medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Genotype-specific effects of physical activities on self-harm behavior in depressed patients: findings from the UK Biobank

Jaegyun Jung, KMD[1,3]; Sangyeon Lee, MS[2,3]; Jeong Ho Lee, MD, PhD[1]; Doheon Lee, PhD[2,3]

[1] Graduate School of Medical Science and Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea

[2] Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea

[3] Bio-Synergy Research Center, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea

Word count: 2,958 words

Corresponding author: Doheon Lee, PhD

Affiliation: Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea

Postal/mail address: #505, Chung Moon-Soul Building (E16), KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea

E-mail address: dhlee@kaist.ac.kr

Tel: +82-42-350-8681 (Office), +82-42-350-4316 (Direct)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

Background

Physical activities are widely implemented for non-pharmacological intervention in depression. However, there is little evidence supporting their genotype-specific effectiveness in reducing the risk of intentional self-harm in midlife and late-life depression patients.

Aims

To assess whether the specific type and amount of physical activity is associated with the risk of intentional self-harm behavior among patients with depression phenotype.

Method

We developed the bidirectional analytic model to investigate the genotype-specific effectiveness (BAIGE) on UK Biobank. After the genetic stratification on the depression phenotype cohort using hierarchical clustering, multivariable logistic regression models and Cox proportional hazard models were built to investigate the associations between physical activity and the risk of intentional self-harm behavior.

Results

A total of 28,923 subjects with depression phenotype were included in the study. In retrospective cohort analysis, the moderate and highly active groups were at lower risk of hospitalizations (moderate: adjusted odds ratio [OR], 0.77 [95% CI, 0.61-0.98], high: OR, 0.76 [95% CI, 0.60-0.97]). In the follow-up prospective cohort analysis, light-intensity physical activity (LIPA) was associated with a lower risk of hospitalizations in one genetic cluster (adjusted hazard ratio [HR],

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

0.26 [95% CI, 0.08 – 0.88]), which was distinguished by three genetic variants: rs1432639, rs4543289, and rs11209948. Compliance with the guideline-level moderate-to-vigorous physical activities was not significantly related to the risk of hospitalizations.

Conclusions

Physical activities, particularly a certain amount of light-intensity physical activity, mitigate the risk of hospitalizations due to self-harm in depression patients with a specific genotype.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Introduction

Depressive disorder is one of the most prevalent mental disorders, affecting over 300 million people.[1] The prevalence of depressive disorder is positively related to age, which makes middle-aged and older adults a population group at higher risk.[2] Midlife and late-life depression are not only associated with cognitive impairment and dementia but also with intentional self-harm (ISH) behavior, including suicidal attempts. It is a widely held view that depression is the most common 'antecedent' of ISH. It is notable that self-harm is less prevalent in older adults, but when they commit ISH, it more often led to suicidal attempts. Older people commit suicide proportionally more than any other age group, and most of them suffer from depression.[3,4] In addition, ISH is a common reason for hospitalizations. As suicide prevention is being a priority in public health strategy in many countries, there have been several attempts to investigate the impacts of hospitalizations due to intentional self-harm (HISH). According to the previous studies, HISH cost the total of 162 million GBP in the UK in 2017 and direct cost of 60 million USD for Denmark, and it is even increasing. Furthermore, these costs for inpatient care accounts for a large portion of these expenses.[5,6] Taken together, the public health impact and use of medical resource highlights the need for study on HISH and its prevention.

Although the mechanism of ISH in patients with depression is still unclear, it has commonly been assumed that gene-environment interactions (GxEs) play a vital role in developing the condition.[5] In consideration of this, non-pharmacological intervention, which might affect the environmental factors has begun to be widely used, along with the concern about the effectiveness of antidepressants.[6,7]

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Physical activity, which refers to activity that involves bodily movement produced by skeletal muscles that requires energy expenditure, is one of the most commonly implemented treatment options in depression.[8,9] A considerable amount of literature has reported the effectiveness of physical activity in reducing depressive symptoms, including suicidal ideation, while it varies by genotype.[10-12]

Physical activity may be classified into four main categories according to the absolute intensity, or rate of energy expenditure: sedentary, light, moderate, and vigorous activity.[13] Despite its wide clinical use, there are still no clear recommendations on the level and intensity of physical activity tailored for depression patients to prevent serious consequences, such as HISH. For example, up-to-date guidelines from the World Health Organization (WHO) mostly focused on moderate-to-vigorous physical activity (MVPA) for the general population with little considerations on emerging and under-studied concepts such as light intensity physical activity (LIPA) or genotype-specific effect of physical activity.[9]

To address these evidence gaps, we examined the beneficial effect of physical activity on the risk of HISH in probable depression phenotypes. By utilizing the longitudinal nature and comprehensive data ranging from sociodemographic information to electronic health records (EHR) and accelerometry data, we sought to determine the optimal range and intensity of physical activity for midlife to late-life depression patients.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Method

UK Biobank

The UK Biobank project is a large prospective cohort data comprising 502,492 participants from England, Scotland, and Wales. Since 2006, extensive data including sociodemographic variables, online questionnaires, lifestyle factors, physical activity data, and genetic data, were collected on the participants who visited 22 assessment centers across the United Kingdom. The participants were aged 40-69 in recruitment, and the follow-up data until 2021 was included in this study. All the participants provided written informed consent, and the project has approval from the North West Multi-centre Research Ethics Committee (MREC) as a Research Tissue Bank (RTB) approval (REC reference: 21/NW/0157, IRAS project ID: 299116). Our research was conducted under application number 86585.

