



OPEN

## Using a two-sample mendelian randomization analysis to explore the relationship between physical activity and Alzheimer's disease

Bowen Zhang, Xiaowen Huang, Xiliang Wang, Xiaorui Chen, Caifang Zheng, Weihao Shao, Gaili Wang & Weidong Zhang

Evidence from previous epidemiological studies on the effect of physical activity on the risk of Alzheimer's disease (AD) is conflicting. We performed a two-sample Mendelian randomization analysis to verify whether physical activity is causally associated with AD. This study used two-sample Mendelian randomization (MR) analysis to estimate the association between physical activity (including overall activity, sedentary behavior, walking, and moderate-intensity activity) and AD. Genetic instruments for physical activity were obtained from published genome-wide association studies (GWAS) including 91,105 individuals from UK Biobank. Summary-level GWAS data were extracted from the International Genomics of Alzheimer's Project IGAP (21,982 patients with AD and 41,944 controls). Inverse Variance Weighted (IVW) was used to estimate the effect of physical activity on AD. Sensitivity analyses including weighted median, MR-Egger, MR-PRESSO, and leave-one-out analysis were used to estimate pleiotropy and heterogeneity. Mendelian randomization evidences suggested a protective relationship between walking and AD (odds ratio (OR) = 0.30, 95% confidence interval (CI), 0.13–0.68,  $P = 0.0039$ ). Genetically predicted overall activity, sedentary behavior, and moderate-intensity activity were not associated with AD. In summary, this study provided evidence that genetically predicted walking might associate with a reduced risk of AD. Further research into the causal association between physical activity and AD could help to explore the real relationship and provide more measures to reduce AD risk.

Alzheimer's disease (AD) is the most prevalent cause of dementia<sup>1</sup>. AD is a kind of brain disease that happens to elder people and is responsible for the slow and gradual deterioration in thinking, memory, and language competence, and could also change the personality<sup>1</sup>. AD is the most prevalent neurodegenerative disorder in the world<sup>2</sup>.

There are 44 million individuals living with AD around the world according to the current estimate<sup>3</sup>. The prevalence of AD in the European population is estimated at 5.05%<sup>4</sup>. The incidence of African Americans is approximately twice that among European Americans<sup>5</sup>. The United States, especially in the West and Southeast might experience the largest growth of AD cases and the number of AD patients could grow to 13.8 million by 2060 years<sup>6</sup>. The growing number of AD cases might cause a huge social-economic burden. Prevention of AD targeting risk factors modification has the potential to curb the increasing number of people living with AD and the increasing economic burden.

Observational studies had identified thirty-three percent of AD cases around the world might be due to several potentially modifiable risk factors after accounting for nonindependence between several risk factors, and the incidence might be decreased by using effective methods targeting to reduce the prevalence of modifiable risk factors such as physical inactivity<sup>7</sup>.

Compared to sedentary lifestyle, long-term exercise could delay the onset of physiological memory loss in middle-aged trained men<sup>8</sup>. Randomized clinical trials (RCT) also had been conducted to investigate whether interventions targeting inactivity could reduce the risk of AD, dementia, or cognitive function. A large and long-term RCT proved that physical activity could improve or at least maintain cognitive functioning in at-risk elder individuals<sup>9</sup>. An intervention study showed that physical activity could provide a modest improvement in cognition for individuals with subjective memory impairment as well<sup>10</sup>. Furthermore, previous meta-analyses including

Department of Epidemiology, School of Public Health, Zhengzhou University, Zhangzhou 450001, Henan, People's Republic of China. ✉email: imooni@163.com

several RCTs also indicated that physical activity intervention showed a statistically significant improvement in cognition of patients diagnosed with AD or slow down the decline of cognition, and different amounts of physical activity can make different effects<sup>11,12</sup>. Additionally, physical activity might be associated with lowering the risk of AD based on a comprehensive meta-analysis<sup>13</sup>.

Though, other long-term RCTs had suggested that physical activity had no significant effects on cognitive decline or might not decrease the prevalence of all-cause dementia<sup>14,15</sup>. No difference in AD risk between physically active and inactive participants was observed in a meta-analysis including 19 prospective observational cohort studies<sup>16</sup>.

Understanding the relative contributions of physical activity to AD is crucial for designing suitable public health interventions to reduce the incidence of AD and decrease the socioeconomic burden. Using observational studies to determine the association might be challenging due to bias of measurement error, confounding, and reverse causation. Therefore, it remains to be elucidated whether the association between physical activity and AD is causal.

Mendelian randomization (MR) analysis using single nucleotide polymorphisms (SNPs) as instrumental variables could be conducted to explore whether the association between physical activity and AD is causal<sup>17</sup>. The study design is similar to RCT since genes are transferred from parents to offspring randomly<sup>17</sup>. All analyses reported in the present study are two-sample MR analyses since SNP-physical activity and SNP-AD associations were extracted from different rather than overlapping samples<sup>18</sup>. The main advantage of using summary-level data from genome-wide association studies (GWAS) in two-sample MR is the increased statistical power, particularly with testing effects on binary disease outcomes<sup>19</sup>.

In this study, we conducted the two-sample MR analysis using summary-level GWAS data on physical activity and AD to make a thorough inquiry about whether physical activity is causally linked with the risk of AD.

## Method

**Study design.** We conducted a two-sample MR analysis to explore the association between physical activity and AD. Genetic variants associated with physical activity were selected as genetic instruments for each exposure. The design of our study had three contents: (1) the acquisition of summary-level GWAS data for physical activity and AD; (2) the identification of genetic variants to serve as instrumental variables (IVs); and (3) the estimates of the effects of physical activity on AD.

