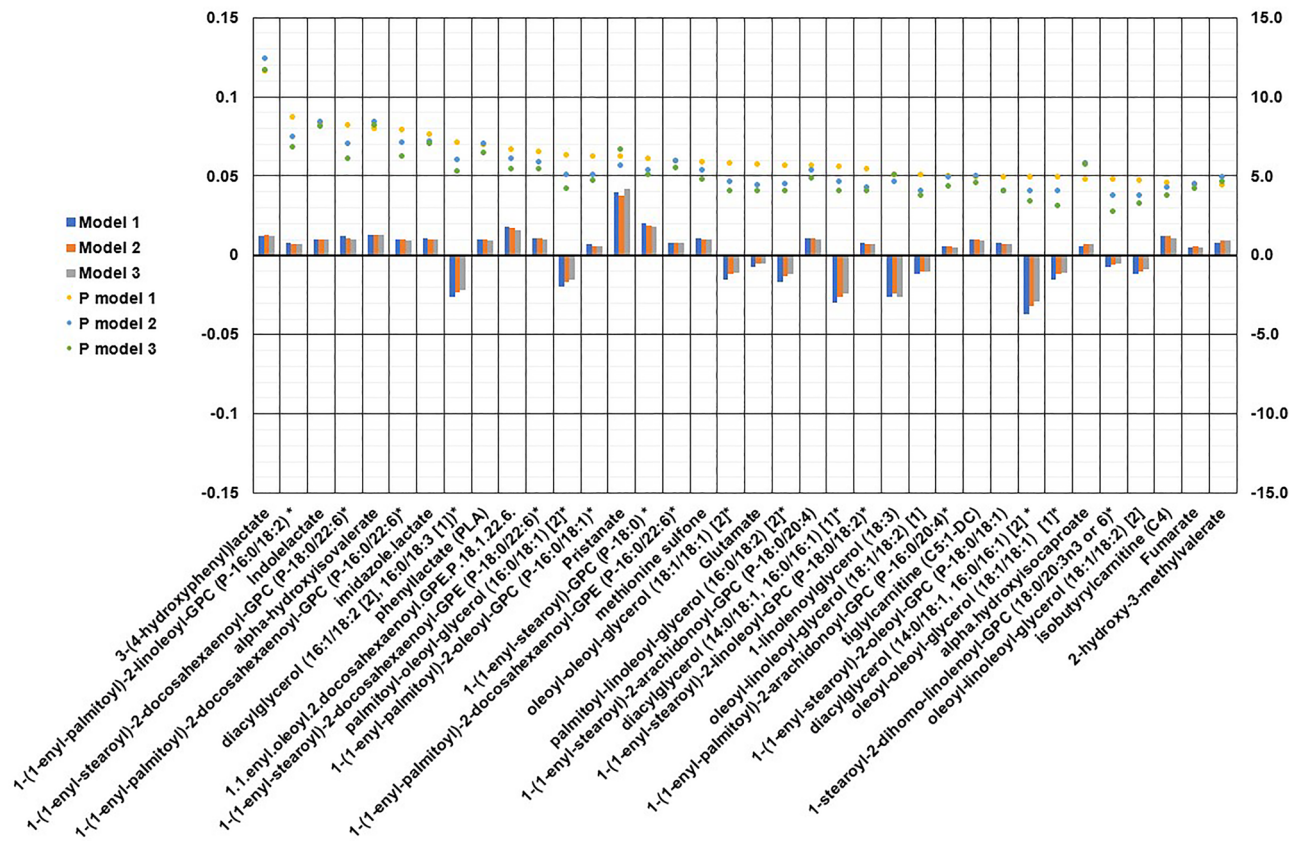


Figure 1. The association of metabolites measured by Metabolon platform with vigorous physical activity in Airwave Study. Asterisk: Model 1: age and sex-adjusted. Double asterisk: Model 2: model 1 with further adjustment for BMI. Triple asterisk: Model 3: model 1 with further adjustment for all covariates [smoking status (current/former/never), systolic and diastolic blood pressure (continuous), alcohol consumption (yes/no), high-density lipoprotein (HDL) and total cholesterol (continuous)]. The bars represent the beta coefficient and the dots the p-value of the each of the three models.

