Original Article

Genetically determined low income modifies Alzheimer's disease risk

Rong-Ze Wang, Yu-Xiang Yang, Hong-Qi Li, Xue-Ning Shen, Shi-Dong Chen, Mei Cui, Yi Wang, Qiang Dong, Jin-Tai Yu

Department of Neurology and Institute of Neurology, Huashan Hospital, State Key Laboratory of Medical Neurobiology and MOE Frontiers Center for Brain Science, Shanghai Medical College, Fudan University, Shanghai, China

Contributions: (I) Conception and design: JT Yu, Q Dong, Y Wang, M Cui; (II) Administrative support: JT Yu, Q Dong, Y Wang, M Cui; (III) Provision of study materials or patients: JT Yu, Q Dong, Y Wang, M Cui; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: RZ Wang, YX Yang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Prof. Jin-Tai Yu, MD, PhD. Department of Neurology and Institute of Neurology, Huashan Hospital, Shanghai Medical College, Fudan University, 12th Wulumuqi Zhong Road, Shanghai 200040, China. Email: jintai_yu@fudan.edu.cn.

Background: Socioeconomic status (SES) is considered to be associated with the prevalence of Alzheimer's disease (AD). However, the causal association remain unclear. Here, we determining whether income has a causal protective effect on the risk of developing AD using Mendelian randomization (MR).

Methods: Single-nucleotide polymorphisms (SNPs) that are strongly associated with household income levels ($P<5\times10^{-8}$) from the UK Biobank (UKB) (n=286,301) were selected as instrumental variables for this study. Confounding instruments were removed through data set browsing. Selected SNPs were then harmonized with results from an AD genome-wide meta-analysis (71,880 cases, 383,378 controls) including both case-control and proxy cases. The analysis was conducted using MR methods, and multiple sensitivity analyses were applied for testing of potential bias.

Results: After confounding instrument removal and clumping, 9 SNPs associated with household income level identified by the UKB were left for the MR analysis. Our results demonstrated that higher household income level was causally related with a lower risk of AD (odds ratio 0.78, 95% confidence interval: 0.69–0.89; P<0.001). Multiple sensitivity analyses suggested no obvious evidence for heterogeneity or pleiotropy of the results.

Conclusions: Under MR assumptions, our results suggest robust evidence of a causal association between household income and AD risk, which may provide potential prevention strategies for this devastating disease.

Keywords: Alzheimer's disease (AD); Mendelian randomization (MR); income; causality

Submitted Jan 21, 2021. Accepted for publication Apr 18, 2021.
doi: 10.21037/atm-21-344
View this article at: https://dx.doi.org/10.21037/atm-21-344

Introduction

Alzheimer's disease (AD) is the most common form of dementia, and it has major implications for worldwide health and social services (1). No effective intervention is currently available for this devastating disease (2). The identification of modifiable risk factors that prevent the disease may be one possible avenue to relieve the burden of AD. Socioeconomic status (SES) corresponds to a complex bundle of social and economic factors including education, income, and occupational status (3) and is considered to be associated with the prevalence of AD (4,5). Studies have demonstrated that higher education level, increasing income, and higher occupational status are associated with decreasing risk of AD (6-9). As an important part of SES, balancing income inequality is expected to reduce the prevalence rate of AD. However, previous

observational studies may be biased due to measurement error, confounding, and reverse causation and thus fail to demonstrate the causal association between income and AD (10,11). Under these circumstances, Mendelian randomization (MR) can be used to determine a causal effect.

MR uses genetic variants that are associated with risk factors as instruments for causality analysis. The results can determine whether an observational association between a risk factor and an outcome is consistent with a causal effect (11). The genetic variants must not be associated with any confounding factor or be directly associated with the outcome (12). MR has already demonstrated causal effects of potential risk factors in the context of multiple diseases and has revealed effects of body mass index, education attainment, and alcohol assumption on AD (13-15). Compared to observational methods, MR is less biased, since it uses single-nucleotide polymorphisms (SNPs) that are randomly allocated at conception as instruments to conduct instrumental variable analysis (16). In this study, we used SNPs selected from a recent genome-wide association study (GWAS) of household income as instruments to perform a 2-sample univariable MR analysis and identify the causal effect of income on the risk of developing AD. We present the following article in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting checklist (available at https://dx.doi.org/10.21037/atm-21-344).