Phenotype and outcome

A comprehensive medical record of each participant, coded clinical events, such as consultations, diagnoses, procedures, and laboratory tests, and registration records, including admission and discharge information, are linked to the UK Biobank project. In this study, primary healthcare data (Category 3000), summary diagnoses (category 2002), and record-level patient data (Category 2006) were used to define the depression cohort and outcome.

The depressive patients cohort was constructed according to a definition provided by Smith et al.: participants who include any of the following ICD-10 diagnoses - F32 (depressive episode), F33

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

(recurrent depressive episode), F34 (persistent mood [affective] disorders), F38 (other mood [affective] disorders), and F39 (unspecified mood [affective] disorder); or participants who were depressed/down for a whole week (Field ID 4598) for at least two weeks (Field ID 4609) and ever seen a GP or psychiatrist for nerves, anxiety or depression (Field ID 2090 and 2100); or participants with anhedonia for a whole week (Field ID 4631) which lasted for at least two weeks (Field ID 5375) and ever seen a GP or psychiatrist for nerves, anxiety, or depression (Field ID 2090 and 2100).[14] Primary outcome of the study was ISH, which was defined using the hospital inpatient records of X60-X84 (intentional self-harm) diagnoses.

Genotyping and imputation

Genome-wide genotype data was collected on all participants using either the UK Biobank Lung Exome Variant Evaluation Axiom or the UK Biobank Axiom array and underwent quality-control procedures performed by Wellcome Trust Centre for Human Genetics, University of Oxford, UK.[15] Genotype data was imputed based on Haplotype Reference Consortium (HRC) and UK10K haplotype resource. Imputation data of over 90 million variants were available through the UK Biobank under approval. We selected 55 risk SNPs of major depressive disorder and depressive symptoms from published sources, which were available in imputed genotype data from the UK Biobank.[16-18] Participants with any missing value in imputation data or mismatch in reported genetic and reported sex were excluded from the analysis.

Physical activity

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Two types of physical activity assessment were used in this study: subjective and objective measurement for the phase 1 and 2, respectively. For phase 1, we used Metabolic Equivalent Task (MET) score based on the latest version of International Physical Activity Questionnaire – Short Form (IPAQ-SF).[19] After having adjusted for activity type (walking, moderate, and vigorous)-specific weight, MET score for each subject was derived and used to classify populations into three categories. Data was processed and categorized following the guideline published by IPAQ Research Committee.[20]

For objective assessment of the physical activity, E-mail invitations were sent to UK Biobank participants with valid e-mail addresses between February 2013 and December 2015. Over 100,000 participants have responded and agreed to provide the data and were sent Axivity AX3 wrist-worn triaxial accelerometer from May 2013, which was designed by Open Lab, Newcastle University. The participants were instructed to wear the device continuously for a whole week to track their entire physical activity.[21]

From measured raw acceleration data at 100Hz with a dynamic range ±8*g*, physical activity data were extracted after calibration procedures. We also classified the raw records of accelerometry data into four types of physical activity – sleep, sedentary behavior, LIPA, and MVPA – following the published methods of machine-learning-based behavior classification.[22]

Statistical analyses

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

This study follows an ambidirectional cohort study design, utilizing the prospective and retrospective nature of the UK Biobank cohort. Sixteen years of longitudinal follow-up time period was divided into two phases: from 2006 to May 2013 for Phase 1 and from June 2013 to 2021 for Phase 2.

Genotype data were binarized, corresponding to the presence of the risk SNPs. Subjects are then clustered based on the Euclidean distance of the binarized genetic vector by hierarchical clustering. Since the number of clusters is unknown, we found the optimal threshold of hierarchical clustering using a silhouette coefficient. We generated decision trees to investigate significant SNPs for each cluster.

Univariate and multiple logistic regression analyses were conducted for the retrospective cohort study in phase 1 (from 2006 to May 2013). Assessing the relationship between the intentional self-harm behavior and self-reported physical activity level measured with IPAQ in each genetic cluster, the multivariable logistic regression model was adjusted for the potential confounders, including age group, sex, ethnicity, current employment status, education level, Townsend deprivation index, smoking status, and alcohol consumption.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The prospective cohort for phase 2 (from June 2013 to 2021) consists of participants without HISH medical records until the beginning of the phase. To assess the effectiveness of physical activity as a non-pharmacological intervention for depressive symptoms, the optimal level of each physical activity had to be decided prior to analysis. Previous guidelines by the World Health Organization (WHO) have indicated the recommended level of MVPA, whereas there is no such recommendation for LIPA.[9] Therefore, the recommended level of LIPA was calculated using the methodology suggested by Chen et al.[23] Then penalized cox-proportional hazard regression analysis was conducted in order to examine the associations between the types and guideline or calculation-based levels of physical activity and incident HISH in each genetic cluster.