**Genetic associations with physical activity.** Over 100,000 participants (mean age 57.52 years) from UK Biobank, a population-based prospective study<sup>20,21</sup>, who had provided valid email addresses were approached with a wear wrist-worn accelerometer, and 103,712 datasets were received with a median wear-time of 6.9 days (Supplementary Table 1)<sup>22</sup>. After quality control, a total of 91,105 participants of European descent remained for subsequent genome-wide association analysis and summary-level GWAS data used in the present analysis had been adjusted for BMI and sex as covariates<sup>23</sup>. A machine-learning model was used to predict four activity conditions including, sleep, sedentary behavior, walking, and moderate-intensity activity. For instance, sedentary behavior was defined as the energy expenditure score of Metabolic Equivalent of Task (MET) being less than 1.5<sup>24</sup>. More information on genotyping and imputation is shown in Supplementary Method 1. Exposures of interest in our MR analysis were overall activity, walking, sedentary behavior, and moderate-intensity activity. We selected SNPs associated with overall activity, sedentary behavior, and walking at the genome-wide significance level ( $p < 5 \times 10^{-8}$ ). Only one SNP (rs568974867) was related to moderate-intensity activity at a genome-wide significance level ( $p < 5 \times 10^{-8}$ ). However, rs568974867 was not present in summary-level GWAS data of AD. Therefore, we set the threshold at a suggestive significance level ( $p < 5 \times 10^{-6}$ ) for moderate-intensity activity. We eliminated IVs which were associated with other genetic variants ( $r^2$  threshold  $< 0.1$  and  $kb = 5,000$ ).

**Genetic associations with Alzheimer's disease.** Genetic associations with AD were obtained from the meta-analyses of GWAS on individuals of European ancestry ( $n_{\text{case}} = 21,982$  and  $n_{\text{control}} = 41,944$ ) contributed by the International Genomics of Alzheimer's Project (IGAP). Details of IGAP were published elsewhere<sup>25</sup>. In brief, in stage 1, IGAP used genotyped and imputed data on 11,480,632 SNPs to meta-analyses GWAS datasets consisting of four consortia: the Alzheimer Disease Genetics Consortium (ADGC); the European Alzheimer's disease Initiative (EADI); the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium (CHARGE); and the Genetic and Environmental Risk in AD Consortium Genetic and Environmental Risk in AD/Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium (GERAD/PERADES). Details of the above were shown in Table 1 and Supplementary Method 1. We inferred positive strand alleles by using allele frequencies to harmonize physical activity and AD data. We removed palindromic SNPs with intermediate allele frequency. More details on harmonization were shown in Supplementary Method 2. We did not use proxy SNPs in this research.

**Statistical power calculation.** Power calculations were conducted using an online web tool for the binary outcome available at <https://sb452.shinyapps.io/power/><sup>26</sup>. The statistical power for our MR relied on several parameters, including type I error of 1.25% after multiple testing corrections, the proportion of variance ( $R^2$ ) in the exposure explained by genetic instruments, the "true" causal effect of the physical activity on AD, and the ratio of cases to controls (1 to 1.908). More information is shown in Supplementary Method 3.

**Statistical analysis.** In the primary analysis, the fixed-effect inverse variance weighted (IVW) method was used to calculate the overall effects in the absence of heterogeneity<sup>27</sup>. However, if there is heterogeneity between the causal estimates of genetic variances, a random-effects model IVW would be conducted. Cochran's Q sta-

Alzheimer's disease	cases			controls		
	N	Percent female	Mean AAO (s.d)	N	Percent female	Mean AAE (s.d)
ADGC	14,428	59.3	71.1 (17.3)	14,562	59.3	76.2 (9.9)
CHARGE	2137	67.3	82.6 (12)	13,474	55.8	76.7 (8.2)
EADI	2240	65	75.4 (9.1)	6631	60.6	78.9 (7.0)
GERAD	3177	64	73.0 (0.2)	7277	51.8	51.0 (0.1)

**Table 1.** Description of datasets in genome-wide association studies. AAO, age at onset; AAE, age at examination; AGDC, the Alzheimer Disease Genetics Consortium; CHARGE, the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium; EADI, the European Alzheimer's disease Initiative; GERAD, the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium.

tistic would be computed during MR analysis to estimate whether heterogeneity exists or not. The amount of heterogeneity also was estimated by  $I^2$  statistic<sup>28</sup>.

**Sensitivity analysis.** MR was based on three key assumptions<sup>29</sup>. Sensitivity analyses were performed to examine the robustness of the MR results against any potential violation of the three key assumptions. Firstly, genetic variants must be associated with the exposure of interest. The strength of each instrument was measured using the  $F$  statistics calculated by the formula:  $F = R^2(N - K - 1)/K(1 - R^2)$ , where  $R^2$  represented the variance in exposures explained by the genetic variance,  $K$  represented the number of instruments, and  $N$  meant the sample size of the GWAS for the association of the SNP-activity factors<sup>30</sup>. The instrumental variables should be strongly associated with exposures and  $F$  statistics of all the genetic variants should be larger than the empirical threshold of 10<sup>31</sup>. Secondly, genetic variants must not be associated with potential confounders of the association between the exposure and the outcome, which meant no horizontal pleiotropy. MR-Egger regression, which provided a statistical test for directional pleiotropy according to Egger intercept, was used to assess the potential presence of horizontal pleiotropy<sup>32</sup>. Though, if the Instrument Strength Independent of Direct Effect (InSIDE) assumption was violated, this estimate might increase type I error rates. Therefore, the weighted median estimator was used for sensitivity analysis. This estimate was consistent if up to half of the genetic variants (or variants comprising 50% of the weight for a weighted analysis) were valid IVs<sup>33</sup>. Additionally, we also conducted MR pleiotropy residual sum and outlier (MR-PRESSO) analysis to explore the robustness of our results<sup>34</sup>. MR-PRESSO was a method to detect and correct outliers in IVW linear regression, which might reduce the heterogeneity in the estimates of causal effect by removing outliers. If the MR-PRESSO analysis indicated significant outliers exist, we would remove the outlier variants (with a  $P$  value less than the threshold in the MR-PRESSO outlier test) and conduct the MR analysis again. At the end of the analysis, a leave-one-out analysis was computed to test the robustness of MR estimates and whether anyone single variant was driving the causal association between exposure and outcome<sup>35</sup>. Thirdly, genetic variants should be associated with AD only through their effects on exposure, not through other pathways. The “negative control” population might be useful to evaluate the exclusion assumption<sup>36</sup>. Sensitivity analyses were not performed for walking as there were only 2 genetic variants.