Methods

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Income data and instruments

SNPs associated with income were obtained from a recently published GWAS on household income in a sample of 286,301 UK Biobank (UKB) participants with White British ancestry (17). There were 138,425 male participants included, and the mean age was 56.5 (SD =8.0 years). Total household income before tax was originally collected using a 5-point scale through the UKB, where 1 denoted less than £18,000, 2 £18,000–£29,999, 3 £30,000–£51,999, 4 £52,000–£100,000, and 5 greater than £100,000. The level of household income was subjected to a regression analysis using income as the outcome. The 40 genetic principal components, genotyping array, batch, age, and sex were used as predictors. Among 30 income-related loci identified in this study, 18 SNPs at a genome-wide significance level ($P<5\times10^{-8}$) were extracted for further analysis.

AD

AD data were obtained from a recently published genome-wide meta-analysis study, which used both traditional clinically diagnosed AD and AD-by-proxy participants (based on parental diagnoses reported by children of parents) from the UKB (18). The AD-by-proxy phenotype showed a strong genetic correlation with AD (r_g=0.81). A sample of 71,880 cases consisting of both AD and AD-by-proxy participants and 383,378 controls was included in the meta-analysis, and the test statistics per SNP per GWAS were used to calculate the P value. The results were used as the outcome data set in this MR analysis.

Statistical analysis

All statistical analyses were conducted using R software (version 3.6.3, R Foundation for Statistical Computing), the MR analysis was performed using the "TwoSampleMR" package (version 0.5.4) (19).

Using the GWAS catalog (20) and the PhenoScanner GWAS database (21,22), we excluded those SNPs that were associated with confounding factors, like educational attainment, or that were directly associated with AD. After removal of SNPs that reached genome-wide significance ($P<5\times10^{-8}$) based on an assessment of confounding factors, 9 SNPs associated with household income level remained. The strength of the instruments was estimated on the basis of the *F* statistic, with an *F* statistic >10 commonly cited as a value that can avoid bias in MR analysis (12) and a lower mean *F* statistic corresponding to a greater bias (23). The selected SNPs in the exposure-outcome data set as mentioned above were then harmonized. The coefficients of SNP-exposure and SNP-outcome were combined in a fixed-effects meta-analysis, and the inverse-variance-weighted (IVW) approach was used to detect the overall estimated causal effect (24). This method is equal to a weighted regression of SNP-outcome coefficients on the SNP-exposure coefficients, and the intercept is assumed to be zero. The result is unbiased if there is no horizontal pleiotropy or if the horizontal pleiotropy is balanced. Power calculations were performed using the mRnd power calculation tool to test whether the sample size was adequate

Table 1 Instrument single-nucleotide polymorphisms for Mendelian randomization analysis and their effects on household income

SNP	Chromosome	Effect allele	Other allele	β value	Standard error	P value
rs159365	5	G	A	0.016	0.002	9.04E-15
rs2332719	3	G	A	−0.013	0.002	1.16E-10
rs537160	6	G	A	−0.013	0.002	2.34E-10
rs12119149	1	T	C	0.013	0.002	5.24E-10
rs12954483	18	G	A	−0.013	0.002	5.80E-10
rs7597007	2	T	G	−0.012	0.002	1.46E-09
rs37976	7	T	C	0.012	0.002	3.14E-09
rs306755	20	C	T	0.012	0.002	3.82E-09
rs2563332	5	T	C	−0.012	0.002	9.90E-09

SNP, single-nucleotide polymorphism.

to reject the null hypothesis (25). Other MR methods including MR-Egger, weighted median, simple mode-based, and weighted mode-based were conducted to control potential bias attributable to violation of the assumptions of the IVW MR analysis. These methods are more robust for horizontal pleiotropy at the cost of reduced statistical power (26-28).