Replication. We searched for the identified metabolites in an untargeted UPLC-MS dataset of 2971 participants in the Airwave Study and found 82 data points that represented 23 identified metabolites. Three metabolites including 1-(1-enyl-palmitoyl)-2-oleoyl-GPC, 1-(1-enyl-palmitoyl)-2-linoleoyl-GPC, and 1-stearoyl-2-dihomo-linolenoyl-GPC were associated with vigorous physical activity with the same direction (p -value < 0.05) (Table 2).

Pathway enrichment analysis. We conducted an enrichment analysis to examine whether the associated metabolites are clustered in a certain pathway or metabolite set. The analysis was performed with the MetaboAnalyst online tool using the Human Metabolome Database (HMDB) metabolite identification codes¹⁶. Several metabolite sets from pathways including arginine and proline metabolism, mitochondrial electron transport chain, and oxidation of branched-chain fatty acids, were enriched with the associated metabolites (Fig. 3).

Mendelian randomisation. We conducted Mendelian randomisation analysis to investigate the potential causal role of physical activity on each of the 36 associated metabolites. Using the IVW method, we did not find a causal association between physical activity and any of the identified metabolites after Bonferroni adjustment (Supplementary Table 3).

We further investigated the potential causal role of obesity on these 36 metabolites. The association of genetic instruments for BMI and 5 of 36 metabolites were taken from the previous GWAS on metabolites in Twins UK and KORA study²¹. IVW estimates were nominally significant for two metabolites, annotated as 3-(4-hydroxyphenyl)lactate and alpha-hydroxy isovalerate. For the remaining 31 metabolites, genetic associations from the Airwave Study were used, resulting in a nominally significant IVW estimate for 1-(1-enyl-stearoyl)-2-docosahexaenoyl-GPC (P-18:0/22:6) * (Supplementary Table 4).

Discussion

In this study, we examined the association of self-reported physical activity with 1148 metabolic biomarkers among 2217 individuals from the Airwave study and replicated the findings in 2971 independent study participants. Vigorous physical activity was found to be associated with 36 blood metabolites (mainly amino acid, plasmalogens and diacylglycerols), of which 28 associations survived adjustment for BMI. Mendelian randomisation

