INVESTIGATING THE EFFECTS OF PHYSICAL ACTIVITY ON LEVELS OF GLYCOPROTEIN ACETYLS

By

SYDNEY THERESE WHITEMAN

A Thesis Submitted to The W.A. Franke Honors College

In Partial Fulfillment of the Bachelors degree

With Honors in

Biology

THE UNIVERSITY OF ARIZONA

M A Y 2 0 2 3

Approved by:

Dr. Yann Klimentidis
Department of Epidemiology and Biostatistics

Abstract

Levels of biomarkers in the blood have been known to signal whether abnormal or normal processes are occurring in the body. Glycoprotein acetyls are a relatively new inflammatory biomarker that has been linked to early cardiovascular disease and other inflammatory diseases. Understanding the relationship of glycoprotein acetyls with certain exposures can help inform us of ways to lower inflammatory disease risk. In this study, physical activity was investigated as the exposure. Summary statistics from genome-wide association studies (GWAS) of self-reported and wrist-worn devices from the UK Biobank were used. We found that physical activity and levels of glycoprotein acetyls had an inverse relationship, such that higher levels of genetically-predicted physical activity were correlated with lower levels of glycoprotein acetyls. Identifying this relationship expands our understanding of how physical activity can affect health. This information can allow us to utilize physical activity as a preventative health measure, and to lower risks of inflammatory and cardiovascular disease.

Introduction

A person's health and wellbeing can be attributed to many factors. Physical activity has been identified as one of the most important lifestyle factors to improve overall health, and to lower risk for a wide range of diseases, including chronic diseases, obesity, cardiovascular disease, and stroke (Pratt et. al, 1999). Furthermore, studies have also been conducted to assess the benefits of physical activity with aging, with exercise showing an association with reduction in age-related bone loss, an increase in muscle mass, and improvement in muscle strength and coordination (Galloway & Jokl, 2000). In addition, studies have shown how one's immune response also benefits from regular exercise, with the total number of sick days being lower in people who engage in regular vigorous physical activity (Miles, 2007). However, while these benefits may be known, only 27% of US adults meet recommended levels of moderate or vigorous physical activity, with 29.2% reporting no regular physical activity outside of work (Pratt et. al, 1999). Differences in physical activity can also be observed amongst genders and races, with men and White Americans getting the most exercise (Pratt et. al, 1999). Differences in physical activity also exist among different levels of socioeconomic status, likely contributing to prominent health disparities (Gordon-Larsen et. al, 2006). Considering the benefits of physical activity for health and disease prevention, the lack of physical activity in the U.S population is concerning on a public health level.

Levels of certain biomarkers circulating in the blood can be used as indicators of normal or abnormal processes in the body (National Cancer Institute, n.d). Previous studies have looked at the relationship between physical activity and different biomarkers, including HDL cholesterol, vitamin D, apolipoprotein B, and hemoglobin A1c. A beneficial effect was observed

from physical activity on these biomarkers (Chomistek et. al., 2011) Another study found that physical activity helped to decrease the levels of risk factors for cardiovascular disease (CVD), such as inflammatory/hemostatic biomarkers (hsCRP, fibrinogen, and sICAM-1), and traditional lipid biomarkers such as total, LDL, and HDL cholesterol (Mora et. al., 2007) In addition, studies have also concluded that men who reported greater vigorous-intensity physical activity had a lower BMI and more favorable biomarker profiles compared to men reporting less activity (Chomistek et. al., 2011). Additionally, both BMI and physical activity are significantly associated in a linear manner with most biomarker levels. In this study, we will focus on whether levels of glycoprotein acetyls are potentially lowered through physical activity. Glycoprotein acetyl (GlycA) level is a relatively new biomarker of inflammation that has been linked to early cardiovascular disease in the young, and has proven to be stable compared to other biomarkers such as C-reactive protein or CRP (Chiesa, 2022).