Results were presented as log odds ratio for comparing with the rest ( $MET \leq 1$ ). The milli-gravity (mg) was used as a unit to report accelerometer-measured physical activity levels. An MR effect was considered significant at a Bonferroni-corrected  $P$ -value of  $0.05/4 = 0.0125$ . A  $P$ -value  $< 0.05$  but above the adjusted  $P$ -value was considered a suggestive association.

Strengthening the Reporting of Observational Study in Epidemiology (STROBE) MR checklist is provided in Supplementary Information<sup>37</sup>. All analyses were performed with Two-Sample MR package 0.5.5 and MR-PRESSO package 1.0 in R version 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria) (<http://www.r-project.org/>)<sup>34,38,39</sup>.

**Ethical issues.** Additional ethical approval was not required, since this study was based on MR analysis and depended on summary-level GWAS data rather than individual-level data, and the databases used in this analysis had been published or shared before.

## Results

In total, we used 33 SNPs in this multi-instrument MR analysis. Detailed information on genetic variants used in this analysis was shown in Supplementary Table 2.  $F$  statistics for instruments for overall activity, sedentary behavior, and walking were larger than the empirical threshold of 10, demonstrating the small possibility of weak instrumental variable bias (Table 2).  $F$  statistics for the instrument for moderate-intensity physical activity was only 2.19, thus suggesting a weak instrument. The statistical power was only 45.3% for sedentary behavior, and the statistical power was between 80% to 1 for another three exposures, including overall activity, walking, and moderate-intensity activity (Table 2).

For overall activity, the random-effect IVW analysis was conducted later due to the heterogeneity ( $Q_{p-val} = 0.024$ ). There was no significant association between genetically predicted overall activity and AD according to the result of IVW analysis (odds ratio (OR) = 0.62; 95% confidence interval (CI), 0.17–2.32,  $P = 0.4805$ ). However, the results computed by the weighted median method proved a suggestive association (OR = 0.35; 95% CI 0.13–0.92,  $P = 0.0362$ ). There was an absence of directional pleiotropy according to MR-Egger regression (intercept = 0.19, 95% CI 0.05–0.33,  $P = 0.234$ ). Besides, the MR estimate of ME-Egger was very different

Exposures	N	K	R <sup>2</sup> (%)	F statistics	Power <sup>a</sup> (%)
Overall activity	91,105	3	0.06	18.23	83.4
Sedentary behavior	91,105	3	0.06	18.23	45.3
Walking	91,105	2	0.06	27.35	100.0
Moderate-intensity activity ( $P < 5e^{-6}$ )	91,105	25	0.06	2.19	97.7

**Table 2.** Results of F statistic and statistical power for the Mendelian randomization analysis of genetic instruments and the risk of AD. N, sample size; K, number of genetic instruments; R<sup>2</sup>, the proportion of variance in the exposure explained by each genetic instrument. <sup>a</sup>Statistical power was calculated based on a sample size of 63,926 participants, type I error 1.25% and a ratio of case to control (1 to 1.908).

Exposure	Outcome	Outcome units	Method	SNPs	P	OR	95%CI	Q	Q_df	Q p-val	I <sup>2</sup>	MR Egger			MR-PRESSO				
												Intercept	95%CI	Intercept pval	Global test P	SNPs	OR	P	
Overall activity	Alzheimer's disease	Log odds ratio	MR Egger	3	0.2207	0.00	(0.00,0.14)	0.67	1	0.412	0.00	0.19	(0.05,0.33)	0.234	0.004	2	1.02 (0.16, 6.58)	0.996	
			IVW	3	0.4805	0.62	(0.17,2.32)	7.43	2	0.024	0.73								
			IVW (mre)	3	0.4805	0.62	(0.17,2.32)												
			Weighted median	3	0.0362	0.35	(0.13,0.92)												
Sedentary behavior	Alzheimer's disease	Log odds ratio	MR Egger	3	0.8147	1.40	(0.06,41.95)	0.01	1	0.922	0.00	-0.01	(-0.10,0.09)	0.916	0.185				
			IVW	3	0.4131	1.33	(0.68,2.60)	0.00	2	0.987	0.00								
			IVW (mre)	3	0.0000	1.33	(1.23,1.43)												
			Weighted median	3	0.5103	1.30	(0.60,2.80)												
Walking	Alzheimer's disease	Log odds ratio	IVW	2	0.0039	0.30	(0.13,0.68)	0.68	1	0.411	0.00	NA	NA	NA	NA				
			IVW (mre)	2	0.0004	0.30	(0.15,0.59)												
Moderate-intensity activity	Alzheimer's disease	Log odds ratio	MR Egger	25	0.8918	1.06	(0.49,2.28)	26.06	23	0.298	0.12	-0.01	(-0.03,0.01)	0.450	0.165				
			IVW	25	0.1371	0.80	(0.59,1.07)												
			IVW (mre)	25	0.1371	0.80	(0.59,2.07)	26.73	24	0.317	0.10								
			Weighted median	25	0.9210	0.98	(0.65,1.46)												