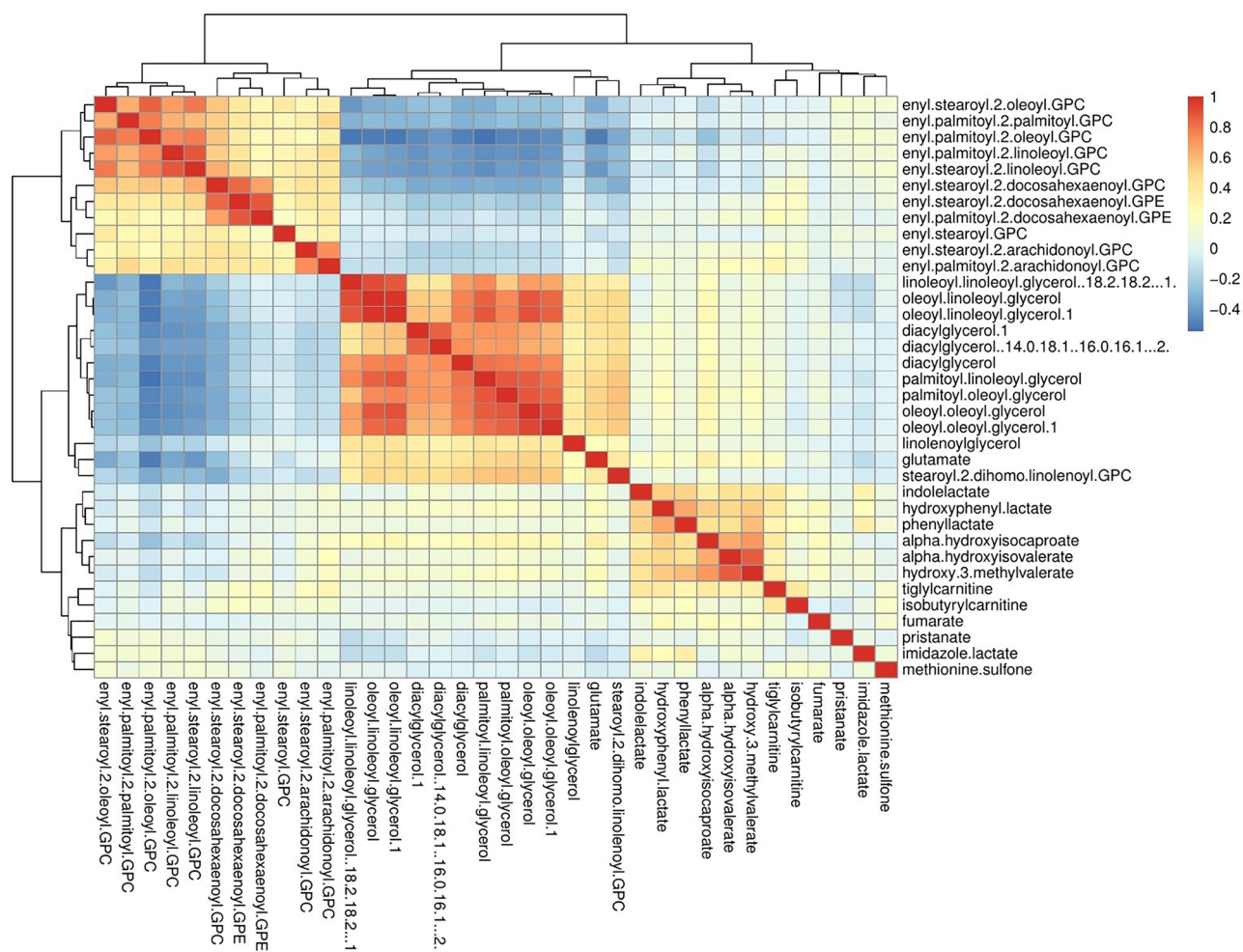


Figure 2. Heat map showing the correlation between the annotated metabolites significantly associated with vigorous physical activity energy expenditure. Colour indicates the Pearson correlation coefficient.

Biochemical	Feature	Beta (replication)	P-value (replication)
1-(1-enyl-palmitoyl)-2-oleoyl-GPC (P-16:0/18:1)*	HPOS_744.5883_3.9884	0.044	1.8×10^{-5}
1-(1-enyl-palmitoyl)-2-linoleoyl-GPC (P-16:0/18:2)*	HPOS_742.5760_3.9933	0.045	4.9×10^{-5}
1-stearoyl-2-dihomo-linolenoyl-GPC (18:0/20:3n3 or 6)*	HPOS_812.6154_3.9711	-0.041	0.0001
1-stearoyl-2-dihomo-linolenoyl-GPC (18:0/20:3n3 or 6)*	HPOS_812.6049_4.1231	-0.035	0.0003
1-stearoyl-2-dihomo-linolenoyl-GPC (18:0/20:3n3 or 6)*	HPOS_812.6046_4.2519	-0.031	0.0012
1-stearoyl-2-dihomo-linolenoyl-GPC (18:0/20:3n3 or 6)*	HPOS_812.6164_4.1712	-0.026	0.0077

Table 2. Association of serum metabolomic features with *m/z* values close to the identified plasma metabolites associated with vigorous physical activity. HPOS means the features were found in the HILIC Positive dataset. Asterisk indicates compounds that have not been officially confirmed based on a standard, but Metabolon Inc. is confident in its identity.

did not show evidence for a causal effect of physical activity on metabolites, however, BMI is likely to affect three of the identified metabolites.