Pleiotropic genetic variants may affect the outcome independently and bias the result. Leave-one-out sensitivity analysis was used to explore SNP heterogeneity, where the random-effects IVW was performed again, but each SNP was left out in turn. This analysis determined if 1 or more SNPs were invalid instrumental variables or if the causal association was driven by a single SNP.

Furthermore, the Cochran's *Q* test, MR-Egger intercept test, MR Pleiotropy RESidual Sum and Outlier (MR-PRESSO) global test were used to investigate potential pleiotropy. Cochran's *Q* was employed as a test for the heterogeneity of the MR estimate, and a Cochran's *Q* P value<0.05 was considered an indicator of heterogeneity (29). The MR-Egger method assumes no intercept term in the model, and the intercept is considered to be zero. When the P value is larger than 0.05 in the MR-Egger intercept test, it can provide evidence for absence of pleiotropic bias (26). MR Pleiotropy RESidual Sum and Outlier (MR-PRESSO) compares the difference between the residuals for each genetic variant in the variable, and thus pleiotropic effects can be detected and outliers identified with this method. MR-PRESSO re-analyzes the association without the outliers, correcting for possible pleiotropic effects (30).

Results

Causal effect of household income on AD risk

After instrument selection and SNP exposure-outcome data harmonization were performed as described above, 9 SNPs were left for MR analysis (*Table 1*). All SNPs passed the clumping with a clumping window of 10,000 kb and a linkage disequilibrium (LD) r^2>0.001. All SNPs had an *F* statistic value greater than 10, and the mean *F* statistic value was 156.54, which demonstrated a low risk of instrument bias.

Using an IVW approach, we found that the AD and AD-by-proxy cohort demonstrated a causal association between genetically predicted household income and AD: the odds ratio (OR) was 0.78 with a 95% confidence interval (CI) between 0.69 and 0.89 (P<0.001; *Figure 1*). Moreover, the effect of household income on the risk of developing AD in weighted median (OR 0.80, 95% CI, 0.68–0.93; P=0.005), simple mode-based (OR 0.71, 95% CI, 0.56–0.91; P=0.027), and weighted mode-based analyses (OR 0.76, 95% CI, 0.59–0.96; P=0.053) demonstrated a similar causal effect as that found with IVW methods, while the result of MR-Egger analysis was in opposition to other methods, maintaining a positive coefficient value (*Figure 1*).

Sensitivity analysis

The results of leave-one-out sensitivity analysis are displayed in a forest plot (*Figure 2*). All error bars of the meta-analysis with 1 SNP removed are on the left side of the zero line, demonstrating there was no single SNP driving the causal link, and the conclusion was stable. The Cochran's