We used Python 3.10.9 and SciPy 1.10.1 for genetic stratification, and 'survival' 3.5.5, 'rms' 6.7-1, and 'CutpointsOEHR' 0.1.2 with a R version 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria) for other analyses. Mean and standard deviation (SD) are reported for continuous baseline characteristics except for the Townsend index (of which statistics are reported with median and interquartile range), and percentages are reported to present categorical variables. Odds ratios and hazard ratios are presented with a 95% confidence interval.

Results

Among 64,826 participants who met the criteria for the probable depression phenotype, 36,238 were eligible for the imputed genotype-based stratification. After excluding the subjects with missing data in covariates, 28,923 depressive participants were included in the retrospective

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

cohort study (eFigure 1 in Supplement). Participants without records of HISH until May 2013 and provided valid accelerometer data comprised the prospective cohort in Phase 2. Sociodemographic details and descriptive statistics of the cohort are presented in Table 1.

Genetic stratification

We found the three largest clusters from the result of hierarchical clustering; each consists of 6,681, 8,905, and 4,401 participants (Figure 1). Also, genetic features that distinguish each cluster were found through the decision tree algorithm. Following cohort studies were conducted targeting these three clusters.

Phase 1: Retrospective cohort study

In phase 1, we sought to examine the association between the IPAQ physical activity group and HISH risk in each genetic cluster. Overall, in unadjusted logistic regression models, participants in moderate or IPAQ physical activity groups were associated with significantly lower odds of HISH. There were similar findings in the models stratified with summed MET minutes per week for all activity. Multivariable models which were adjusted for covariates showed that more active IPAQ groups were at the lower risk odds of HISH in the full cohort (OR for the moderately active group: 0.771[0.609-0.977], OR for the highly active group: 0.764[0.602-0.971]), although there were mixed findings from each genetic cluster (Table 2).

Phase 2: Prospective cohort study

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Beginning in June 2013, we assessed the relationships between the accelerometer-measured physical activity levels and incident HISH in 7,137 participants during a mean (SD) follow-up period of 8.55 (0.44) years. After the classification of the accelerometer data, the cohort was stratified using the guideline-based thresholds for MVPA (≥150 minutes per week). Since the amount of LIPA had a U-shaped relationship with the log-relative hazards, the 'goldilocks' level for LIPA (between 2.38 and 9.53 hours a day) was derived by calculated optimal cut points (Figure 2).

The results of penalized cox-proportional hazard regression analysis are shown in Table 3. 'Goldilocks' amount of LIPA was associated with a lower risk of HISH in genetic cluster 2 (HR, 0.23; 95% CI, 0.07 – 0.76), and it was significant after adjusting for possible confounders (adjusted HR, 0.26; 95% CI, 0.08 – 0.88), although there were no significant differences observed in other clusters. No significant differences in risk of HISH were found between the levels of MVPA.

Discussion

The current study sought to examine the preventive effect of physical activity on the HISH of a population with depression phenotype. The findings from the retrospective cohort study in phase 1 revealed that depressive participants who engage in an average or higher level of physical activity tend to be at lower risk of HISH. In the subgroup analysis, the relationship between the overall physical activity level and incident HISH differed across the genetic clusters. A possible explanation for the mixed results and discrepancies across the genetic clusters is the genotype-

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

specific effect of physical activity, which in turn heightens the need for detailed research on the detailed classification of physical activity and its impact in groups with different genotypes.

Analysis of the prospective cohort settings in phase 2 showed that participation in machine-learning derived 'goldilocks' level of LIPA reduces the risk of HISH in genetic cluster 2. Since the effect was not significant in other clusters, it can thus be suggested that the calculated level of LIPA could be recommended for patients with specific genetic variants for cluster 2.

Meanwhile, the data from this study did not confirm the significant effect of guideline-based MVPA thresholds in the prevention of HISH. These findings are broadly consistent with recent studies that found the beneficial effect of LIPA on depressive symptoms rather than MVPA. Although existing studies and recommendations have focused on MVPA or the overall amount of physical activity, recent reports have suggested that objectively measured LIPA is more effective than another type of physical activity in alleviating depressive symptoms.[24-27] Because of their limitations in compliance and adherence to the physical activity prescription, LIPA could be a more appropriate option for the middle-aged and older population.[25]

Although the molecular mechanism of LIPA's preventive effect on depressive symptoms is not clear, one can assume that LIPA affects the pathophysiology of depression in the way general physical activity does. According to our analyses, there were three SNPs that distinguish cluster 2: rs1432639, rs4543289, and rs11209948. Previous results from GWAS reported that two of the three variants, rs1432639 and rs11209948, are related to obesity/BMI and body fat percentage,