The association of self-reported physical activity and metabolites was previously studied by Ding et al., in the Nurses' Health Study and the Health Professional Follow-up Study in the US⁸. Ding et al. examined 337 metabolites measured by liquid chromatography-mass spectrometry (LC-MS) in plasma and reported 20 metabolites associated with physical activity. Of those metabolites, we replicated the association with glutamate in our study. Moreover, Xiao et al. investigated the association of plasma metabolites measured by the Metabolon platform with physical activity in a prospective cohort study of 277 Chinese adults⁷ and identified 11 metabolites associated with physical activity, among which alpha-hydroxy isovalerate and glutamate overlapped with our findings. Although the association with glutamate was found in the same direction, the direction of the effect for

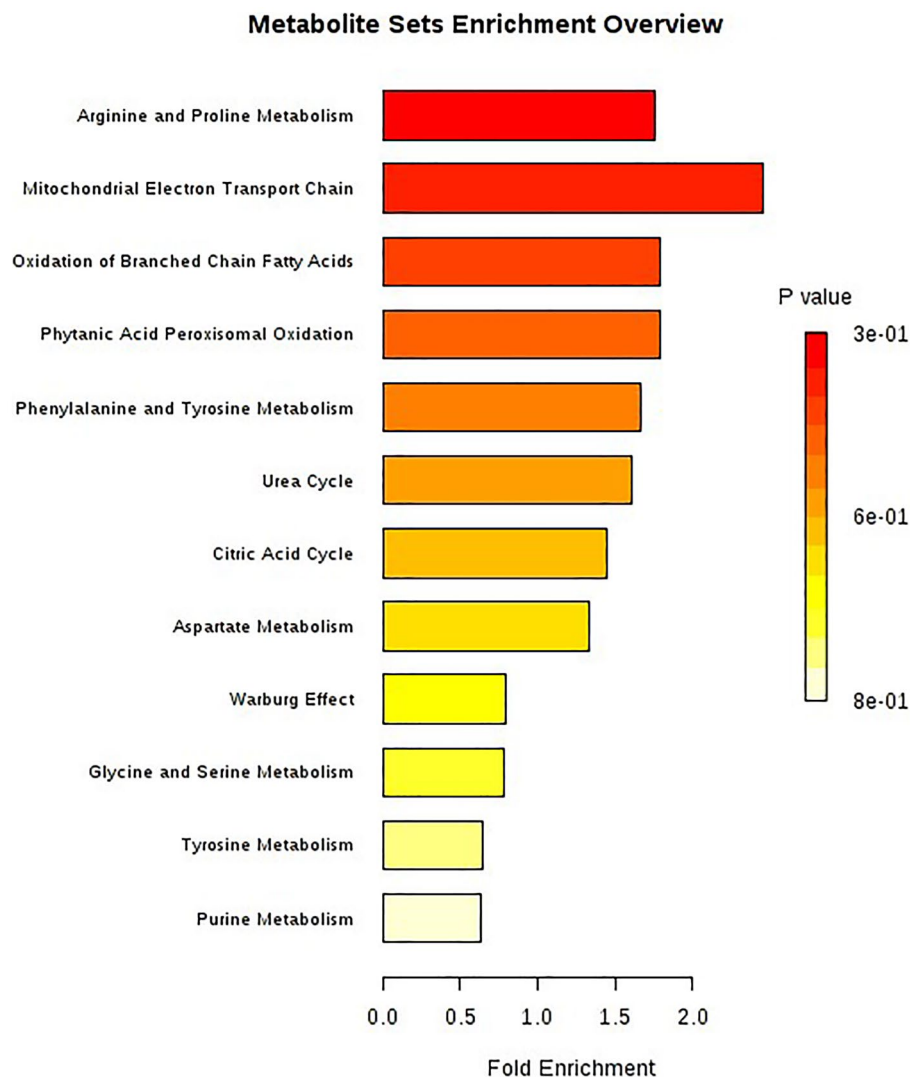


Figure 3. Metabolite set enrichment pathway showing the significantly associated pathways related to vigorous physical activity.

alpha-hydroxy isovalerate was different. They also found (borderline) significance for tiglylcarnitine and both 2-hydroxy-3-methylvalerate and 4-methyl-2-oxopentanoate (branched-chain fatty acids biochemically very similar to our finding alpha-hydroxyisocaproate, the latter its oxidised—ketone—version). The study by Xiao et al. had some differences from ours. First, Xiao et al. used an accelerometer to measure physical activity while we used a questionnaire to calculate physical activity levels. Nevertheless, the questionnaire is considered a valid proxy to calculate MET. Second, Xiao et al. considered overall physical activity levels, while in our study the found metabolites were associated only with vigorous-intensity activity. The difference in the measured exercise intensity may explain the inconsistency between our findings and Xiao et al. as it has been previously shown by Devlin et al., that intense physical activity can increase amino acid synthesis²⁵.