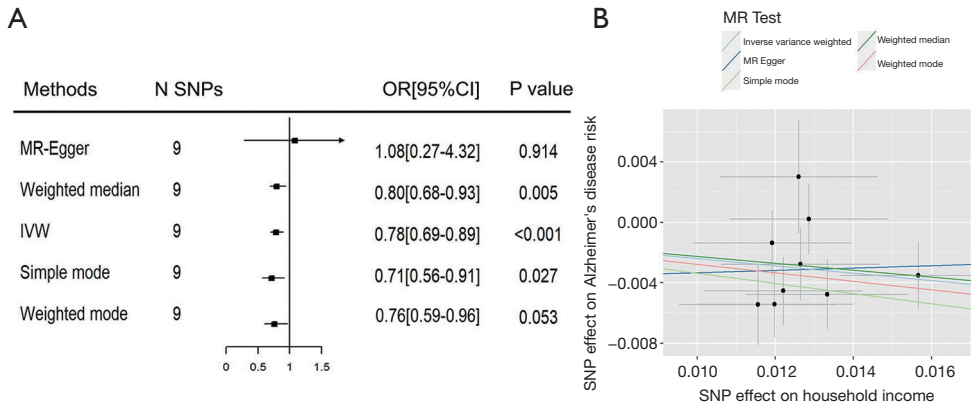

Figure 1 Summary Mendelian randomization (MR) estimates of association between household income and Alzheimer's disease (AD). (A) The causal association between household income and AD was derived from the inverse-variance-weighted, MR-Egger, weighted median, simple mode-based, and weighted mode-based methods. Inverse-variance-weighted, number, odds ratio, and confidence interval. (B) Scatterplot demonstrated single-nucleotide polymorphisms' potential effects of household income on AD. The slope of each line corresponds to the estimated MR effect per method, and the black dots correspond to the intersection points of the effects of exposure and outcome.

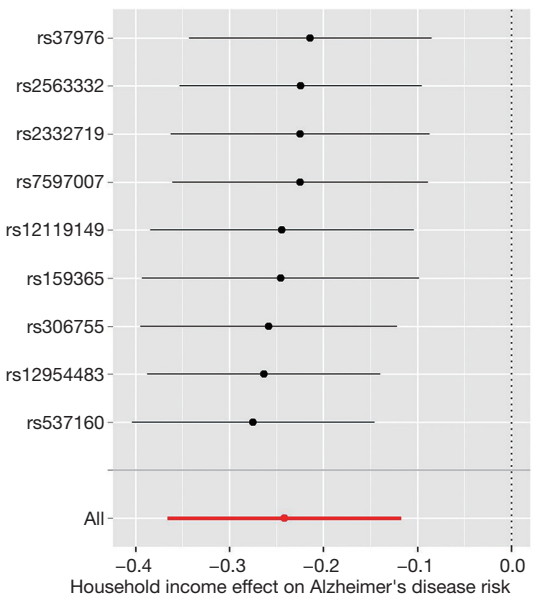

Figure 2 Leave-one-out sensitivity analysis. Forest plot of leave-one-out analysis to investigate whether the causal association was driven by a unique single nucleotide polymorphism. All error bars of the meta-analysis with 1 single-nucleotide polymorphism (SNP) removed are on the left side of the zero line.

Q P value was equal to 0.663, suggesting no remarkable heterogeneity of instruments. The MR-Egger intercept test P value was 0.321, and MR-PRESSO identified no outliers with a global test P value of 0.484. These tests provided evidence for the absence of heterogeneity or pleiotropy. The power was 1 with a sample size of 455,258, and the proportion of cases was equal to 0.158.

Discussion

In this study, we used MR analysis with selected genetic instruments to demonstrate that increasing household income had a protective role on AD risk. The results were robust to heterogeneity and pleiotropy. The causal association estimates using weighted median, simple mode-based, and weighted mode-based analyses were concordant with IVW analysis results, which confirmed the result's robustness to violation of MR assumptions.

Our results were similar to the conclusions of other previous studies in relation to income, where lower household income and occupational social class were associated with higher risk of AD and dementia mortality (8,31,32). However, all prior study results were limited by selection bias and the heterogeneous nature of the comparison groups. The strength of using MR to study the causal associations between income and AD is that instruments are randomly allocated genotypes from birth, which reduces confounding and allows for a more robust causal effect. Moreover, as an unbiased, imprecise estimate, MR is preferable to assess causal association compared with a precise, biased estimate generated by observational studies (33).

The causal associations between income and AD suggest that social status has an important impact on cognitive disease. One possible mechanism underlying this causal effect may be the cognitive reserve hypothesis, according to which individuals with higher education and income levels have learned how to compensate for the disease process and can stave off the effects of brain disease longer (34). This study suggests that people with low SES are more vulnerable to dementia and that more attention should be paid to this vulnerable group to prevent dementia.