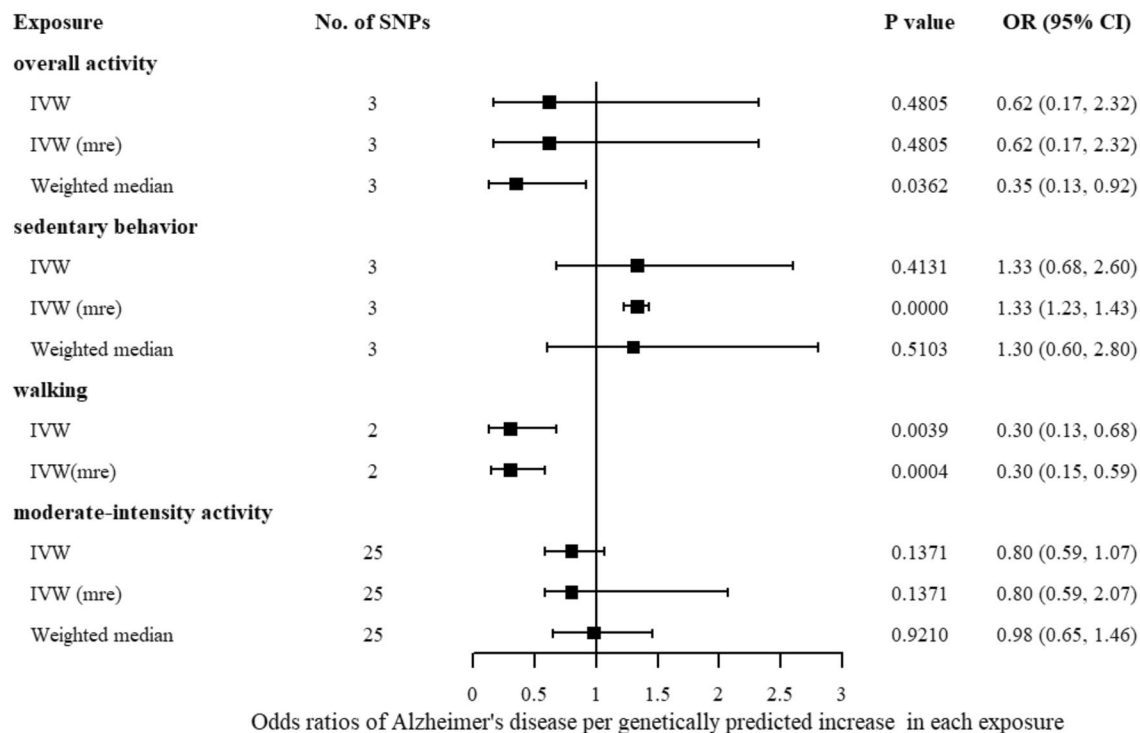
**Table 3.** MR Estimates for the effects of physical activity on AD risk in IGAP datasets. SNPs, single nucleotide polymorphisms; CI, confidence interval; OR, odds ratio; Q, Cochran Q statistics; IVW, inverse variance weighted; WM, weighted median; MR-Egger, Mendelian randomization-Egger; MR-PRESSO, MR pleiotropy residual sum and outlier.  $P_{het}$ : P-values of Cochran's Q test for heterogeneity were shown.  $P_{int}$ : P-values of MR-Egger regression test on the intercept.

from others (OR = 0.00; 95% CI 0.00–0.14,  $P = 0.2207$ ). The results of the null relationship were consistent before and after removing the outlier using MR-PRESSO (Table 3). Plots of leave-one-out analysis (Supplementary Fig. 1) demonstrated that there was potentially influential SNP driving the link between overall activity and AD. Therefore, we need to carefully interpret the result and draw a cautious conclusion. A Scatter plot was shown in Supplementary Fig. 2.

The result of random-effect IVW proved sedentary behavior might be associated with a higher risk of AD (OR = 1.33; 95% CI 1.23–1.43,  $P = 2.0188 \times 10^{-12}$ ) (Table 3, Fig. 1). Though fixed-effect IVW and weighted median analysis did not show evidence of a causal link. The MR-Egger intercept and MR-PRESSO did not show evidence of pleiotropy with sedentary behavior. Plots of leave-one-out analysis demonstrated no potentially influential SNPs driving the link between sedentary behavior and AD (Supplementary Fig. 1). A Scatter plot of SNP effects on exposure versus their effects on AD was shown in Supplementary Fig. 2.

The IVW regression showed that an increase in walking time might decrease the risk of AD (OR = 0.30, 95% CI 0.13–0.68,  $P = 0.0039$ ). The result was consistent with the random-effect IVW method (OR = 0.30; 95% CI 0.15–0.59,  $P = 0.0004$ ). There was still a significant association after multiple testing corrections. The results were shown in Fig. 1. There was no heterogeneity according to Cochran's Q test ( $P = 0.411$ ) (Table 3). However, no other sensitivity analyses were conducted due to the limited number of IVs.

As for moderate-intensity activity, using 25 genetic instruments, linear MR analysis demonstrated no causal effects of moderate-intensity activity on AD risk (OR = 0.80; 95% CI 0.59–1.07,  $P = 0.1371$ ), which was consistent with results in the weighted median method. Our analysis suggested no significant evidence of horizontal pleiotropy (as indicated by MR-Egger regression intercept close to zero and  $P = 0.450$ , Supplementary Fig. 2). The result of MR-PRESSO analysis indicated that no outliers existed (Table 3). The result of the leave-one-out



**Figure 1.** Forest plot for MR analysis of the causal effect of physical activity on Alzheimer's disease. SNPs, single nucleotide polymorphisms; No. of SNPs, number of SNP; OR, odds Ratio; 95% CI, 95% confidence interval; IVW, inverse variance weighted.

analysis (Supplementary Fig. 1) showed potentially influential SNP driving the link between moderate-intensity activity. Results were shown in Fig. 1.