In this study, we used Mendelian randomization (MR) to assess the evidence for a causal role of physical activity in levels of GlycA. MR is an investigative tool that uses genetic information to assess the causal relationship between an exposure and an outcome. MR utilizes genetic variants that have been randomly assigned at conception and have remained unchanged throughout one's lifetime (Roberts, 2018). This approach acts on three main assumptions: the genetic variant is associated with the exposure, the variant does not associate with the confounder, and the genetic variant influences the outcome through the exposure (Lee, 2020). Typically, in these studies, single nucleotide polymorphisms (SNPS) are used as the genetic variant. MR is a useful tool for studying the causes and effects of diseases, where manipulation of the exposure in a randomized trial might not be practical nor ethical. In this

project, data was collected from genome-wide association studies ('GWAS') from UK Biobank regarding five different types of physical activity: moderate to vigorous activity, vigorous activity, strenuous sports, accelerometer-based physical activity, and average acceleration.

Methods

UK Biobank

In this study, we used data from the UK Biobank, which is a large-scale biomedical database that has genetic and health information from half a million participants, aged 40-69, living in the United Kingdom. The data from UK Biobank consists mainly of people of European descent. All participants provided informed consent, and ethical approval was given to the UK Biobank study by the North West Multicentre Research Ethics Committee, the National Information Governance Board for Health & Social Care, and the Community Health Index Advisory Group.

Exposure Measures

We used the results of a previously conducted GWAS of physical activity (Klimentidis et al., 2018). Briefly, physical activity measurements were obtained from self-reported data, specifically from a touchscreen questionnaire that asked questions about levels of moderate and vigorous physical activity. In the questionnaire, participants were asked questions regarding how many days they exercise in a moderate (cycling at normal pace, carrying light loads) or vigorous (fast cycling, heavy lifting) way. Physical activity was then classified as vigorous (VPA) or moderate-to-vigorous (MVPA) physical activity. Individuals who preferred not to answer or did

not know were excluded. Also, those who reported more than 16 hours of either MVPA or VPA were also excluded. In addition, questions were asked about the specific types of physical activity. Individuals who spent 2-3 days doing strenuous sports or other exercises (SSOE) were then grouped together. In addition, approximately 100,000 participants wore an Axivity AX3 wrist-worn accelerometer for seven days. The two measures derived from this data were overall acceleration average and the fraction of accelerations > 425 milli-gravities (mg). Participants that had less than three days of data, or those missing data in each one-hour period of the 24-hour cycle, were excluded.

Outcome Measures

The measures of GlycA were part of a metabolomics project in the UK Biobank. An NMR metabolomics platform (Nightingale Health Plc; biomarker quantification version 2020) was used to conduct metabolomic profiling in the plasma samples collected at the baseline visit of ~120,000 UKB participants who were chosen at random. This platform was able to quantify 249 metabolic biomarkers (168 directly measured, and 81 ratios of those measures) as well as lipoprotein subclass distribution, particle size and composition (Bragg et. al, 2022). This Nightingale NMR platform has obtained regulatory approval, with 37 biomarkers approved for diagnostics use (Julkunen et. al, 2021).

Statistical Analysis

To examine the association of genetically predicted levels of physical activity with GlycA, Mendelian randomization was utilized. Physical activity GWAS summary statistics were

downloaded from the GWAS catalog, and the GWAS of GlycA was obtained from the MRC IEU OpenGWAS project (https://gwas.mrcieu.ac.uk/). Selection of the exposure SNPs was done at two different p-value thresholds: p<5E-6 and p<5E-8. Additionally, the *TwoSampleMR* package in R was utilized. The summary statistics from the GWAS of physical activity were read into R, and restricted to those with a p-value less than 5E-6 or 5E-8. The results of the exposure SNPs in the GlycA outcome GWAS were then harmonized so that the beta-coefficients referred to the same allele in both data sets. We performed an MR analysis using the inverse-variance weighted (IVW), MR-Egger, weighted median, simple mode, and weighted mode methods. Each type of MR varies in what is taken into account, as MR-Egger accounts for pleiotropy and the IVW does not. Using all five types of MR helps to fully understand the potential effects that physical activity has on GlycA. The effect measure that we primarily used was the IVW beta-coefficient and its 95% confidence interval. All analyses were performed using R version 4.2.1.