Another cohort metabolic profiling study on 1193 Japanese individuals by Fukai et al. examined associations of physical activity with polar metabolites and found that a higher level of physical activity is associated with higher concentration levels of piperolate and lower concentration levels of amino acids⁶. Fukai also reported that a lower glutamate level is associated with physical activity. Glutamate was measured using capillary electrophoresis time-of-flight mass spectrometry (CE-TOFMS).

We found that the association of vigorous physical activity and glutamate was independent of BMI. Previous studies have shown that glutamate is associated with insulin resistance²⁶, cardiovascular disease²⁷, liver disease²⁸, cancer and HIV infectious^{29,30}. Increased glutamate level has also been reported in the pathogenesis of immunosuppression³¹. Hyperglutamatemia has also been found in amyotrophic lateral sclerosis³², Parkinson's disease³³, epilepsy³⁴, autism³⁵, migraine³⁶, and depression³⁷. Although elevated levels of glutamate in different diseases have various causes, inflammation due to oxidative stress is a common pathologic pathway in all instances. We also found an inverse association between vigorous physical activity and diacylglycerols, which

were previously shown to be elevated in the muscle of obese individuals and type 2 diabetes patients, indicating substantial evidence in the development of insulin resistance^{38,39}.

We found 36 metabolites associated with vigorous activity, while only one metabolite was associated with moderate activity. The difference might indicate the gap between metabolomic changes and the health effect of moderate versus vigorous physical activity. It has been shown previously that vigorous-intensity physical activity has a greater effect on preventing disease and provides extra benefit compared to moderate-intensity^{40–42}. However, obesity could be a confounding factor. Overweight or obese individuals are less likely to take part in intense physical activity and therefore the association could have been confounded. Adiposity affects a wide range of metabolites and adipose tissue is known to release a wide range of cytokines that affects the metabolome. To investigate the potential confounding effect of obesity, we adjusted the associations for BMI and observed that many associations became non-significant or the effect estimates were diminished.

Insulin resistance plays a crucial role in pathogenesis of type 2 diabetes. Our study highlighted two metabolites with inverse association with vigorous physical activity, both belonging to the diacylglycerol family. Previous studies have shown that diacylglycerols could induce insulin resistance due to an imbalance between fatty acid delivery and intracellular fatty acid oxidation and storage⁴³. It is widely known that insulin resistance plays a crucial role in the pathogenesis of type 2 diabetes⁴⁴, and our current findings might indicate evidence for the positive effect of vigorous physical activity on reducing type 2 diabetes. Fumarate is a final product of the tricarboxylic acid energy cycle⁴⁵. Importantly our study identified an elevated plasma concentration of fumarate in highly active individuals, indicating a greater potential for releasing energy during catabolism.

About 4% of all metabolites measured in this study were associated with vigorous physical activity. The results of pathway analysis indicated that the associated metabolites are clustered in several pathways. The arginine and proline pathways include the biosynthesis and metabolism of several amino acids, including arginine, ornithine, proline, citrulline and glutamate, as well as the mitochondrial electron transport chain pathway used for the aerobic production of energy by mitochondria. This pathway is highly efficient in releasing energy compared to anaerobic glycolysis and therefore can have a meaningful relationship with vigorous long-duration physical activity. Other enriched pathways include oxidation of branched-chain fatty acid and phytanic acid peroxisomal oxidation to increase aerobic capacity during exercise⁴⁶.

This study has several strengths. We used the Metabolon platform which covers a wide range of metabolites including 391 lipid metabolites which are important in relation to physical activity. We replicated our findings in a sizable replication panel. Moreover, the Mendelian randomisation approach was used to assess the possible causal relationship between physical activity and metabolites. We also applied Mendelian randomisation to assess the effect of obesity, a major potential confounder, on the studied metabolites. Finally, we conducted a pathway analysis which helped us to highlight potential underlying pathways. Certain limitations should also be mentioned. Physical activity in our study was measured using a self-reported questionnaire. While this method is considered a valid approach to calculate MET, objective methods such as accelerometer might be more accurate⁴⁷. We were not able to identify genetic instruments for all metabolites and the instruments for some of the metabolites might be weak, therefore, we cannot rule out a possibility of an undetected causal effect of physical activity on the metabolites.