## Discussion

In this two-sample MR analysis, we examined the effects of physical activity, including overall activity, sedentary behavior, walking, and moderate-intensity activity, on the risk of AD. Genetically predicted overall activity, sedentary behavior, and moderate-intensity activity were not associated with AD. Higher levels of walking might be beneficial for preventing AD.

For overall activity, the results of the IVW analysis (both fixed-effect and random-effect) suggested the absence of an association between exposure and AD. Nevertheless, the weighted median demonstrated a suggestive relationship. Besides, there was a huge difference between the effect sizes from IVW and ME-Egger analyses. While the MR-Egger estimate might be imprecise and influenced by outlying genetic variants, MR-PRESSO analysis did suggest outlier SNPs existed. There was still a null association between overall activity and AD when we excluded the potential outlier. Based on the different results drawn by sensitivity analyses, we should be cautious to conclude that overall activity does not influence the risk of AD.

There was no heterogeneity existing for sedentary behavior according to Cochran's Q statistic. We mainly focused on the MR estimate of the fixed-effect IVW method, which revealed a positive relationship between sedentary behavior and AD. Higher levels of sedentary behavior seemed related to a higher risk of AD, which indicated a deleterious effect of sedentary behavior on AD. However, consistent with the results of sensitivity analyses, the association was not significant. Sedentary behavior was defined as an activity that had a MET energy expenditure score of  $\leq 1.5$  and usually occurs at sitting, lying, and reclining status<sup>24</sup>. Nonetheless, other behaviors such as driving (MET < 2.5) and some instances of non-desk work were categorized as sedentary behavior in the original GWAS analysis which might be the explanation for the absence of association in our MR analysis<sup>23</sup>.

We found that the risk of AD decreases by approximately 30% for each 1-SD increase in genetically predicted walking levels. Although we only used 2 SNPs as instrumental variables for walking, these 2 SNPs explained a 0.12% variance in walking and they fulfilled the criterion of not being weak instrumental variables ( $F$  statistics > 10)<sup>31</sup>. Besides, the statistical power of the MR analysis for walking was sufficient. Therefore, low statistical power and weak instruments are unlikely to explain our results. However, sensitivity analysis could not be conducted to verify the independence and exclusion restriction assumption which might influence our results due to the limited number of SNPs.

The results of MR analysis did not demonstrate a relationship between moderate-intensity activity and AD. The null association might be influenced by a bias of weak instruments ( $F$  statistic = 2.19) due to violations of instrumental assumption. Secondly, even when using large samples ( $n = 91,105$ ), results using weak instruments might be biased towards the null in the two-sample MR analysis<sup>36</sup>. That was also a well-powered MR result since moderate-intensity activity showed significant power (> 80%).



Our results were consistent with previous studies which suggested that physical activity was associated with a decreased AD risk<sup>13,40–42</sup>. The result from the previous analysis indicated exercise was related to a 28% decreased risk of AD was similar to our results<sup>13</sup>. The meta-analysis, which included 3345 participants had point estimates that suggested a protective effect of physical activity. The result of this meta-analysis finding an OR for the development of AD was 0.60 (95% CI 0.51–0.71) compared to inactive individuals<sup>40</sup>. Our study confirmed the finding of a clinical trial that suggested physical activity was associated with a reduced risk of AD (OR = 0.48; 95% CI 0.17–0.85)<sup>43</sup>.

The mechanism might be explained as follows: both animal and human research had shown that exercise resulted in beneficial changes to brain structure and function accompanying cognitive function<sup>44,45</sup>. Physical activity might modulate amyloid  $\beta$  ( $A\beta$ ) turnover, inflammation, the synthesis and release of neurotrophins, and cerebral blood flow (CBF)<sup>46,47</sup>. Besides, activity can influence cognition through different pathophysiological processes such as the immune system, endothelial function, cerebrovascular insufficiency, oxidative stress, and neurotoxicity<sup>48</sup>.

It was worth mentioning that our results indicated walking rather than moderate-intensity activity is related to the reduced AD risk. The possible explanation might be the small sample size and the young age of subjects from previous research<sup>49</sup>. For AD, symptoms usually occurred after age 60<sup>1</sup>. Secondly, the beneficial effect of an exercise intervention on cognitive function was driven by the types of activity<sup>50</sup>. Previous research might mix the effect of moderate-intensity activity with light and vigorous activity making it difficult to isolate the effects of moderate-intensity activity<sup>51,52</sup>. Further, moderate-intensity activity was defined by energy expenditure score in our MR analysis. The classification criteria for moderate-intensity activity in the previous research were inconsistent with the present MR analysis<sup>45</sup>. Apart from that, physical activity intervention in other research included not only aerobic exercise but also strength, and balance exercise<sup>42,53,54</sup>. The beneficial effect of moderate-intensity activity on cognitive function might pair with other elements. Our conclusion was supported by previous studies, which indicated that walking might enhance cognitive performance and suggested walking as little as 1.5 h weekly was beneficial for cognition in healthy elderly<sup>55</sup>. Furthermore, walking was not only positively correlated with emotional health, and memory, but also was beneficial for reducing depression scores, anxiety, and lower psychological stress<sup>56–59</sup>. Walking could even stabilize the progressive cognitive dysfunctions in AD patients<sup>60,61</sup>. Thus, we suggested that walking might be an effective strategy for AD prevention.

The results contrasted the finding of other studies that reported no difference in AD risk between physically active and inactive participants. The reasons for the difference between our MR analysis result and other research might be explained as follows. Firstly, previous studies could not exclude the potential effects caused by some biases, including unmeasurable confounders, reverse causality, and recall biases. Besides, for different randomized controlled trials, the intervention of physical activity or aerobic exercise and follow-up periods might be quite different. Moreover, there might be heterogeneity among these studies, especially in the different methods of measurement used for cognitive function. Secondly, most previous studies focused on the outcome of the cognitive improvement in people with AD or mild cognitive impairment, not the risk of incidence<sup>62,63</sup>. Thirdly, there were limited subjects in previous RCT studies<sup>64,65</sup>. Therefore, it would be hard to compare the effects of physical activity on AD risk.