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

respectively.[18,28] The nearest gene of the two variants is neuronal growth regulator 1 (NEGR1), which is suggested to be a central hub that links depression and obesity.[29] When considering the proposed mechanism of neuroplasticity of physical activity in depression, the two variants possibly involved in mediating neural plasticity along with NEGR1.[30] Taken together with the role of rs4543289 in severity and symptoms of depression, the complex interaction between these genetic components could provide a probable explanation for the greater effect of physical activity on HISH in cluster 2.[17]

One of the distinguishing features of our research is that we utilized various study designs and measurement tools available to extend the current findings. For example, throughout our study, the amount of physical activity was largely associated with a relatively lower risk of incidence HISH. Although the association was inconsistent in the retrospective cohort study, which made use of categorization derived from self-reported questionnaires, we took advantage of objectively measured physical activity and confirmed effectiveness in the prospective cohort settings. Recent evidence, ranging from the Mendelian randomization study and meta-analyses, has reported the possible discrepancies between subjective and objective measurement of physical activity, and objective measurement (i.e., accelerometer data) is more likely to be related to the improved outcomes in depression and related outcomes.[31-33] This discrepancy could be attributed to the bias of self-reported assessment tools. In spite of its convenience and cost-effectiveness, subjective assessment of physical activity is vulnerable to measurement error and various types of bias. These limitations could be more consequential when it comes to recall and social desirability bias in respondents in older age, or with mental disorders.[24,34] Moreover, it has been reported that participation in LIPA is especially prone to recall bias, which in turn provides a

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

possible explanation for the lack of previous research on it and emphasizes the need for epidemiological studies with objective physical activity assessment.[35]

To the best of our knowledge, this is the first study to examine the genotype-specific effect of LIPA on HISH in midlife and late-life depression. While depression in midlife and late-life is becoming more prevalent, serious consequences of the disorder, such as ISH, is recognized as a major public health concern. Furthermore, research on personalized intervention for depression has been mostly restricted to pharmacological treatment.[36] The present findings have implications for developing personalized recommendations for LIPA, while further research on more genotypes and psychophysiological factors is needed for a better understanding of the non-pharmacological treatment options for depression.

Limitations

There are several limitations to this study. First, this study was based on a single database, which in turn could be affected by possible biases such as ethnicity in the original dataset. Second, most of the variables were collected at a specific time point, which means that the variables might not fully represent the characteristics of the participants. For example, participants might have changed their physical activity pattern or lifestyle factors during the observation period or after the measurement. Third, the genetic stratification was based on previously reported risk SNPs, so there might have been other genetic variants that we have yet to account for, and would have uniquely found in the UK Biobank dataset. Fourth, since the sum of hours during a day is fixed, the effects of specific intensity and level of activity might partly be due to change in other types

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

of activities. For example, previous studies and guidelines reported the benefit of displacing the sedentary activity to LIPA or MVPA. Fifth, because the statistical model used in the analysis presupposes a 'U-shaped' relationship between the continuous variable and the log-relevant hazard ratio, the optimal range could not be calculated for MVPA, of which the relationship with the hazard ratio was not 'U-shaped.' However, the findings of this study are consistent with previous reports when applying guideline-based thresholds.[24]

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

References

1. Organization WH. *Depression and other common mental disorders: global health estimates*. 2017.
2. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*. 2020;396(10258):1204-1222. doi:10.1016/S0140-6736(20)30925-9
3. Skegg K. Self-harm. *The Lancet*. 2005;366(9495):1471-1483. doi:10.1016/S0140-6736(05)67600-3
4. Rodda J, Walker Z, Carter J. Depression in older adults. *BMJ*. 2011;343:d5219. doi:10.1136/bmj.d5219
5. Althoff RR, Hudziak JJ, Willemsen G, Hudziak V, Bartels M, Boomsma DI. Genetic and environmental contributions to self-reported thoughts of self-harm and suicide. *Am J Med Genet B Neuropsychiatr Genet*. Jan 2012;159b(1):120-7. doi:10.1002/ajmg.b.32010
6. Gartlehner G, Wagner G, Matyas N, et al. Pharmacological and non-pharmacological treatments for major depressive disorder: review of systematic reviews. *BMJ Open*. Jun 14 2017;7(6):e014912. doi:10.1136/bmjopen-2016-014912
7. Farah WH, Alsawas M, Mainou M, et al. Non-pharmacological treatment of depression: a systematic review and evidence map. *Evidence Based Medicine*. 2016;21(6):214-221. doi:10.1136/ebmed-2016-110522
8. López-Torres Hidalgo J, Aguilar Salmerón L, Boix Gras C, et al. Effectiveness of physical exercise in the treatment of depression in older adults as an alternative to antidepressant drugs in primary care. *BMC Psychiatry*. 2019/01/14 2019;19(1):21. doi:10.1186/s12888-018-1982-6
9. Fiona CB, Salih SA-A, Stuart B, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine*. 2020;54(24):1451. doi:10.1136/bjsports-2020-102955
10. Dotson VM, Hsu FC, Langaee TY, et al. Genetic Moderators of the Impact of Physical Activity on Depressive Symptoms. *J Frailty Aging*. 2016;5(1):6-14. doi:10.14283/jfa.2016.76
11. Ben S, Timothy O, Rachel C, et al. Effectiveness of physical activity interventions for improving depression, anxiety and distress: an overview of systematic reviews. *British Journal of Sports Medicine*. 2023;57(18):1203. doi:10.1136/bjsports-2022-106195
12. Vancampfort D, Hallgren M, Firth J, et al. Physical activity and suicidal ideation: A systematic review and meta-analysis. *Journal of Affective Disorders*. 2018/01/01/ 2018;225:438-448. doi:https://doi.org/10.1016/j.jad.2017.08.070
13. Piercy KL, Troiano RP, Ballard RM, et al. The Physical Activity Guidelines for Americans. *JAMA*. 2018;320(19):2020-2028. doi:10.1001/jama.2018.14854
14. Smith DJ, Nicholl BI, Cullen B, et al. Prevalence and Characteristics of Probable Major Depression and Bipolar Disorder within UK Biobank: Cross-Sectional Study of 172,751 Participants. *PLOS ONE*. 2013;8(11):e75362. doi:10.1371/journal.pone.0075362
15. Clare B, Colin F, Desislava P, et al. Genome-wide genetic data on ~500,000 UK Biobank participants. *bioRxiv*. 2017:166298. doi:10.1101/166298