Conclusion

In this study, we identified several metabolites associated with vigorous but not moderate physical activity, a finding that may have implications for the recommended intensity level of physical activity aiming to improve health. Larger studies and better genetic instruments are needed to investigate a wider range of metabolites. Our findings may also have implications as biomarkers for different exercise levels.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 13 June 2022; Accepted: 14 December 2022

Published online: 09 February 2023

References

1. Erik, D. X. *et al.* Physical activity and performance impact long-term quality of life in older adults at risk for major mobility disability. *Am. J. Prev. Med. (Elsevier Inc.)* **56**(1), 141–146 (2019).
2. Patti, G. J., Yanes, O. & Siuzdak, G. Innovation: Metabolomics: The apogee of the omics trilogy. *Nat. Publ. Gr. (Nature Publishing Group)* **13**(4), 263–269 (2012).
3. Jacob, M., Lopata, A. L., Dasouki, M. & Abdel Rahman, A. M. Metabolomics toward personalized medicine. *Mass Spectrom. Rev.* **38**(3), 221–238 (2019).
4. Johnson, C. H., Ivanisevic, J. & Siuzdak, G. Metabolomics: Beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell Biol.* **17**, 451–459 (2016).
5. Brennan, A. M. *et al.* Plasma metabolite profiles in response to chronic exercise. *Med. Sci. Sports Exerc.* **50**(7), 1480–1486 (2018).
6. Fukai, K. *et al.* Metabolic profiling of total physical activity and sedentary behavior in community-dwelling men. *PLoS ONE* **11**(10), 1–14 (2016).
7. Xiao, Q. *et al.* Objectively measured physical activity and plasma metabolomics in the Shanghai Physical Activity Study. *Int. J. Epidemiol.* **45**(5), 1433–1444 (2016).
8. Ding, M. *et al.* Metabolome-wide association study of the relationship between habitual physical activity and plasma metabolite levels. *Am. J. Epidemiol.* **188**(11), 1932–1943 (2019).
9. Smith, G. D. & Hemani, G. Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.* **23**(1), 89–98 (2014).

