Biological Psychiatry

Archival Report

# A Phenome-wide Association and Mendelian Randomization Study for Alzheimer's Disease: A Prospective Cohort Study of 502,493 Participants From the UK Biobank

Shi-Dong Chen, Wei Zhang, Yu-Zhu Li, Liu Yang, Yu-Yuan Huang, Yue-Ting Deng, Bang-Sheng Wu, John Suckling, Edmund T. Rolls, Jian-Feng Feng, Wei Cheng, Qiang Dong, and Jin-Tai Yu

## ABSTRACT

**BACKGROUND:** Considerable uncertainty remains regarding associations of multiple risk factors with Alzheimer's disease (AD). We aimed to systematically screen and validate a wide range of potential risk factors for AD.

**METHODS:** Among 502,493 participants from the UK Biobank, baseline data were extracted for 4171 factors spanning 10 different categories. Phenome-wide association analyses and time-to-event analyses were conducted to identify factors associated with both polygenic risk scores for AD and AD diagnosis at follow-up. We performed two-sample Mendelian randomization analysis to further assess their potential causal relationships with AD and imaging association analysis to discover underlying mechanisms.

**RESULTS:** We identified 39 factors significantly associated with both AD polygenic risk scores and risk of incident AD, where higher levels of education, body size, basal metabolic rate, fat-free mass, computer use, and cognitive functions were associated with a decreased risk of developing AD, and selective food intake and more outdoor exposures were associated with an increased risk of developing AD. The identified factors were also associated with AD-related brain structures, including the hippocampus, entorhinal cortex, and inferior/middle temporal cortex, and 21 of these factors were further supported by Mendelian randomization evidence.

**CONCLUSIONS:** To our knowledge, this is the first study to comprehensively and rigorously assess the effects of wide-ranging risk factors on AD. Strong evidence was found for fat-free body mass, basal metabolic rate, computer use, selective food intake, and outdoor exposures as new risk factors for AD. Integration of genetic, clinical, and neuroimaging information may help prioritize risk factors and prevention targets for AD.

https://doi.org/10.1016/j.biopsych.2022.08.002

Alzheimer's disease (AD) is the leading contributor to the global burden of disease because of its high prevalence and disabling consequences (1). Many risk factors have been identified as associated with AD. Recent publications have focused on validating limited sets of risk factors, and data from large-scale cohorts are increasingly used to acquire knowledge on actionable strategies that target those factors and prevent disease onset (2,3). However, these studies have been mainly carried out with hypothesis-driven designs. Additional factors may remain overlooked or unknown.

AD is a hereditary disease. The heritability is estimated to be between 60% and 80% (4). Recent genome-wide association studies (GWASs) conducted by Schwartzentruber *et al.* (5) and Bellenguez *et al.* (6) have, respectively, identified genetic variants spanning 37 and 75 risk loci. By using summary statistics from these large-scale GWASs, multiple factors were discovered to share genetic architecture with AD and possibly be involved in the pathology of the disease (7–10). This revealed that genetic association is an effective tool for identifying AD risk factors and identifying those that may have causal effects on AD.

A phenome-wide association study (PheWAS) is a type of hypothesis-free analysis in which a broad range of phenotypes can be examined in genetic association with a disease. Through application of a polygenic risk score (PRS) of AD as a proxy for AD risk, the associations of a wide array of both established and undiscovered nongenetic factors with AD can be systematically screened (11). This approach has received relatively little attention owing to a lack of resources with sufficient variety and volume. Large-scale datasets such as the UK Biobank (UKB) now provide an unparalleled opportunity for this approach (12). Such efforts have a significant advantage compared to traditional observational studies that require long-term follow-up and large sample sizes, as statistical power can be dramatically increased by exploiting genetic and phenotypic information. Other analytic techniques can be subsequently performed to

SEE COMMENTARY ON PAGE 756

 
 ISSN: 0006-3223

rigorously assess the putative associations and reduce false positive findings, including traditional longitudinal analysis and Mendelian randomization (MR) analysis (13).