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

16. Okbay A, Baselmans BML, De Neve J-E, et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics*. 2016/06/01 2016;48(6):624-633. doi:10.1038/ng.3552
17. Hyde CL, Nagle MW, Tian C, et al. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nature Genetics*. 2016/09/01 2016;48(9):1031-1036. doi:10.1038/ng.3623
18. Wray NR, Ripke S, Mattheisen M, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*. 2018/05/01 2018;50(5):668-681. doi:10.1038/s41588-018-0090-3
19. Craig CL, Marshall AL, SjÖStrÖM M, et al. International Physical Activity Questionnaire: 12-Country Reliability and Validity. *Medicine & Science in Sports & Exercise*. 2003;35(8)
20. Committee IR. Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ)-short and long forms. *http://wwwipaqkise/scoringpdf*. 2005;
21. Doherty A, Jackson D, Hammerla N, et al. Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. *PLOS ONE*. 2017;12(2):e0169649. doi:10.1371/journal.pone.0169649
22. Rosemary W, Shing C, Karl S-B, et al. Reallocation of time between device-measured movement behaviours and risk of incident cardiovascular disease. *British Journal of Sports Medicine*. 2022;56(18):1008. doi:10.1136/bjsports-2021-104050
23. Chen Y, Huang J, He X, et al. A novel approach to determine two optimal cut-points of a continuous predictor with a U-shaped relationship to hazard ratio in survival data: simulation and application. *BMC Medical Research Methodology*. 2019/05/09 2019;19(1):96. doi:10.1186/s12874-019-0738-4
24. Kandola A, Lewis G, Osborn DPJ, Stubbs B, Hayes JF. Depressive symptoms and objectively measured physical activity and sedentary behaviour throughout adolescence: a prospective cohort study. *The Lancet Psychiatry*. 2020;7(3):262-271. doi:10.1016/S2215-0366(20)30034-1
25. Ku P-W, Steptoe A, Liao Y, Sun W-J, Chen L-J. Prospective relationship between objectively measured light physical activity and depressive symptoms in later life. *International Journal of Geriatric Psychiatry*. 2018/01/01 2018;33(1):58-65. doi:https://doi.org/10.1002/gps.4672
26. Imai A, Kurihara T, Kimura D, Tanaka N, Sanada K. Association between non-locomotive light-intensity physical activity and depressive symptoms in Japanese older women: A cross-sectional study. *Mental Health and Physical Activity*. 2020/03/01/ 2020;18:100303. doi:https://doi.org/10.1016/j.mhpa.2019.100303
27. Helgadóttir B, Forsell Y, Hallgren M, Möller J, Ekblom Ö. Long-term effects of exercise at different intensity levels on depression: A randomized controlled trial. *Preventive Medicine*. 2017/12/01/ 2017;105:37-46. doi:https://doi.org/10.1016/j.ypmed.2017.08.008
28. Hübel C, Gaspar HA, Coleman JRI, et al. Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 2019/09/01 2019;180(6):428-438. doi:https://doi.org/10.1002/ajmg.b.32709

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

29. Noh K, Park J-C, Han J-S, Lee SJ. From Bound Cells Comes a Sound Mind: The Role of Neuronal Growth Regulator 1 in Psychiatric Disorders. *Exp Neurobiol*. 2020;29(1):1-10. doi:10.5607/en.2020.29.1.1