Results

Each physical activity measurement was tested using two different levels of statistical significance. Each measurement was also run with five types of Mendelian randomization analyses with the exception of the fraction of accelerations > 425 milli-gravities (mg), for which we only could calculate the IVW estimate due to a low number of exposure SNPs.

Self-Reported Physical Activity Measurements

The self-reported physical activity measurements included moderate to vigorous physical activity (MVPA), vigorous physical activity (VPA), and strenuous sports and other

exercises (SSOE). For each of these measurements, a negative correlation was observed. For MVPA, the beta-coefficient was less than zero for the significance level of 5E-8 for each type of MR, with the exception of MR Egger and inverse variance weighted (IVW). MR Egger had a beta-coefficient of 2.05 (SE: 0.5), while IVW had a beta-coefficient of 0.05 (SE: 0.14). At the p-value of 5E-6, MR Egger had a beta-coefficient of 0.2 (SE: 0.214). As seen in Figure 1, the confidence intervals for both p-values of 5E-6 and 5E-8 cross the red line located at 0 for the different types of MR. This suggests that there is little evidence that MVPA has a significant effect on GlycA.

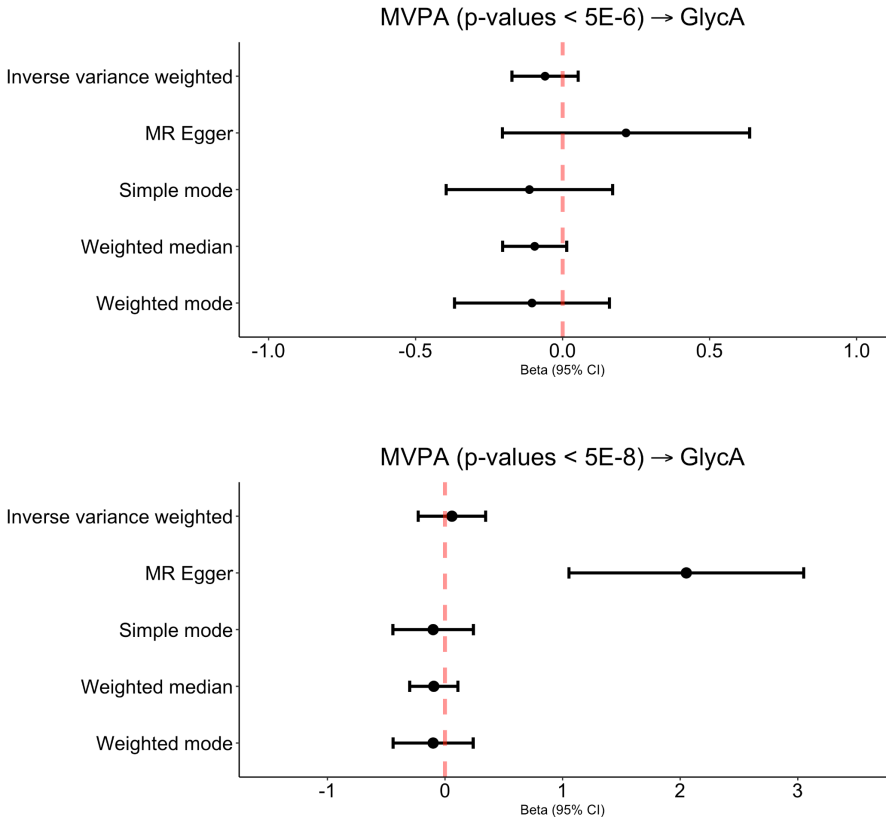

Figure 1. Forest plot of moderate to vigorous physical activity (MVPA) measurements and glycoprotein acetyls.

For VPA at the significance level of 5E-8, all beta-coefficients were less than 0, potentially suggesting an overall negative correlation. At the significance level of 5E-6, the beta-coefficient at MR Egger was above 0, with a beta-coefficient of 0.277 (SE: 0.3). This can be seen in Figure 2.