10. Elliott, P. *et al.* The Airwave Health Monitoring Study of police officers and staff in Great Britain : Rationale, design and methods. *Environ Res. (Elsevier)* **134**, 280–285 (2014).
11. Evans, A. M., Dehaven, C. D., Barrett, T., Mitchell, M. & Milgram, E. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal. Chem.* **81**(16), 6656–6667 (2009).
12. Lewis, M. R. *et al.* Development and application of ultra-performance liquid chromatography-TOF MS for precision large scale urinary metabolic phenotyping. *Anal. Chem.* **88**(18), 9004–9013 (2016).
13. Sarafian, M. H. *et al.* Objective set of criteria for optimization of sample preparation procedures for ultra-high throughput untargeted blood plasma lipid profiling by ultra performance liquid chromatography–mass spectrometry. *Anal. Chem.* **86**(12), 5766–5774 (2014).
14. Bauman, A. E. *et al.* International Physical Activity Questionnaire : 12-country reliability and validity. *Med. Sci. Sport Exerc.* **35**(8), 1381–1395 (2000).
15. Ainsworth, B. E. *et al.* Compendium of physical activities: An update of activity codes and MET intensities. *Med. Sci. Sport Exerc.* **32**(12), 498–504 (2000).
16. Xia, J., Sinelnikov, I. V., Han, B. & Wishart, D. S. MetaboAnalyst 3.0-making metabolomics more meaningful. *Nucleic Acids Res.* **43**(W1), W251–W257 (2015).
17. Minelli, C. *et al.* The use of two-sample methods for Mendelian randomization analyses on single large datasets. *Int. J. Epidemiol.* **50**(5), 1651–1659 (2021).
18. Klimentidis, Y. C. *et al.* Genome-wide association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple variants including CADM2 and APOE. *Int. J. Obes.* **42**(6), 1161–1176 (2018).
19. Roshchupkin, G. V., Adams, H. H. H., Vernooij, M. W., Hofman, A. & Van Duijn, C. M. OPEN HASE : Framework for efficient high-dimensional association analyses. *Nat. Publ. Gr. (Nature Publishing Group)* **6**, 1–8 (2016).
20. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body mass index in ~ 700 000 individuals of European ancestry. *Hum. Mol. Genet.* **27**(20), 3641–3649 (2018).
21. Shin, S. *et al.* An atlas of genetic influences on human blood metabolites. *Nat. Genet.* **46**(6), 543–550 (2014).
22. Bowden, J., Smith, G. D., Haycock, P. C. & Burgess, S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator genetic epidemiology. *Genet. Epidemiol.* **40**(4), 304–314 (2016).
23. Bowden, J., Greco, D., Minelli, C., Smith, G. D. & Thompson, J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat. Med.* **36**(11), 1783–1802 (2017).
24. Bowden, J., Smith, G. D. & Burgess, S. Mendelian randomization methodology Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**(2), 512–525 (2015).
25. Devlin, J. T., Brodsky, I., Scrimgeour, A., Fuller, S. & Bier, D. M. Amino acid metabolism after intense exercise. *Am. J. Physiol. Endocrinol. Metab.* **258**(2), 249–255 (1990).
26. Greenfield, J. R. *et al.* Oral glutamine increases circulating glucagon-like peptide 1, glucagon, and insulin concentrations in lean, obese, and type 2 diabetic subjects 1–4. *Am. J. Clin. Nutr.* **1**, 106–113 (2009).
27. Zheng, Y. *et al.* Metabolites of glutamate metabolism are associated with incident cardiovascular events in the PREDIMED PRE-vencción con Dieta MEDiterránea (PREDIMED) trial. *J. Am. Heart Assoc.* **5**(9), 003755 (2016).
28. Morgan, M. Y., Marshall, A. W., Milsom, J. P. & Sherlock, S. Plasma amino-acid patterns in liver disease. *Gut* **23**(5), 362–370 (1982).
29. Eck, H. P., Drings, P. & Dröge, W. Plasma glutamate levels, lymphocyte reactivity and death rate in patients with bronchial carcinoma. *J. Cancer Res. Clin. Oncol.* **115**(6), 571–574 (1989).
30. Dröge, W., Eck, H. P., Betzler, M. & Näher, H. Elevated plasma glutamate levels in colorectal carcinoma patients and in patients with acquired immunodeficiency syndrome (AIDS). *Immunobiology* **174**(4–5), 473–479 (1987).
31. Dröge, W. *et al.* Plasma glutamate concentration and lymphocyte activity. *J. Cancer Res. Clin. Oncol.* **114**(2), 124–128 (1988).
32. Andreadou, E. *et al.* Plasma glutamate and glycine levels in patients with amyotrophic lateral sclerosis: The effect of riluzole treatment. *Clin. Neurol. Neurosurg.* **110**(3), 222–226 (2008).
33. Iwasaki, Y., Ikeda, K., Shiojima, T. & Kinoshita, M. Increased plasma concentrations of aspartate, glutamate and glycine in Parkinson's disease. *Neurosci. Lett.* **145**(2), 175–177 (1992).
34. Paraskevas, G. P. *et al.* Add-on lamotrigine treatment and plasma glutamate levels in epilepsy: Relation to treatment response. *Epilepsy Res.* **70**(2–3), 184–189 (2006).
35. Aldred, S., Moore, K. M., Fitzgerald, M. & Waring, R. H. Plasma amino acid levels in children with autism and their families. *J. Autism Dev. Disord.* **33**(1), 93–97 (2003).
36. Alam, Z., Coombes, N. H., Waring, R., Williams, A. C. & Steventon, G. B. Plasma levels of neuroexcitatory amino acids in patients with migraine or tension headache. *J. Neurol. Sci.* **156**(1), 102–106 (1998).
37. Mitani, H. *et al.* Correlation between plasma levels of glutamate, alanine and serine with severity of depression. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **30**(6), 1155–1158 (2006).
38. Bergman, B. C., Hunerdosse, D. M., Kerege, A., Playdon, M. C. & Perreault, L. Localisation and composition of skeletal muscle diacylglycerol predicts insulin resistance in humans. *Diabetologia* **55**(4), 1140–1150 (2013).
39. Dubé, J. J. *et al.* Effects of weight loss and exercise on insulin resistance, and intramyocellular triacylglycerol, diacylglycerol and ceramide. *Diabetologia* **54**(5), 1147–1156 (2013).
40. Pavey, T. G., Peeters, G., Bauman, A. E. & Brown, W. J. Does vigorous physical activity provide additional benefits beyond those of moderate?. *Med. Sci. Sports Exerc.* **45**(10), 1948–1955 (2013).
41. Swain, D. P. & Franklin, B. A. Comparison of cardioprotective benefits of vigorous versus moderate intensity aerobic exercise. *Am. J. Cardiol.* **97**(1), 141–147 (2006).
42. Rankin, A. J., Rankin, A. C., Macintyre, P. & Hillis, W. S. Walk or run? Is high-intensity exercise more effective than moderate-intensity exercise at reducing cardiovascular risk?. *Scott Med. J.* **57**(2), 99–102 (2012).
43. Erion, D. M. & Shulman, G. I. Diacylglycerol-mediated insulin resistance. *Nat. Med.* **16**(4), 400–402 (2013).
44. Henriksen, E. J., Diamond-Stanic, M. K. & Marchionne, E. M. Oxidative stress and the etiology of insulin resistance and type 2 diabetes. *Free Radic. Biol. Med. (Internet, Elsevier Inc.)* **51**(5), 993–999. (2011).
45. Roa Engel, C. A., Straathof, A. J. J., Zijlmans, T. W., Van, G. W. M. & Van Der, W. L. A. M. Fumaric acid production by fermentation. *Appl. Microbiol. Biotechnol.* **78**(3), 379–389 (2008).
46. Kainulainen, H., Hulmi, J. J. & Kujala, U. M. Potential role of branched-chain amino acid catabolism in regulating fat oxidation. *Exerc. Sport Sci. Rev.* **41**(4), 194–200 (2013).
47. Kozey, S. L., Lyden, K., Howe, C. A., Staudenmayer, J. W. & Freedson, P. S. Accelerometer output and MET values of common physical activities. *Med. Sci. Sports Exerc.* **42**(9), 1776–1784 (2010).