30. Kandola A, Ashdown-Franks G, Hendrikse J, Sabiston CM, Stubbs B. Physical activity and depression: Towards understanding the antidepressant mechanisms of physical activity. *Neuroscience & Biobehavioral Reviews*. 2019/12/01/ 2019;107:525-539. doi:https://doi.org/10.1016/j.neubiorev.2019.09.040

31. Choi KW, Chen C-Y, Stein MB, et al. Assessment of Bidirectional Relationships Between Physical Activity and Depression Among Adults: A 2-Sample Mendelian Randomization Study. *JAMA Psychiatry*. 2019;76(4):399-408. doi:10.1001/jamapsychiatry.2018.4175

32. Michalak J, Niemeyer H, Tschacher W, Baumann N, Chi Zhang X, Adolph D. Subjective and Objective Measures of Activity in Depressed and Non-depressed Individuals in Everyday Life. *Journal of Experimental Psychopathology*. 2022/04/01 2022;13(2):20438087221092582. doi:10.1177/20438087221092582

33. Vancampfort D, Firth J, Schuch FB, et al. Sedentary behavior and physical activity levels in people with schizophrenia, bipolar disorder and major depressive disorder: a global systematic review and meta-analysis. *World Psychiatry*. 2017/10/01 2017;16(3):308-315. doi:https://doi.org/10.1002/wps.20458

34. Heesch KC, van Uffelen JGZ, Hill RL, Brown WJ. What do IPAQ questions mean to older adults? Lessons from cognitive interviews. *International Journal of Behavioral Nutrition and Physical Activity*. 2010/05/11 2010;7(1):35. doi:10.1186/1479-5868-7-35

35. Matthews CE, Moore SC, George SM, Sampson J, Bowles HR. Improving Self-Reports of Active and Sedentary Behaviors in Large Epidemiologic Studies. *Exercise and Sport Sciences Reviews*. 2012;40(3)

36. Vos CF, ter Hark SE, Schellekens AFA, et al. Effectiveness of Genotype-Specific Tricyclic Antidepressant Dosing in Patients With Major Depressive Disorder: A Randomized Clinical Trial. *JAMA Network Open*. 2023;6(5):e2312443-e2312443. doi:10.1001/jamanetworkopen.2023.12443

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure legends

Figure 1. Dendrogram for the hierarchical agglomerative clustering

Figure 2. The 'U-shaped' relationship between the daily average hours of light-intensity physical activity and hazard ratios (HRs) of HISH

* Black, blue, and yellow lines are for the genetic cluster 1, 2, and 3, respectively. Red dotted line depicts a base level of HR (=1.0). At the point above the red dotted line, relative hazard is larger than 1, which means subjects in that point is at increased risk of HISH.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Tables

Table 1. Sample characteristics of the participants according to the IPAQ physical activity group

Baseline characteristics		**No. (%)**		
		Low (n = 3,432)	**Moderate (n = 6,419)**	**High (n = 6,135)**
Age		55.6 (7.9)	56.4 (8.0)	56.2 (8.2)
Sex	**Male**	1,366 (39.8)	2,287 (35.6)	2,393 (39.0)
	Female	2,066 (60.2)	4,132 (64.4)	3,742 (61.0)
Ethnicity	**White British**	3,184 (92.8)	5,880 (91.6)	5,685 (92.7)
Townsend deprivation index, median (interquartile range)		-1.6 (4.6)	-1.7 (4.3)	-1.8 (4.4)
Employment	**Paid/self-employment**	1,751(51.0)	3,316(51.7)	3,248 (53.0)
	Unemployed	1,681(49.0)	3,103 (48.3)	2,887 (47.0)
Educational attainment (years)	**Below 10**	1,250 (36.4)	2,057 (32.0)	2,053 (33.5)
	11-19	1,152 (33.6)	2,110 (32.9)	2,200 (35.9)
	Over 20	1,030 (30.0)	2,252 (35.1)	1,882 (30.7)
Smoking status	**Never smoked**	1,609 (46.9)	3,073 (47.9)	2,976 (48.5)
	Previous smoker	1,234 (36.0)	2,468 (38.4)	2,298 (37.5)
	Current smoker	589 (17.2)	878 (13.7)	861 (37.5)
Alcohol consumption	**Never**	392 (11.4)	608 (9.5)	539 (14.0)
	Special occasions only	552 (16.1)	849 (13.2)	738 (11.5)
	Moderate	1,841 (53.6)	3,640 (56.7)	3,595 (56.0)
	Heavy	647 (18.9)	1,322 (20.6)	1,263 (19.7)

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 2. Unadjusted and adjusted odds ratios (ORs) for risk of HISH according to the IPAQ physical activity group in the three largest genetic clusters