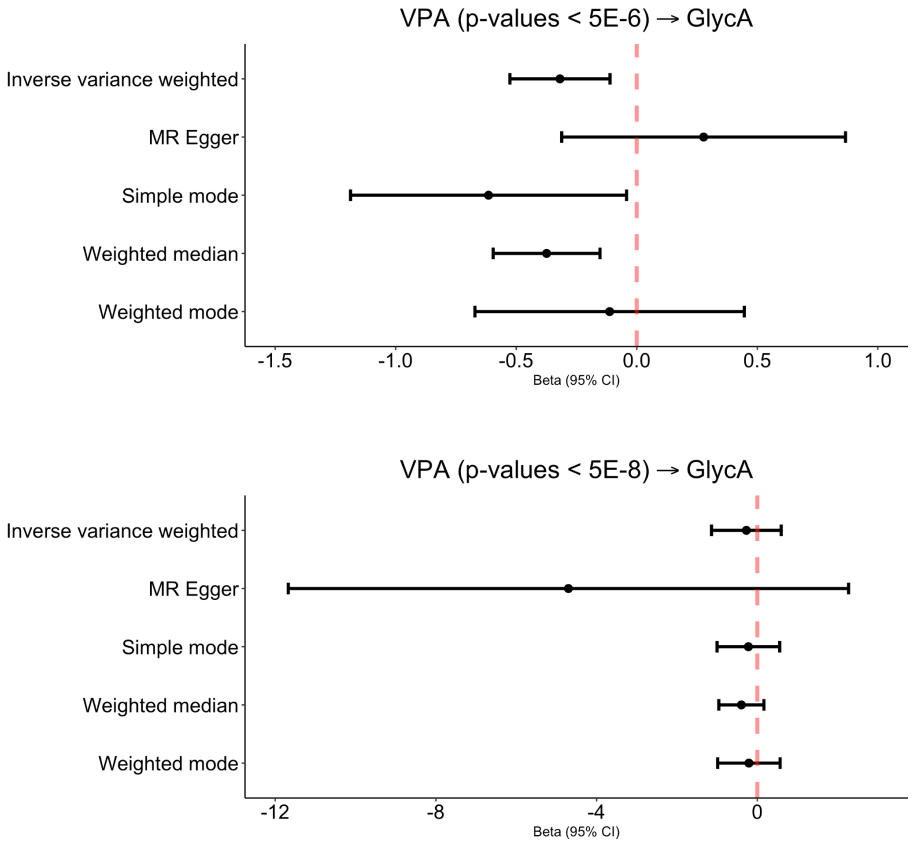

Figure 2. Forest plot of vigorous physical activity measurements (VPA) and glycoprotein acetyls.

For SSOE, we found that there was a majority of negative correlation. Lower GlycA levels are correlated with high levels of strenuous sports and other exercises. For MR Egger, the beta coefficient is slightly larger than 0 for both p-value of 5E-6 and 5E-8, at 0.042 (SE: 0.42) and 0.012 (SE: 1.8). Figure 3 summarizes these results.