The significant results in IVW and other analyses attest to our findings' robustness, while the adverse result of the MR-Egger approach raise some concerns. Although the MR-Egger method is more sensitive to detecting violations of the instrumental variable assumptions, the influence of strong variants might have biased the results and have inflated type 1 error rates (35). In that case, the results of IVW analysis are more reliable.

Our study has several limitations. First, income was measured at the household level instead of the individual level. However, previous GWAS have shown that household income has a genetic correlation of 0.90 (SE =0.04) with educational attainment measured on an individual level, indicating that household-level effects are likely to be generalizable to individuals (36). Second, participants of AD-by-proxy group are from the UKB, meaning an overlap was present between exposure and outcome. Using the overlap calculation tool for 2-sample MR, bias caused by overlap was equal to zero, and the type I error rate was equal to 0.05 with an overlap rate of 0.8, suggesting no significant bias of the result (37). Third, although we had retrieved literature and applied heterogeneity and pleiotropy tests, the biological effects of these significant SNPs are not available to fully filter out pleiotropy. Fourth, the GWAS data set used was mainly derived from populations of European ancestry to avoid confounding due to population stratification; thus, the present results may not be applicable to other ethnic groups, and further research is needed to better understand how these findings may generalize to other populations. Finally, nonrandom selection into the analytical cohorts of the UKB (38), where individuals with higher income levels and general health are more likely to participate in the follow-up visit, might have biased the results.

Conclusions

In this study, we found robust evidence for the causal association between household income and AD risk. The higher income group demonstrated a reduced risk of developing AD, which is consistent with previous studies. This finding clarifies the effects of income on AD occurrence and may provide potential prevention strategies for vulnerable people.

Acknowledgments

This work was made possible by the generous sharing of GWAS summary statistics. The authors thank the investigators of the original income study; detailed data can be accessed at https://www.lothianbirthcohort.ed.ac.uk/content/gwas-summary-data. We also thank the participants, researchers, and staff of the Psychiatric Genomics Consortium (PGC) for sharing additional data on the PGC GWAS sample used in this study; it can be accessed at https://ctg.cncr.nl/software/summary_statistics. We thank the Zhanjiang Lab; Tianqiao and Chrissy Chen Institute; and the State Key Laboratory of Neurobiology and Frontiers Center for Brain Science of Ministry of Education, Fudan University for their help.

Funding: This study was supported by grants from the National Natural Science Foundation of China (No. 91849126) and the Shanghai Municipal Science and Technology Major Project (No. 2018SHZDZX01).

Footnote

Reporting Checklist: The authors have completed the STROBE reporting checklist. Available at https://dx.doi.org/10.21037/atm-21-344