In this study, using phenotypic and genomic data from 502,493 UKB participants, we first generated AD-PRS and conducted a PheWAS to investigate the effect and significance of associations between AD-PRS and all the available risk factors. A total of 84 factors were found to be associated with AD-PRS and followed by time-to-event analyses to evaluate their associations with clinically significant AD at follow-up. Next, two-sample MR analysis was performed to evaluate potential causal relationships between the identified factors and AD. Finally, we conducted an imaging analysis to explore the underlying biological mechanisms for the factors associated with both AD-PRS and incident AD.

## METHODS AND MATERIALS

### Study Population

The UKB is a population-based cohort of more than 500,000 participants in the United Kingdom recruited between 2006 and 2010 (14). The UKB has research tissue bank approval from the North West Multi-centre Research Ethics Committee (https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us/ethics) and provided oversight for this study. Participation is voluntary, and participants are free to withdraw at any time without giving any reason. Written informed consent was obtained from all participants. Genetic and phenotypic data were obtained at baseline. Clinical outcomes including AD diagnoses were available over a follow-up period from 2007 to 2020 via hospital inpatient records, death certificates, primary care records, and self-reports. The initial sample included 502,493 participants between 38 and 73 years of age. Data acquisition and analyses in this study were conducted under UKB Application No. 19542. AD-PRS generation, PheWAS, and MR analysis only included populations of European ancestry to reduce the impact of population structures on genetic data analysis. To ensure adequate power, all participants, regardless of ethnicity, were included in time-to-event and imaging analyses.

### AD-PRS Generation

Genotype data were available for all 502,493 participants in the UKB. Detailed genotyping and quality control procedures by the UKB are available in a previous publication (15). We excluded single nucleotide polymorphisms with call rates $<95\%$, minor allele frequency $<0.1\%$, and deviation from the Hardy-Weinberg equilibrium with $p < 1 \times 10^{-10}$. We selected subjects who were estimated to have recent British ancestry based on self-report information and principal component analyses of the genotypes and had no more than 10 putative third-degree relatives in the kinship table. After the quality control procedures, we obtained a total of 591,050 single nucleotide polymorphisms and 337,199 participants.

We calculated the PRS with the summary statistics from a meta-analysis of GWASs from 4 large AD consortia (16), which included a total of 7,055,881 single nucleotide polymorphisms and 54,062 individuals (17,008 cases and 37,054 control subjects).

PRSice was the software (http://www.PRSice.info) used for the calculation of AD-PRS. We used $p$ value–informed clumping with a cutoff of $r^2 = 0.1$ in a 250-kb window in the analysis (17). $P$ thresholds for scoring were set at $p < .0005$, $p < .001$, $p < .005$, $p < .01$, $p < .05$, $p < .1$, $p < .5$, and $p < 1$ (18).

### Risk Factors

The factors included in the PheWAS consisted of 10 broad categories (containing 4171 variables), which were 1) sociodemographic, 2) physical measures, 3) lifestyle and environment, 4) health conditions, 5) mental health, 6) medications and operations, 7) cognitive function, 8) sex-specific factors, 9) employment, and 10) early-life factors. These variables were from 4 categories (population characteristics, assessment center, online follow-up, and health-related outcomes) in the UKB showcase and were recategorized to a small extent based on the framework of the showcase. For further details, see Figure 1; Supplemental Data Tables S1 and S2 in Supplement 2.