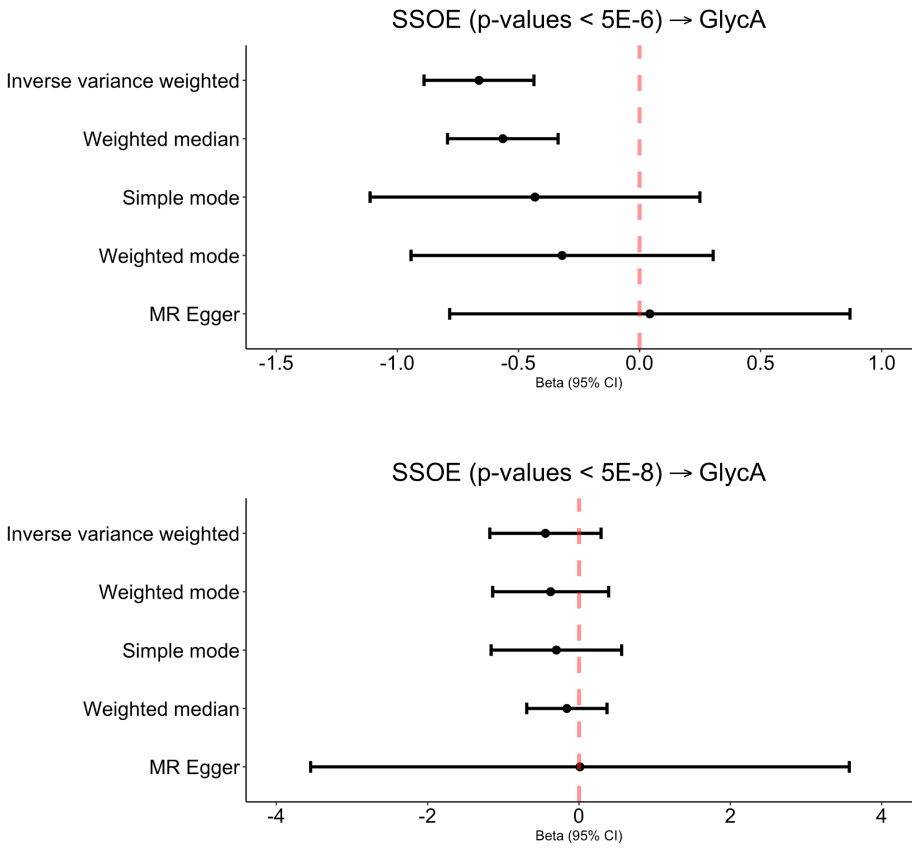

Figure 3. Forest plot of data collected from the group for strenuous sports and other exercises (SSOE) and glycoprotein acetyls.

Accelerometer Physical Activity Measurements

The accelerometer-derived physical activity measurements included average acceleration and the fraction of accelerations > 425 milli-gravities (mg) (ACC > 425mg). For average acceleration, all beta-coefficients were below zero at the significance level of 5E-6. For 5E-8, three out of the five beta-coefficients were slightly larger than 0. For MR Egger the beta-coefficient was 0.046 (SE: 0.05), for simple mode it was 0.009 (SE: 0.02), and for weighted mode it was 0.01 (SE: 0.02). This is summarized in Figure 4.

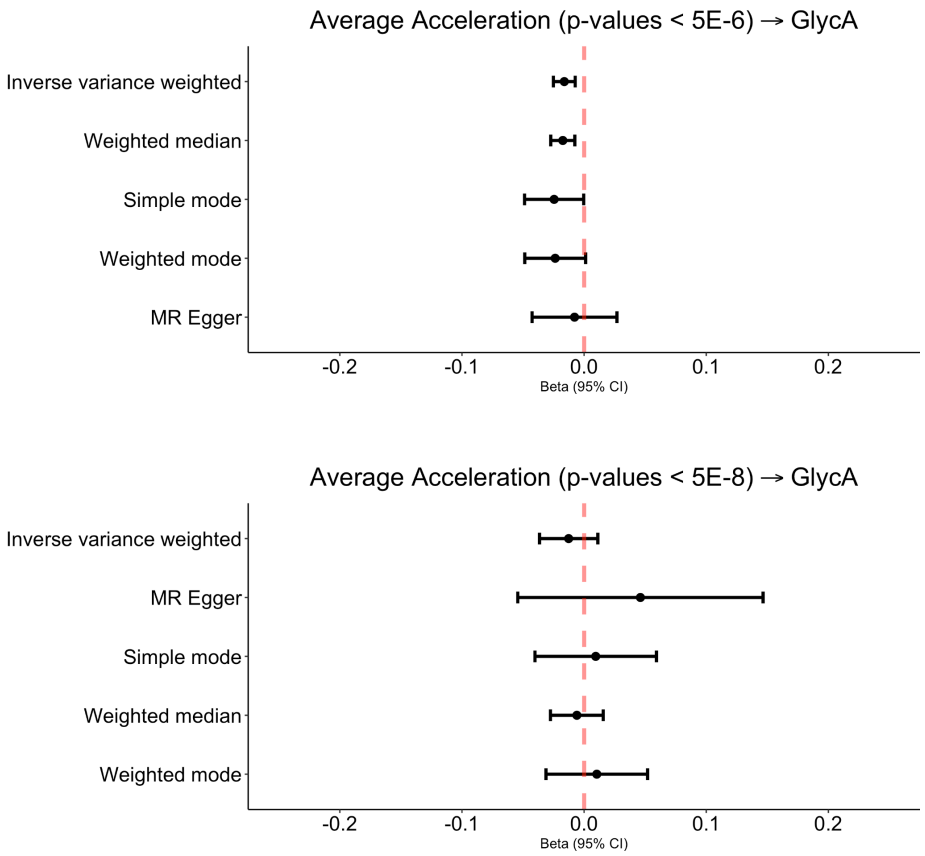

Figure 4. Forest plot for average acceleration and glycoprotein acetyls.