Genetic cluster	**N**	**IPAQ group**	**Model 1 OR (95% CI)**	**Model 2 aOR (95% CI)**[a]
1	5,365	Low	Reference	
		Moderate	0.60 (0.41 – 0.88)	0.71 (0.47 – 1.05)
		High	0.65 (0.45 – 0.96)	0.77 (0.52 – 1.15)
2	7,086	Low	Reference	
		Moderate	0.57 (0.39 – 0.82)	0.70 (0.48 – 1.01)
		High	0.70 (0.49 – 1.00)	0.85 (0.59 – 1.22)
3	3,535	Low	Reference	
		Moderate	0.84 (0.53 – 1.36)	1.03 (0.64 – 1.69)
		High	0.49 (0.29 – 0.83)	0.61 (0.35 – 1.04)
Total	15,986	Low	Reference	
		Moderate	0.64 (0.52 – 0.80)	0.77 (0.61 – 0.98)
		High	0.64 (0.50 – 0.80)	0.76 (0.60 – 0.97)

a. The aORs from Model 2 are adjusted for all covariates listed in Table 1.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 3. Unadjusted and adjusted hazard ratios (HRs) for HISH in group within recommended level of light-intensity physical activity (LIPA) and moderate-to-vigorous physical activity (MVPA)

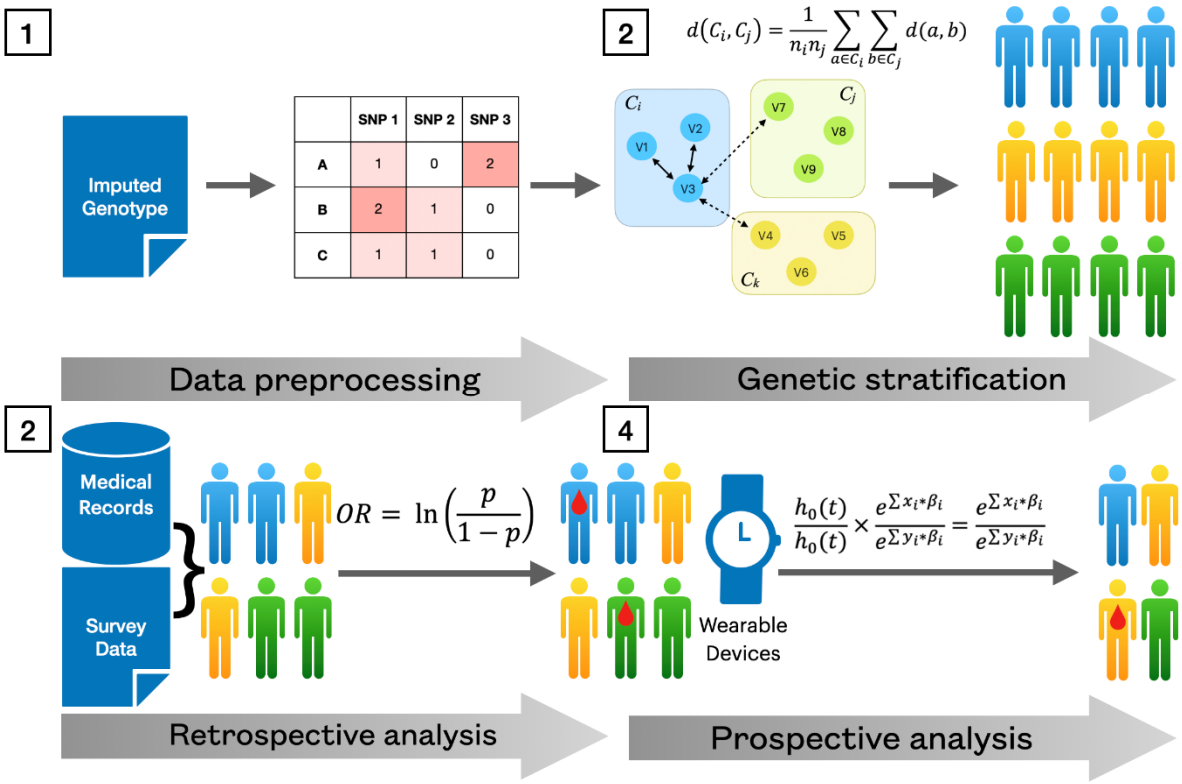
Genetic Cluster	**No. of events (%)**	**No. of participants**	**Model 1 HR (95% CI)**	**Model 2 adjusted HR (95% CI)**
Light-intensity Physical Activity[a]				
1	13 (1.1)	1,204	0.61 (0.14 – 2.69)	0.65 (0.15 - 2.93)
2	11 (0.7)	1,563	0.26 (0.08 – 0.84)	0.26 (0.08 - 0.88)
Moderate-to-vigorous Physical Activity (MVPA)				
1	10 (1.3)	746	1.53 (0.52 – 4.49)	1.18 (0.39 – 3.59)
2	9 (0.9)	983	1.42 (0.48 – 4.24)	1.38 (0.44 – 4.30)
3	7 (1.3)	528	1.14 (0.33 – 3.89)	0.85 (0.24 – 2.97)

a. The results for cluster 3 was not applicable due to the insufficient number of events. No significant difference was found between the group with or within LIPA recommendation in the log-rank test (p = 0.3).

b**.** The adjusted HRs from Model 2 are adjusted for all covariates listed in Table 1.