Data Sharing Statement: Available at https://dx.doi.org/10.21037/atm-21-344

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/atm-21-344). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. 2020 Alzheimer's disease facts and figures. Alzheimers Dement 2020. [Epub ahead of print]. doi: 10.1002/alz.12068.
2. Ang ET, Tai YK, Lo SQ, et al. Neurodegenerative diseases: exercising toward neurogenesis and neuroregeneration. Front Aging Neurosci 2010;2:25.
3. Farah MJ. The Neuroscience of Socioeconomic Status: Correlates, Causes, and Consequences. Neuron 2017;96:56-71.
4. Kaplan GA, Turrell G, Lynch JW, et al. Childhood socioeconomic position and cognitive function in adulthood. Int J Epidemiol 2001;30:256-63.
5. Marden JR, Tchetgen Tchetgen EJ, Kawachi I, et al. Contribution of Socioeconomic Status at 3 Life-Course Periods to Late-Life Memory Function and Decline: Early and Late Predictors of Dementia Risk. Am J Epidemiol 2017;186:805-14.
6. Evans DA, Hebert LE, Beckett LA, et al. Education and other measures of socioeconomic status and risk of incident Alzheimer disease in a defined population of older persons. Arch Neurol 1997;54:1399-405.
7. Chaaya M, Phung K, Atweh S, et al. Socio-demographic and cardiovascular disease risk factors associated with dementia: Results of a cross-sectional study from Lebanon. Prev Med Rep 2018;9:1-5.
8. Deckers K, Cadar D, van Boxtel MPJ, et al. Modifiable Risk Factors Explain Socioeconomic Inequalities in Dementia Risk: Evidence from a Population-Based Prospective Cohort Study. J Alzheimers Dis 2019;71:549-57.
9. Larsson SC, Traylor M, Malik R, et al. Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. BMJ 2017;359:j5375.
10. Barrowman MA, Peek N, Lambie M, et al. How unmeasured confounding in a competing risks setting can affect treatment effect estimates in observational studies. BMC Med Res Methodol 2019;19:166.
11. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 2003;32:1-22.
12. Lawlor DA, Harbord RM, Sterne JA, et al. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med 2008;27:1133-63.
13. Mukherjee S, Walter S, Kauwe JSK, et al. Genetically predicted body mass index and Alzheimer's disease-related phenotypes in three large samples: Mendelian randomization analyses. Alzheimers Dement 2015;11:1439-51.
14. Andrews SJ, Goate A, Anstey KJ. Association between alcohol consumption and Alzheimer's disease: A Mendelian randomization study. Alzheimers Dement 2020;16:345-53.
15. Anderson EL, Howe LD, Wade KH, et al. Education, intelligence and Alzheimer's disease: evidence from a multivariable two-sample Mendelian randomization study. Int J Epidemiol 2020;49:1163-72.
16. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013;37:658-65.
17. Hill WD, Davies NM, Ritchie SJ, et al. Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. Nat Commun 2019;10:5741.
18. Jansen IE, Savage JE, Watanabe K, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet 2019;51:404-13.
19. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018;7:34408.
20. MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res 2017;45:D896-901.
21. Staley JR, Blackshaw J, Kamat MA, et al. PhenoScanner: a database of human genotype-phenotype associations. Bioinformatics 2016;32:3207-9.
22. Kamat MA, Blackshaw JA, Young R, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics 2019;35:4851-3.
23. Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. Int J Epidemiol 2011;40:755-64.
24. Burgess S, Scott RA, Timpson NJ, et al. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. Eur J Epidemiol 2015;30:543-52.

25. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. Int J Epidemiol 2013;42:1497-501.
26. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44:512-25.
27. Bowden J, Davey Smith G, Haycock PC, et al. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 2016;40:304-14.
28. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol 2017;46:1985-98.
29. Bowden J, Del Greco MF, Minelli C, et al. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med 2017;36:1783-802.
30. Verbanck M, Chen CY, Neale B, et al. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018;50:693-8.
31. Fernández-Blázquez MA, Noriega-Ruiz B, Ávila-Villanueva M, et al. Impact of individual and neighborhood dimensions of socioeconomic status on the prevalence of mild cognitive impairment over seven-year follow-up. Aging Ment Health 2021;25:814-23.
32. Korhonen K, Einiö E, Leinonen T, et al. Midlife socioeconomic position and old-age dementia mortality: a large prospective register-based study from Finland. BMJ Open 2020;10:e033234.
33. Bautista LE, Smeeth L, Hingorani AD, et al. Estimation of bias in nongenetic observational studies using "mendelian triangulation". Ann Epidemiol 2006;16:675-80.
34. Stern Y. Cognitive reserve. Neuropsychologia 2009;47:2015-28.
35. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol 2017;32:377-89.
36. Hill WD, Hagenaars SP, Marioni RE, et al. Molecular Genetic Contributions to Social Deprivation and Household Income in UK Biobank. Curr Biol 2016;26:3083-9.
37. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genet Epidemiol 2016;40:597-608.
38. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. Am J Epidemiol 2017;186:1026-34.

Cite this article as: Wang RZ, Yang YX, Li HQ, Shen XN, Chen SD, Cui M, Wang Y, Dong Q, Yu JT. Genetically determined low income modifies Alzheimer's disease risk. Ann Transl Med 2021;9(15):1222. doi: 10.21037/atm-21-344