For accelerometer data, only inverse variance weighted was used at the significance level of 5E-8, with a beta-coefficient less than 0. At the significance level of 5E-6, all beta-coefficients were below 0, suggesting a negative correlation between GlycA and the fraction of accelerations. Figure 5 summarizes these results.

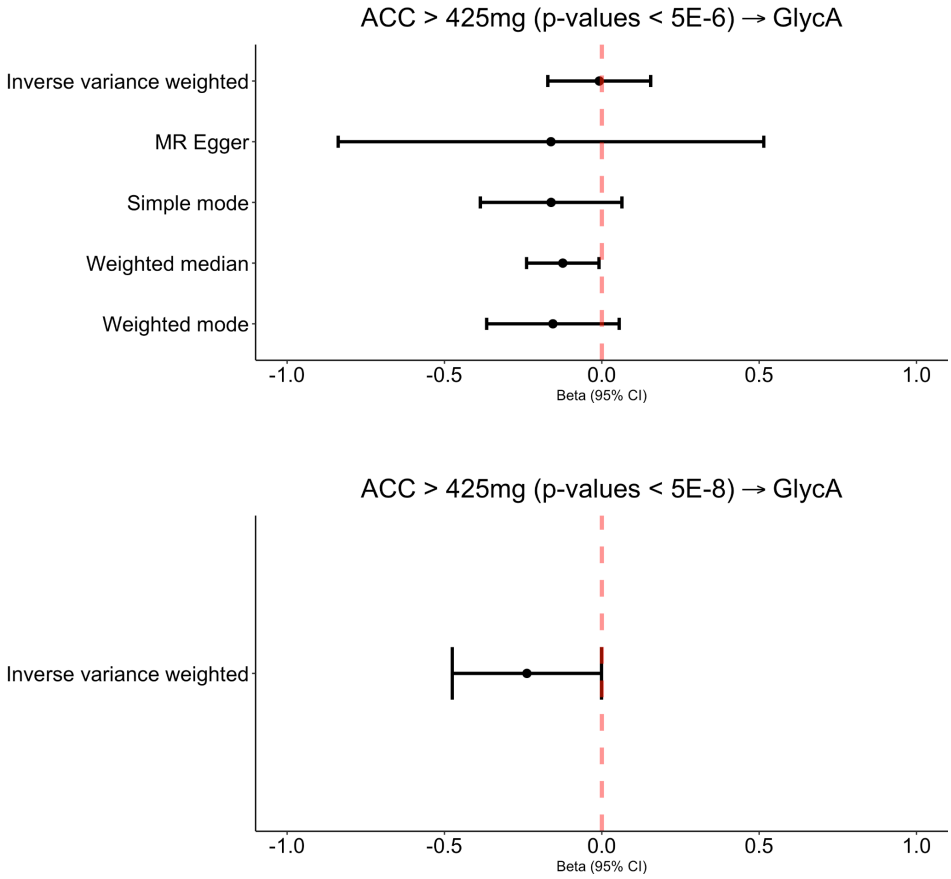

Figure 5. Forest plot for the fraction of accelerations (ACC > 425 milli-gravities) and glycoprotein acetyls.

Discussion

This study aimed to understand the effects of different types of physical activity on the inflammatory biomarker GlycA. Self-reported measurements of physical activity showed a largely negative correlation with GlycA, with the beta-coefficients landing lower than zero. Moderate to vigorous activity was an exception, showing little evidence for a significant effect. This suggests that higher levels of physical activity, such as vigorous activity as well strenuous sports and other exercises, are associated with lower levels of GlycA. Accelerometer-based physical activity data also suggested the same relationship. As seen in figures 1-4, MR Egger

produced a wide 95% confidence interval with the beta-coefficient landing above 0. This can be attributed to the nature of that type of MR, as MR Egger can be less powerful, but more conservative, than the other methods.

