The Causal Relationship between Genetically Predicted Biological Aging, Alzheimer’s Disease and Cognitive Function: A Mendelian Randomisation Study

Yitong Hao 1, Weibo Tian1, Bo Xie 1, Xiaoshu Fu 1, Shufeng Wang 2 and Yu Yang 1,\*

1 *Department of Neurology and Neuroscience Center, The First Hospital of Jilin University, Jilin 130021, China.*

2 *Jilin Province Branch of ChinaUnicom, Jilin 130000, China.*

*\*Correspondence: yang\_yu@jlu.edu.cn*

**Supplementary document 1. Definition, limitations, and resolution of Mendelian randomization (MR) study core assumptions**

1. Core assumptions of MR study.

(1) Assumption of association: genetic variation (instrumental variables) is reliably associated with exposure factors;

(2) Independence assumption: genetic variance (instrumental variable) is uncorrelated with any known and unknown confounders;

(3) Exclusivity assumption: genetic variation (instrumental variable) affects the outcome only through exposure factors.

2. Limitations and resolution

(1) Weak instrumental variable violates the assumption of association. In a two-sample MR, weak instrumental variable may introduce regression dilution bias and thus underestimate the exposure-outcome association. To avoid this, the data used in this study were from GWAS studies with large sample sizes, and a series of screening methods were used to ultimately obtain SNPs that were strongly associated with exposure factors. For example, the Cragg-Donald F statistic was used to evaluate the strength of instrumental variables in this study. The F-statistic for the strength of instrumental variables is at least 10, and when F is less than 10, the estimates of causal effects are severely biased. In addition, the F-statistic corresponds to the p-value of the significance of the instrumental variable-exposure association, which is usually used as a screening threshold for instrumental variables. In order to ensure that the instrumental variables in the model are sufficiently strong, in this study, the genome-wide significance level (P < 5 × 10-6 ~ P < 5 × 10-8) was used as the threshold to screen the instrumental variables, so as to avoid potential weak instrumental variable bias.

(2) Horizontal pleiotropy of genetic variants violates the assumption of independence and the assumption of exclusivity. Horizontal pleiotropy can be understood as genetic variation influencing the occurrence of an outcome factor through pathways other than exposure. To avoid this, we first identified and excluded instrumental variables with pleiotropy by MR-Pleiotropy Residual Sum and Outlier method, and then tested whether there was pleiotropy in causal effects using the MR-Egger intercept method.

(3) Linkage disequilibrium violates the assumption of independence and the assumption of exclusivity. Genetic variants with similar genomic positions tend to be co-inherited, which is called linkage disequilibrium. To avoid this, strict screening conditions (R2 < 0.01 ~ 0.001, window size = 5000 ~10000 kb) were used in this study to remove SNPs that might be linkage disequilibrium.

(4) Population stratification violates the assumptions of independence and exclusivity. Population stratification refers to the fact that the frequency of genetic variants varies among populations with different genetic backgrounds, leading to spurious associations between genetic variants and outcomes. Therefore, to avoid this one problem, the GWAS data used in this study were all Europeans with the same genetic background.