Our results were contradictory to a previous MR study which suggested that physical activity did not affect the risk of development of AD<sup>66</sup>. The possible reason might be as follows: previous MR studies used moderate-to-vigorous physical activity (MET  $\geq$  6) as exposure, which might be different from the effects of daily activity. Physical activities were divided into more detailed categories, including sedentary behavior, walking, and moderate-intensity activity in the present MR analysis.

This study had several strengths. Firstly, daily physical activities including sedentary behavior and walking were recorded by activity trackers and were more precise than self-report data. Genetic variants identified in GWAS analysis based on self-reported data were at least partly influenced by reporting bias<sup>67,68</sup>. Secondly, based on the design of the present two-sample MR analysis, there might be less bias from confounding which would have been the case if both physical activity and AD GWAS were performed in the same sample. Thirdly, the influence of confounding factors on the studied association was reduced due to the strict selection of IVs. Fourthly, the previous MR analysis used average accelerations and the fraction of accelerations > 425 milligravities as exposure measures, the latter meaning vigorous physical activity (METs  $\geq$  6) only<sup>66</sup>. More physical activity statuses were analyzed in our study. Besides, various methods, including IVW, MR-Egger, weighted-median, and MR-PRESSO were performed in this MR analysis to robust our conclusion. Finally, population stratification could be decreased in the present study as the AD dataset was restricted to participants of European ancestry.

Our study had certain limitations. Firstly, the GWAS of exposure consisted of participants aged 55–74 years with higher socioeconomic and physical health status, which might affect the results<sup>23</sup>. Besides, our results assumed that the sample used to define the IVs for physical activity and the samples from the IGAP consortium used to estimate the genetic association with AD are both from European ancestry. Hence, our results might not be suited to all populations. Thirdly, we cannot exclude the possibility of the sample overlap since both physical activity and AD used individuals from Europe. Therefore, that might generate bias in our analysis. A further limitation is that the number of genetic variants used for walking was limited for sensitivity analysis. Another limitation is that the studies included in IGAP used different diagnostic criteria for AD, but all cases met standard criteria for possible, probable, or definite AD. Hence certain misclassification might be inevitable. Sixthly, these exposure datasets recorded by wrist-worn accelerometer co-occur and interact across daily life. Multivariable MR analysis could be conducted to explore the relationship between physical activity and AD to make the results more precise<sup>69</sup>. Seventhly, the association between moderate-intensity activity and AD might be biased by weak instruments ( $F$  statistic = 2.19). Besides, the final MR assumption (the exclusion assumption) had not been evaluated. We did not search through the PhenoScanner to screen genetic variants which might be related to confounding factors<sup>70</sup>. We just conducted MR-PRESSO analysis to detect and remove outliers. We

additionally calculated statistical power for our MR analysis. We had sufficient power to detect the effects of exposures on AD, except for sedentary behavior. Therefore, we should draw our conclusion with caution. Finally, the two-sample MR analysis assumed a linear relationship between physical activity and AD. In this research, we could not investigate the nonlinear effects of physical activity. Nevertheless, there could be a nonlinear relationship, such as U- or J- shaped association, between physical activity and AD. Therefore, further research could be needed to detect the causal factors of AD.

In summary, this study provided evidence that genetically predicted walking might associate with a reduced risk of AD. The elderly could use walking as the main way of commuting in daily life to reduce the risk of AD. Further research into the causal association between risk factors and AD could help to explore the real relationship and provide more measures to reduce AD risk. Mechanisms between physical activity and AD might need more investigation. More high-quality RCTs with fewer methodological issues and heterogeneity also are needed to back up our present findings.

### Data availability

The datasets analyzed during the current study can be downloaded from [<https://doi.org/10.5287/bodleian:yjp6zZmdj>] and GWAS Catalog [[http://ftp.ebi.ac.uk/pub/databases/gwas/summary\\_statistics/GCST007001-GCST08000/GCST007511/](http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST007001-GCST08000/GCST007511/)].