GlycA has been previously studied in relation to cardiovascular disease, and a similar relationship between GlycA and physical activity has been found. In a study produced by the American Heart Association, GlycA was studied in two groups, one aged 15 and one aged 24. It was found that there was an association between GlycA and markers of early cardiovascular risk. They also found that in individuals as young as 15 years old, there was observed evidence of increased GlycA levels for those who had low levels of physical activity (Chiesa et. al, 2022). This aligns with the relationship found in our analysis.

The strengths of this study include a large sample size of participants from the UK Biobank. Another strength is that this study included measurements from both self-reported and accelerometer-based data, allowing for two different sources of physical activity measurement. A limitation of this study is that the majority of the participants are older adults, with largely European ancestry. Participants also volunteer to be a part of this study, which could make the group less representative of the entire population. Another limitation, as seen in the accelerometer-based data, is that there are not many SNPs associated with the physical activity measurement, making this instrument not as strong as others.

Further research could be conducted to include participants of different ancestry backgrounds, and to see if additional SNPs associated with physical activity are found. Further research could also be conducted with different inflammatory biomarkers, such as cytokines or

pro-atherogenic enzymes, to further strengthen the evidence for a causal protective effect from physical activity.

Conclusion

In conclusion, this analysis suggests that higher levels of physical activity are correlated with lower levels of GlycA. As GlycA is an inflammatory biomarker, studying the effects of physical activity on inflammation can help to further understand its importance on overall and preventive health. GlycA can not only be used to help signify risks of cardiovascular disease, but other inflammatory diseases as well. Solidifying the effect physical activity has on inflammation can help us to use exercise in a more targeted approach, and to help reduce the risk of inflammatory diseases as well.

References:

Assah, F., Ekelund, U., Brage, S., Mbanya, J., Wareham, N. (2011). Urbanization, Physical Activity, and Metabolic Health in Sub-Sahrana Africa. *Diabetes Care* 1 February 2011; 34 (2): 491–496. https://doi.org/10.2337/dc10-0990

Ahmend, H., Blaha, M., Nasir, K., Riveria, J., Blumenthal, R. (2012). Effects of Physical Activity on Cardiovascular Disease. The American Journal of Cardiology, 109(2), 288-295. https://www.sciencedirect.com/science/article/abs/pii/S0002914911027597?casa_token=S7mil79jHgUAAAAA:0lTkIzwHZBQPiXGRoVpv2WMvV5LK54HMcbmr7Ylcb9WM-T4uIbLJa636nssQ5n0x_6E0iKuhal4

Center for Disease Control. (2018, August 6*) How is Mendelian Randomization and How Can it be Used as a Tool for Medicine and Public Health? Opportunities and Challenges*. CDC. https://www.cdc.gov/genomics/events/precision_med_pop.htm#:~:text=Mendelian%20randomization%20is%20a%20method,on%20disease%20in%20observational%20studies

Center for Disease Control (2023, Feburary 14). Physical Activity Boosts Brain Health. Center for Disease Control and Prevention. https://www.cdc.gov/nccdphp/dnpao/features/physical-activity-brain-health/index.html#:~:text=Physical%20activity%20can%20help%20you,of%20cognitive%20decline%2C%20including%20dementia.

Chiesa, S. Charakida, M., Georgiopoulos, G., Roberts, J., Stafford, S., Park, C., Mykkanen, J., Kahonen, M., Lehtimaki, T., Ala-Korpela, M., Raitakari, O., Pietlalnen, M., Pussinen, P., Muthurangu, V., Hughes, A., Satta, N., Timpson, N., Deanfield, J. (2022, February 12). Glycoprotein Acetyls: A Novel Inflammatory Biomarker of Early Cardiovascular Risk in the Young. Journal of the American Heart Association, 11(4), https://doi.org/10.1161/JAHA.121.024380