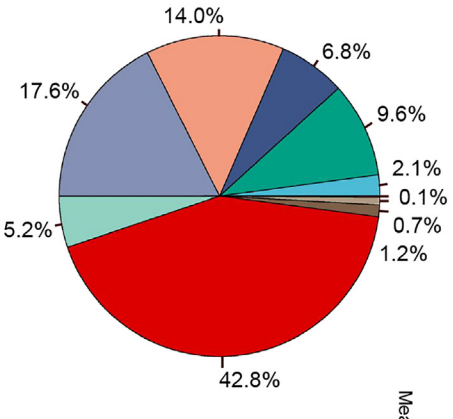
### Imaging Data of Brain Structures

Quality-controlled T1-weighted magnetic resonance imaging data were used for studying associations between factors and brain structures. Details of the processing procedure can be found in the UKB protocol (https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/brain_mri.pdf) and Supplemental Methods in Supplement 1. Surface templates were used to extract imaging-derived phenotypes referring to atlas regions' surface volume (19), and FreeSurfer's aseg was used to extract subcortical regions (20). FreeSurfer aparc (ID = 192) and aseg (ID = 190) atlas corresponding to 68 cortical regions and 41 subcortical regions were applied in this study.

### Statistical Analysis

**Phenome-wide Association Study.** The PHESANT package in R was used to test the PheWAS associations. The PHESANT's automated rule-based method is described in detail in a previous publication (21). In brief, decision rules were based on the variable type, and each variable was categorized as 1 of 4 data types: continuous, ordered categorical, unordered categorical, or binary. Normality of continuous data was ensured by an inverse normal rank transformation prior to testing. In this study, AD-PRSs were set as independent variables, and selected factors were set as dependent variables, with age, gender, genotyping array, the first 10 genetic principal components, and the assessment center as covariates included in the model. Overall, 4171 factors $\times$ 8 AD-PRS = 33,368 tests across factors, and AD-PRS $p$ thresholds were corrected altogether by Benjamini-Hochberg procedure (false discovery rate [FDR] correction) ($q < .05$). Only factors found significantly associated with AD-PRS at a minimum of 4 PRS variant $p$ thresholds would be evaluated subsequently. For direct comparison of the results between linear and logistic regression models, standardized regression coefficients were estimated as effect sizes (β) for both types of models and were reported with log-transformed odds ratio (OR) for binary dependent variables. Two-sided statistical tests were applied in all analyses.

| Category | Number of factors | Sample sizes (Median[Range]) |
|---|---|---|
| Sociodemographics | 89 | 288853 (15102-337128) |
| Employment | 400 | 84309 (1654-216683) |
| Physical measures | 282 | 68609 (839-336568) |
| Lifestyle and environment | 582 | 47727 (1198-337093) |
| Health conditions | 733 | 337128 (513-337128) |
| Mental health | 215 | 109672 (999-337128) |
| Medications and operations | 1784 | 337128 (1694-337128) |
| Cognitive function | 49 | 29667 (7180-336810) |
| Sex-specific factors | 31 | 130686 (14335-180819) |
| Early life factors | 6 | 331789 (255833-336672) |



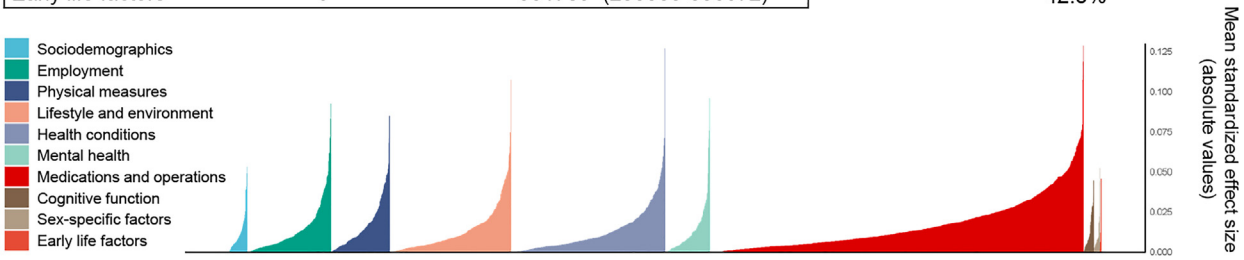




**Figure 1.** Overview of risk factors. There were 4171 factors from 10 categories included. The top panel demonstrates the number of factors with median and range of sample size in each category, with the size of pie chart section indicating the ratio of each category. The bottom panel is the plot illustrating the effect sizes of all included factors. The y-axis represents the mean standard effect size (absolute value) across AD-PRS generated with all 8 *p* thresholds. Colors indicate categories. AD-PRS, Alzheimer's disease-polygenic risk score.