Received: 24 December 2021; Accepted: 21 July 2022

Published online: 28 July 2022

### References

1. Watwood, C. *Alzheimer's Disease*, 39. (DLPS Faculty Publications, 2011).
2. Ballard, C. *et al.* Alzheimer's disease. *Lancet* **377**, 1019–1031 (2011).
3. Lane, C. A., Hardy, J. & Schott, J. M. Alzheimer's disease. *Eur. J. Neurol.* **25**, 59–70 (2018).
4. Niu, H., Álvarez-Álvarez, I., Guillén-Grima, F. & Aguinaga-Ontoso, I. Prevalence and incidence of Alzheimer's disease in Europe: A meta-analysis. *Neurologia* **32**, 523–532 (2017).
5. Rajan, K. B., Weuve, J., Barnes, L. L., Wilson, R. S. & Evans, D. A. Prevalence and incidence of clinically diagnosed Alzheimer's disease dementia from 1994 to 2012 in a population study. *Alzheimers Dement* **15**, 1–7 (2019).
6. Alzheimer's Association. 2021 Alzheimer's disease facts and figures. *Alzheimers Dement: J. Alzheimer's Assoc.* **17**, 327–406 (2021).
7. Norton, S., Matthews, F. E., Barnes, D. E., Yaffe, K. & Brayne, C. Potential for primary prevention of Alzheimer's disease: An analysis of population-based data. *Lancet Neurol.* **13**, 788–794 (2014).
8. De la Rosa, A. *et al.* Long-term exercise training improves memory in middle-aged men and modulates peripheral levels of BDNF and Cathepsin B. *Sci. Rep.* **9**, 3337 (2019).
9. Ngandu, T. *et al.* A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): A randomised controlled trial. *Lancet* **385**, 2255–2263 (2015).
10. Lautenschlager, N. T. *et al.* Effect of physical activity on cognitive function in older adults at risk for Alzheimer disease: A randomized trial. *JAMA* **300**, 1027–1037 (2008).
11. Jia, R. X., Liang, J. H., Xu, Y. & Wang, Y. Q. Effects of physical activity and exercise on the cognitive function of patients with Alzheimer disease: A meta-analysis. *BMC Geriatr.* **19**, 181 (2019).
12. Du, Z. *et al.* Physical activity can improve cognition in patients with Alzheimer's disease: A systematic review and meta-analysis of randomized controlled trials. *Clin. Interv. Aging* **13**, 1593–1603 (2018).
13. Xu, W. *et al.* Meta-analysis of modifiable risk factors for Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry* **86**, 1299–1306 (2015).
14. Andrieu, S. *et al.* Effect of long-term omega 3 polyunsaturated fatty acid supplementation with or without multidomain intervention on cognitive function in elderly adults with memory complaints (MAPT): A randomised, placebo-controlled trial. *Lancet Neurol.* **16**, 377–389 (2017).
15. Moll van Charante, E. P. *et al.* Effectiveness of a 6-year multidomain vascular care intervention to prevent dementia (preDIVA): A cluster-randomised controlled trial. *Lancet* **388**, 797–805 (2016).
16. Kivimäki, M. *et al.* Physical inactivity, cardiometabolic disease, and risk of dementia: An individual-participant meta-analysis. *BMJ* **365**, l1495 (2019).
17. Smith, G. D. & Ebrahim, S. “Mendelian randomization”: Can genetic epidemiology contribute to understanding environmental determinants of disease?. *Int. J. Epidemiol.* **32**, 1–22 (2003).
18. Burgess, S., Davies, N. M. & Thompson, S. G. Bias due to participant overlap in two-sample Mendelian randomization. *Genet. Epidemiol.* **40**, 597–608 (2016).
19. Burgess, S., Scott, R. A., Timpson, N. J., Davey Smith, G. & Thompson, S. G. Using published data in Mendelian randomization: A blueprint for efficient identification of causal risk factors. *Eur. J. Epidemiol.* **30**, 543–552 (2015).
20. Sudlow, C. *et al.* UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
21. U. Biobank. *Repeat Assessment: Participant Characteristics of responders vs. non-responders* (2014).
22. Doherty, A. *et al.* Large scale population assessment of physical activity using wrist worn accelerometers: The UK Biobank Study. *PLoS ONE* **12**, e0169649 (2017).
23. Doherty, A. *et al.* GWAS identifies 14 loci for device-measured physical activity and sleep duration. *Nat. Commun.* **9**, 5257 (2018).
24. Tremblay, M. S. *et al.* Sedentary behavior research network (SBRN)—Terminology Consensus Project process and outcome. *Int. J. Behav. Nutr. Phys. Act* **14**, 75 (2017).
25. Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019).
26. Burgess, S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int. J. Epidemiol.* **43**, 922–929 (2014).
27. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* **37**, 658–665 (2013).
28. Higgins, J. P., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *BMJ* **327**, 557–560 (2003).
29. Greenland, S. An introduction to instrumental variables for epidemiologists. *Int. J. Epidemiol.* **29**, 722–729 (2000).
30. Burgess, S. & Thompson, S. G. Avoiding bias from weak instruments in Mendelian randomization studies. *Int. J. Epidemiol.* **40**, 755–764 (2011).
31. Staiger, D. O. & Stock, J. H. Instrumental variables regression with weak instruments. (1994).

32. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).
33. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* **40**, 304–314 (2016).
34. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* **50**, 693–698 (2018).
35. Burgess, S., Bowden, J., Fall, T., Ingelsson, E. & Thompson, S. G. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. *Epidemiology* **28**, 30–42 (2017).
36. Davies, N. M., Holmes, M. V. & Davey Smith, G. Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ* **362**, k601 (2018).
37. Skrivankova, V. W. *et al.* Strengthening the reporting of observational studies in epidemiology using mendelian randomization: The STROBE-MR statement. *JAMA* **326**, 1614–1621 (2021).
38. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **7**, (2018).
39. Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* **13**, e1007081 (2017).
40. Santos-Lozano, A. *et al.* Physical activity and alzheimer disease: A protective association. *Mayo Clin. Proc.* **91**, 999–1020 (2016).
41. Iso-Markku, P. *et al.* Physical activity as a protective factor for dementia and Alzheimer's disease: Systematic review, meta-analysis and quality assessment of cohort and case-control studies. *Br. J. Sports Med.* **56**, 701–709 (2022).
42. Sattler, C., Erickson, K. I., Toro, P. & Schröder, J. Physical fitness as a protective factor for cognitive impairment in a prospective population-based study in Germany. *J. Alzheimers Dis.* **26**, 709–718 (2011).
43. Rovio, S. *et al.* Leisure-time physical activity at midlife and the risk of dementia and Alzheimer's disease. *Lancet Neurol.* **4**, 705–711 (2005).
44. Parachikova, A., Nichol, K. E. & Cotman, C. W. Short-term exercise in aged Tg2576 mice alters neuroinflammation and improves cognition. *Neurobiol. Dis.* **30**, 121–129 (2008).
45. Marin Bosch, B. *et al.* A single session of moderate intensity exercise influences memory, endocannabinoids and brain derived neurotrophic factor levels in men. *Sci. Rep.* **11**, 14371 (2021).
46. De la Rosa, A. *et al.* Physical exercise in the prevention and treatment of Alzheimer's disease. *J. Sport Health Sci.* **9**, 394–404 (2020).
47. Müller, S. *et al.* Relationship between physical activity, cognition, and Alzheimer pathology in autosomal dominant Alzheimer's disease. *Alzheimers Dement.* **14**, 1427–1437 (2018).
48. López-Ortiz, S. *et al.* Physical exercise and Alzheimer's disease: Effects on pathophysiological molecular pathways of the disease. *Int. J. Mol. Sci.* **22**, 2897 (2021).
49. Nakagawa, T. *et al.* Regular moderate- to vigorous-intensity physical activity rather than walking is associated with enhanced cognitive functions and mental health in young adults. *Int. J. Environ. Res. Public Health* **17**, 614 (2020).
50. Groot, C. *et al.* The effect of physical activity on cognitive function in patients with dementia: A meta-analysis of randomized control trials. *Ageing Res. Rev.* **25**, 13–23 (2016).
51. Scarmeas, N. *et al.* Physical activity, diet, and risk of Alzheimer disease. *JAMA* **302**, 627–637 (2009).
52. Buchman, A. S. *et al.* Total daily physical activity and the risk of AD and cognitive decline in older adults. *Neurology* **78**, 1323–1329 (2012).
53. Bohannon, R. W. & Crouch, R. Minimal clinically important difference for change in 6-minute walk test distance of adults with pathology: A systematic review. *J. Eval. Clin. Pract.* **23**, 377–381 (2017).
54. Lamb, S. E. *et al.* Dementia and physical activity (DAPA) trial of moderate to high intensity exercise training for people with dementia: Randomised controlled trial. *BMJ* **361**, k1675 (2018).
55. Weuve, J. *et al.* Physical activity, including walking, and cognitive function in older women. *JAMA* **292**, 1454–1461 (2004).
56. Zhu, Z. *et al.* Exploring the relationship between walking and emotional health in China. *Int. J. Environ. Res. Public Health* **17**, 8804 (2020).
57. Demnitz, N. *et al.* Cognition and mobility show a global association in middle- and late-adulthood: Analyses from the Canadian Longitudinal Study on Aging. *Gait Posture* **64**, 238–243 (2018).
58. Kelly, P. *et al.* Walking on sunshine: Scoping review of the evidence for walking and mental health. *Br. J. Sports Med.* **52**, 800–806 (2018).
59. Vogt, T., Schneider, S., Brummer, V. & Strüder, H. K. Frontal EEG asymmetry: The effects of sustained walking in the elderly. *Neurosci. Lett.* **485**, 134–137 (2010).
60. Venturelli, M., Scarsini, R. & Schena, F. Six-month walking program changes cognitive and ADL performance in patients with Alzheimer. *Am. J. Alzheimer's Dis. Other Dementias* **26**, 381–388 (2011).
61. Winchester, J. *et al.* Walking stabilizes cognitive functioning in Alzheimer's disease (AD) across one year. *Arch. Gerontol. Geriatr.* **56**, 96–103 (2013).
62. Morris, J. K. *et al.* Aerobic exercise for Alzheimer's disease: A randomized controlled pilot trial. *PLoS ONE* **12**, e0170547 (2017).
63. Fonte, C. *et al.* Comparison between physical and cognitive treatment in patients with MCI and Alzheimer's disease. *Ageing* **11**, 3138–3155 (2019).
64. Yu, F. *et al.* Cognitive effects of aerobic exercise in Alzheimer's disease: A pilot randomized controlled trial. *J. Alzheimer's Dis.* **80**, 233–244 (2021).
65. Vidoni, E. D. *et al.* Effect of aerobic exercise on amyloid accumulation in preclinical Alzheimer's: A 1-year randomized controlled trial. *PLoS ONE* **16**, e0244893 (2021).
66. Baumeister, S. E. *et al.* Physical activity and risk of Alzheimer disease: A 2-sample mendelian randomization study. *Neurology* **95**, e1897–e1905 (2020).
67. 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**, 1161–1176 (2018).
68. Folley, S., Zhou, A. & Hyppönen, E. Information bias in measures of self-reported physical activity. *Int. J. Obes.* **42**, 2062–2063 (2018).
69. Burgess, S. & Thompson, S. G. Multivariable Mendelian randomization: The use of pleiotropic genetic variants to estimate causal effects. *Am. J. Epidemiol.* **181**, 251–260 (2015).
70. Staley, J. R. *et al.* PhenoScanner: A database of human genotype-phenotype associations. *Bioinformatics* **32**, 3207–3209 (2016).

## Acknowledgements

We thank the International Genomics of Alzheimer's Project (IGAP) for providing summary results data for these analyses. The investigators within IGAP contributed to the design and implementation of IGAP and/or provided data but did not participate in the analysis or writing of this report. IGAP was made possible by the generous participation of the control subjects, the patients, and their families. The i-Select chips were funded by the French National Foundation on Alzheimer's disease and related disorders. EADI was supported by the LABEX (laboratory of excellence program investment for the future) DISTALZ grant, Inserm, Institut Pasteur de



Lille, Université de Lille 2, and the Lille University Hospital. GERAD/PERADES was supported by the Medical Research Council (Grant n° 503480), Alzheimer's Research UK (Grant n° 503176), the Wellcome Trust (Grant n° 082604/2/07/Z), and the German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) grant n° 01GI0102, 01GI0711, 01GI0420. CHARGE was partly supported by the NIH/NIA grant R01 AG033193 and the NIA AG081220 and AGES contract N01-AG-12100, the NHLBI grant R01 HL105756, the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA grants: U01 AG032984, U24 AG021886, U01 AG016976, and the Alzheimer's Association grant ADGC-10-196728.

### Author contributions

B.Z.: software, investigation, writing—original draft. X.H.: formal analysis. X.W.: methodology. X.C.: writing—review and editing. C.Z., G.W.: data curation. W.S.: software, visualization. W.Z.: supervision, writing—review and editing.

### 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/s41598-022-17207-x>.

**Correspondence** and requests for materials should be addressed to W.Z.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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) 2022