Chomistek, A. K., Chiuve, S. E., Jensen, M. K., Cook, N. R., & Rimm, E. B. (2011). Vigorous physical activity, mediating biomarkers, and risk of myocardial infarction. *Medicine and science in sports and exercise*, *43*(10), 1884–1890. https://doi.org/10.1249/MSS.0b013e31821b4d0a

Ertek, S., & Cicero, A. (2012). Impact of physical activity on inflammation: effects on cardiovascular disease risk and other inflammatory conditions. *Archives of medical science : AMS*, *8*(5), 794–804.
https://doi.org/10.5114/aoms.2012.31614

Galloway, M. T., & Jokl, P. (2000). Aging successfully: the importance of physical activity in maintaining health and function. *The Journal of the American Academy of Orthopaedic Surgeons*, *8*(1), 37–44. https://doi.org/10.5435/00124635-200001000-00004

Gordon-Larsen, P., Nelson, M., Page P., Popkin, B.; Inequality in the Built Environment Underlies Key Health Disparities in Physical Activity and Obesity. *Pediatrics* February 2006; 117 (2): 417–424. 10.1542/peds.2005-0058

Idle, J., Gonzalez, F. (2007) Metabolomics, *Cell Metabolism,* 6(5), 348-351. ISSN 1550-4131, https://doi.org/10.1016/j.cmet.2007.10.005.
https://www.sciencedirect.com/science/article/pii/S1550413107002987

Julkunen, H., Cichonska, A., Slagboom, P., Wurtz, P. (2021) Metabolic biomarker profiling for identification of susceptibily to sever pneumonia and COVID-19 in the general population. eLife. https://elifesciences.org/articles/63033#:~:text=The%20biomarkers%20span%20multiple%20metabolic,levels%20and%20high%20chronic%20inflammation.

Klimentidis, Y. C., Raichlen, D. A., Bea, J., Garcia, D. O., Wineinger, N. E., Mandarino, L. J., Alexander, G. E., Chen, Z., & Going, S. B. (2018). Genome-wide association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple variants including CADM2 and APOE. *International journal of obesity (2005)*, *42*(6), 1161–1176. https://doi.org/10.1038/s41366-018-0120-3

Lee, Y. (2020). Overview of Mendelian Randomization Analysis, Journal of Rheumatic Diseases, 27(4): 241-246. https://doi.org/10.4078/jrd.2020.27.4.241

Mayo Clinic Staff. (2021, May 6). *Metabolic Syndrome.* Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/metabolic-syndrome/symptoms-causes/syc-20351916#:~:text=Metabolic%20syndrome%20is%20a%20cluster,abnormal%20cholesterol%20or%20triglyceride%20levels.

Miles, L. (2007), Physical activity and health. Nutrition Bulletin, 32: 314-363. https://doi.org/10.1111/j.1467-3010.2007.00668.x

Mora, S., Cook, N., Buring, J. E., Ridker, P. M., & Lee, I. M. (2007). Physical activity and reduced risk of cardiovascular events: potential mediating mechanisms. *Circulation*, *116*(19), 2110–2118. https://doi.org/10.1161/CIRCULATIONAHA.107.729939

NCI Dictionary of Cancer terms. National Cancer Institute. (n.d.). Retrieved March 14, 2023, from https://www.cancer.gov/publications/dictionaries/cancer-terms/def/biomarker#:~:text=(BY%2Doh%2DMAR%2D,molecular%20marker%20and%20signature%20molecule.

Pratt, M., Macera, C., Blanton, C. (1999) *Levels of physical activity and inactivity in children and adults in the United States: current evidence and research issues*. Medicine & Science in Sports & Exercise. https://journals.lww.com/acsm-msse/Fulltext/1999/11001/Levels_of_physical_activity_and_inactivity_in.7.aspx

Roberts R. (2018). Mendelian Randomization Studies Promise to Shorten the Journey to FDA Approval. *JACC. Basic to translational science*, *3*(5), 690–703. https://doi.org/10.1016/j.jacbts.2018.08.001

Sharp, G. (2020, March 10) *Introduction to Mendelian Randomization* [Video]. YouTube. https://www.youtube.com/watch?v=isMZhasFc4M