**Time-to-Event Analysis.** Multivariable Cox proportional hazard regression models were used to examine the association of the PheWAS-selected factors with incident AD. Follow-up was calculated in person-years from the date of recruitment until the date of first incident AD diagnosis, death, loss to follow-up, or the last date of hospital admission data available, whichever came first. The model was adjusted for age, gender, and *APOE* ε4 carrier. The proportional hazards assumption was evaluated using tests of Schoenfeld residuals. In the main analysis, participants with prevalent dementia diagnoses, including AD diagnosis at baseline or before age 50, were excluded and were limited to those with follow-up periods of ≥3 years.

**MR Analysis.** TwosampleMR package in R was used to conduct two-sample MR analyses. GWAS summary data for factors were acquired from the Medical Research Council Integrative Epidemiology Unit OpenGWAS database (https://gwas.mrcieu.ac.uk/), and summary data from a recent large AD GWAS meta-analysis (7) was used as an outcome dataset. The inverse variance weighted (IVW) method was the primary method for conducting MR. MR–pleiotropy residual sum and outlier (MR-PRESSO) was mainly used for detecting potential pleiotropy and correcting IVW estimates. See reverse MR and more details in Supplemental Methods in Supplement 1.

### Associations Between Factors and Brain Structures

Linear regression models were applied to investigate the association of selected factors with brain morphometric measures. Covariates were age, gender, *APOE* ε4 carrier, and imaging scanning site. FDR corrections were conducted for multiple comparisons among cortical and subcortical regions, respectively. Hippocampus (22,23), entorhinal (24,25), inferior temporal (26–28), and middle temporal (26) cortices out of FreeSurfer aparc and aseg atlas were selected as AD-related brain structures based on previous publications, which were also supported with UKB data (see Supplemental Results in Supplement 1).

## RESULTS

A summary of the 4171 factors that entered the analytic pipeline is shown in Figure 1. The analytic pipeline is illustrated in Figure S1 in Supplement 1, with participant inclusion in each analysis step shown in Figure S2 in Supplement 1. An overview of the results from the 4 analytic steps is shown in Figure 2.

### PheWAS and Time-to-Event Analyses Identified 51 Factors Spanning 6 Categories as Consistently Associated With AD

Initially, 84 factors survived PheWAS. This included 8 sociodemographics, 33 physical measures, 18 lifestyle and environment, 5 health conditions, 8 medications and operations, 11 cognitive functions, and 1 sex-specific factor (standardized coefficients $\beta = -0.0758$ to 0.369, $p_{FDR} = 5.60 \times 10^{-55}$ to .050) (Figure 3; Supplemental Data Table S1 in Supplement 2). All significant associations showed an identical effect direction for each of the 84 factors. Population baseline characteristics of PheWAS can be found in Table S1 in Supplement 1. The complete PheWAS results for the 4171 factors are presented in Supplemental Data Table S2 in Supplement 2. In total, 22.4%