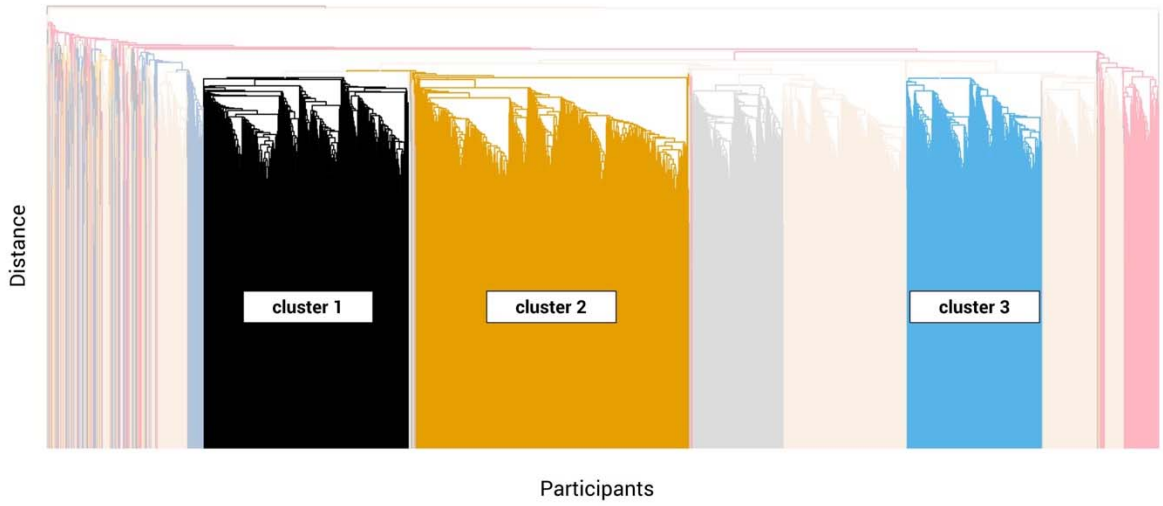
medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure1. Scheme of the bidirectional analytic model to investigate the genotype-specific effectiveness (BAIGE)



medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 2. Dendrogram for the hierarchical agglomerative clustering



medRxiv preprint doi: https://doi.org/10.1101/2024.05.10.24307087; this version posted May 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

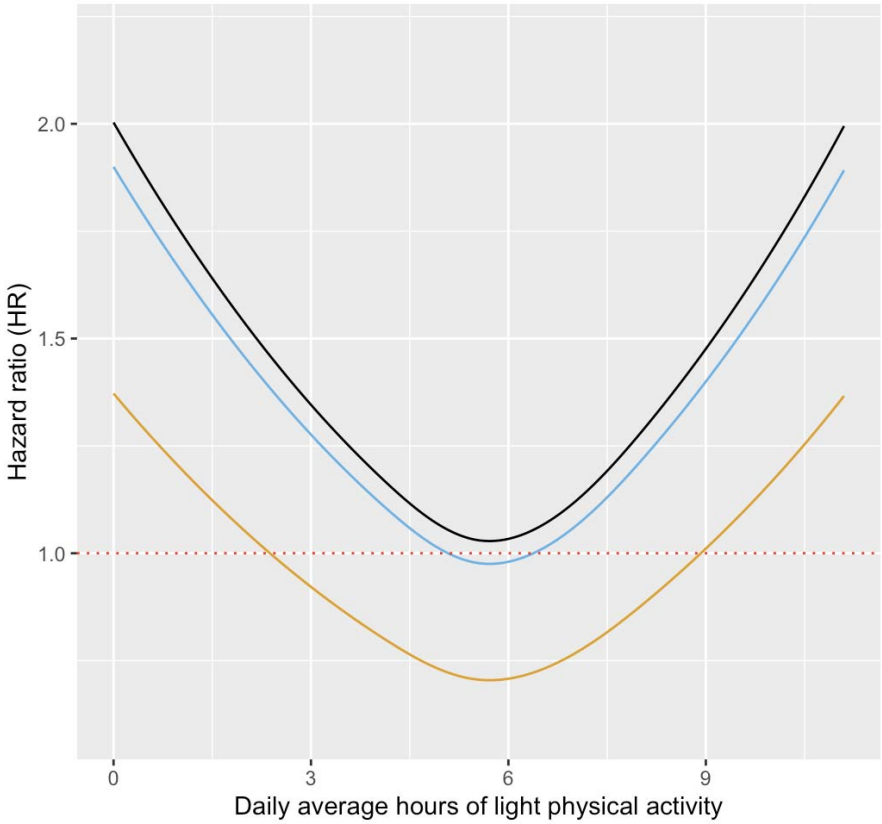



Figure 3. The 'U-shaped' relationship between the daily average hours of light-intensity physical activity and hazard ratios (HRs) of HISH