Author contributions

A.D. and M.K. had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. M.K. analysed the data and wrote the first draft. R.M., R.P. and J.H. contributed to data analysis and interpretation of the results. H.G., P.E. and I.T. contributed to the critical revision of the manuscript. All authors read and approved the final submitted version of the manuscript.

Funding

This work was supported by Medical Research Council (grant number MR/R023484/1), with additional support from the National Institute for Health Research (NIHR) Imperial College Biomedical Research Centre (BRC) in collaboration with Imperial College NHS Healthcare Trust. We thank all participants in the Airwave Health Monitoring Study. The views expressed are those of the authors and not necessarily those of the sponsors. PE acknowledges support from the Medical Research Council and Public Health England (MR/L01341X/1; MR/S019669/1) for the MRC Centre for Environment and Health; and the NIHR Health Protection Research Unit in Chemical and Radiation Threats and Hazards at Imperial College London (NIHR-200922). This work used computing resources of the UK MEDical BIOinformatics partnership (UK MED-BIO supported by the Medical Research Council (MR/L01632X/1). PE is supported by the UK Dementia Research Institute which receives its funding from UK DRI Ltd funded by the UK Medical Research Council, Alzheimer's Society and Alzheimer's Research UK (MC_PC_17114). PE is associate director of the Health Data Research UK London funded by a consortium led by the UK Medical Research Council. RM is sponsored by the President's PhD Scholarships from Imperial College London. AD is funded by the Wellcome Trust seed award (206046/Z/17/Z).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-26377-7>.

Correspondence and requests for materials should be addressed to A.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023