**Figure 2.** Factors associated with AD polygenic risk score across analytic steps. Main results (39 of 84 factors) are shown in bold. Green and red cells indicate decreased and increased risk of AD, while blank cells indicate no results supporting associations between factors and AD. Initially, 84 factors survived PheWAS; then, 51 survived time-to-event analyses; then, 21 of 51 factors were further supported by MR, and 39 factors showed significant results in imaging analysis. Except for nonsignificant associations discovered between factors and AD, other conditions shown by blank cells include the following: [a]Factors are duplicates, thus not entering subsequent analysis steps; [b]Factors have no nonincident AD, thus not entering time-to-event analysis and subsequent steps; [c]Factors show contradictory effect directions between time-to-event analysis and PheWAS; [d]Factors show no associations with brain structures in imaging analysis; and [e]Factors show associations with inconsistent effect directions between imaging analyses and PheWAS/time-to-event analyses. Details can be found in Supplemental Data Tables S1, S4, S6, and S9–S12 in Supplement 2. A/AS, Advanced/Advanced Subsidiary; AD, Alzheimer's disease; avMSE, average mean spherical equivalent; BMI, body mass index; FI, fluid intelligence; IPAQ, International Physical Activity Questionnaire; MET, metabolic equivalent of task; MR, Mendelian randomization; NM, numeric memory; O levels/GCSEs, General Certificate of Secondary Education Ordinary levels; PheWAS, phenome-wide association study; PM, pair matching; SDS, symbol digit substitution; TM, Trail Making; UK, United Kingdom.

of cognitive phenotypes and 11.7% of physical measures showed significant results in PheWAS. The proportions were lower for phenotypes of sociodemographic (8.9%), health conditions (5.95%), sex-specific factors (3.22%), lifestyle and environment (1.01%), and medications and operations (0.45%), and no phenotypes of employment, mental health, or early-life factors showed significant results. After excluding 3 duplicate factors with smaller sample sizes (body mass index, weight, and fluid intelligence score) and 2 factors where all incident AD were from the disease group (vascular dementia and unspecified dementia), 79 factors entered time-to-event analysis.

Of these, 51 factors survived time-to-event analyses, while 26 showed nonsignificant associations with incident AD, and 2 had inconsistent effect directions between PheWAS and time-to-event analyses. Baseline descriptions are presented in

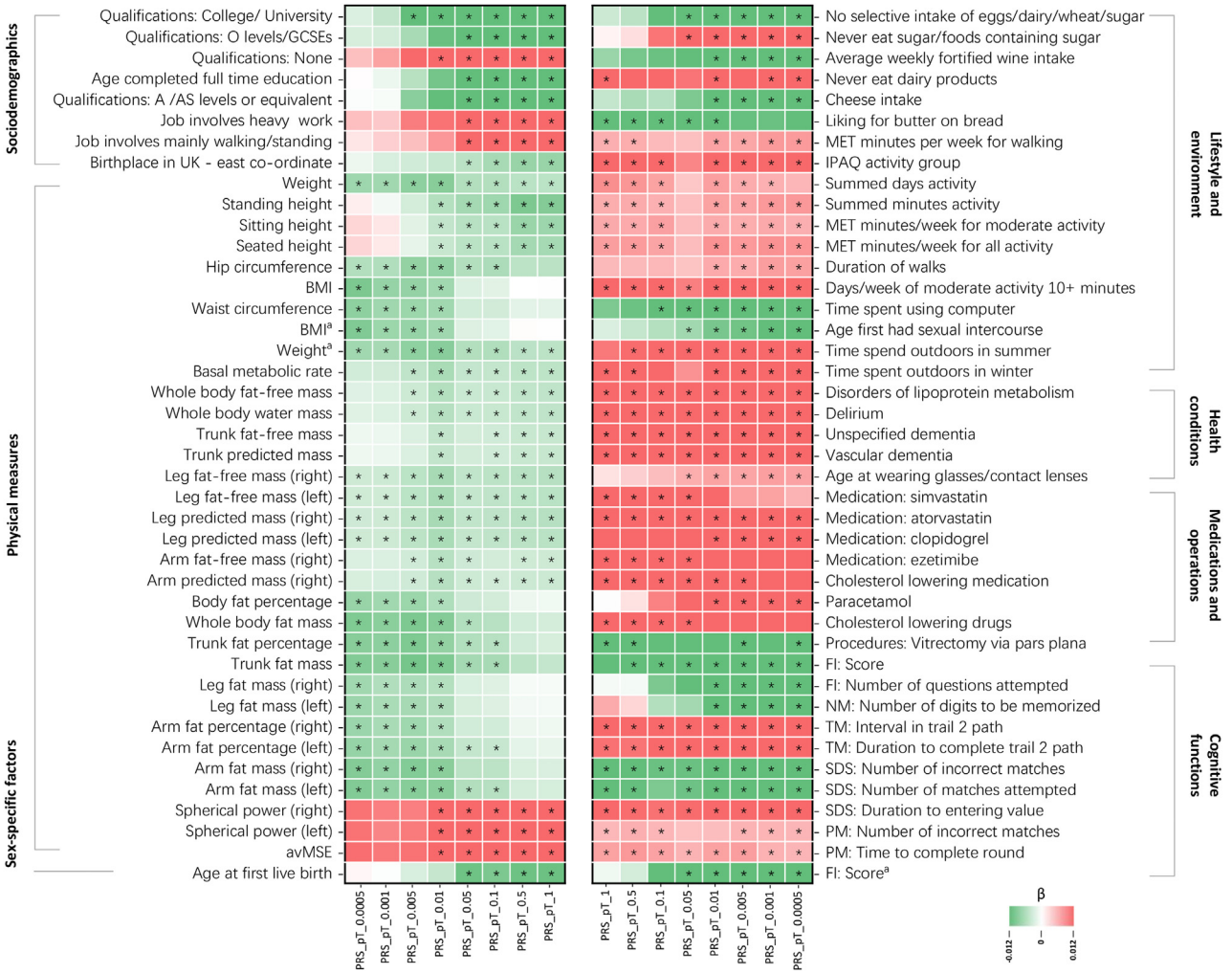



**Figure 3.** Heatmap for the factors significantly associated with AD-PRS. The shown factors were significantly associated with AD-PRS at a minimum of 4 *p* thresholds for AD-PRS. Shades of cells indicate the standardized effect sizes (β) between each AD-PRS and each phenotype with a darker color denoting a larger effect size. Cells with an asterisk indicate significant associations after FDR correction. Factors marked with superscript letter a means duplication. Detail can be found in Supplemental Data Table S1 in Supplement 2. A/AS, Advanced/Advanced Subsidiary; AD, Alzheimer's disease; avMSE, average mean spherical equivalent; BMI, body mass index; FDR, false discovery rate; FI, fluid intelligence; IPAQ, International Physical Activity Questionnaire; MET, metabolic equivalent of task; MR, Mendelian randomization; NM, numeric memory; O levels/GCSEs, General Certificate of Secondary Education Ordinary levels; PheWAS, phenome-wide association study; PM, pair matching; PRS, polygenic risk score; pT, *p* threshold; SDS, symbol digit substitution; TM, Trail Making; UK, United Kingdom.

Supplemental Data Table S3 in Supplement 2. The complete results of time-to-event analysis for the 79 factors are provided in Supplemental Data Table S4 in Supplement 2. We herein reported 22 of the 79 factors with details in main text (Table 1).

Qualifications of college/university degree was associated with both lower AD-PRS (β = −0.0366 to −0.0027, $p_{FDR}$ = $3.25 \times 10^{-18}$ to $6.32 \times 10^{-3}$) and decreased risk of incident AD (hazard ratio = 0.74, 95% CI, 0.659–0.831, $p = 3.81 \times 10^{-7}$). In contrast, no qualification attainment was associated with higher AD-PRS (β = 0.0183 to 0.0392, $p_{FDR} = 6.67 \times 10^{-13}$ to $1.90 \times 10^{-2}$) and increased risk of incident AD (HR = 1.371, 95% CI, 1.248–1.507, $p = 5.80 \times 10^{-11}$). Consistent with these results, older age at completing full-time education was associated with lower AD-PRS (β = −0.0270 to −0.0241, $p_{FDR} = 8.11 \times 10^{-9}$ to $5.37 \times 10^{-7}$) and decreased risk of AD (HR = 0.963, 95% CI, 0.939–0.989, $p = 4.46 \times 10^{-3}$).

Multiple associations were identified for physical measures. Higher levels of global body measures were associated with lower AD-PRS and decreased risk of incident AD, including body mass index (β = −0.0103 to −0.0072, $p_{FDR} = 1.67 \times 10^{-6}$ to $2.68 \times 10^{-3}$; HR = 0.982, 95% CI, 0.972–0.992, $p = 6.02 \times 10^{-4}$), weight (β = −0.0091 to −0.0052, $p_{FDR} = 7.87 \times 10^{-7}$ to $3.52 \times 10^{-2}$; HR = 0.989, 95% CI, 0.985–0.992, $p = 5.59 \times 10^{-10}$), standing height (β = −0.0099 to −0.0043, $p_{FDR} = 4.99 \times 10^{-13}$ to $4.70 \times 10^{-2}$; HR = 0.977, 95% CI, 0.970–0.983, $p = 4.06 \times 10^{-11}$), hip circumference (β = −0.0086 to −0.0058, $p_{FDR} = 4.09 \times 10^{-6}$ to $4.41 \times 10^{-2}$; HR = 0.988, 95% CI, 0.983–0.994, $p = 1.04 \times 10^{-5}$), basal

**Table 1. Results for PheWAS-Selected Factors in Time-to-Event Analysis**

| Factors | Sample Size, *n* | HR | 95% CI | *p* Value |
|---|---|---|---|---|
| Sociodemographics | | | | |
| Qualifications: college/university | 105,450/244,615 | 0.740 | 0.659–0.831 | $3.81 \times 10^{-7}$ |
| Qualifications: none | 65,688/284,377 | 1.371 | 1.248–1.507 | $5.80 \times 10^{-11}$ |
| Age completed full-time education | 243,608 | 0.963 | 0.939–0.988 | $4.46 \times 10^{-3}$ |
| Physical Measures | | | | |
| BMI | 352,351 | 0.982 | 0.972–0.992 | $6.02 \times 10^{-4}$ |
| Weight | 352,487 | 0.989 | 0.985–0.992 | $5.59 \times 10^{-10}$ |
| Standing height | 352,737 | 0.977 | 0.970–0.983 | $4.06 \times 10^{-11}$ |
| Hip circumference | 352,873 | 0.988 | 0.983–0.993 | $1.04 \times 10^{-5}$ |
| Basal metabolic rate | 347,141 | 0.9998 | 0.9997–0.9999 | $1.60 \times 10^{-11}$ |
| Whole body fat-free mass | 347,123 | 0.974 | 0.966–0.981 | $1.73 \times 10^{-11}$ |
| Lifestyle and Environment | | | | |
| No selective intake of eggs/dairy/wheat/sugar | 269,915/82,376 | 0.672 | 0.613–0.738 | $3.89 \times 10^{-17}$ |
| Never eat sugar or foods containing sugar | 61,792/269,915 | 1.353 | 1.218–1.502 | $1.48 \times 10^{-8}$ |
| Time spent using computer | 276,808 | 0.904 | 0.865–0.945 | $6.66 \times 10^{-6}$ |
| Time spent outdoors in summer | 320,857 | 1.022 | 1.001–1.043 | $3.59 \times 10^{-2}$ |
| Time spent outdoors in winter | 279,866 | 1.052 | 1.026–1.079 | $8.92 \times 10^{-5}$ |
| Health Conditions | | | | |
| Disorders of lipoprotein metabolism | 88,272/265,227 | 1.173 | 1.071–1.285 | $6.18 \times 10^{-4}$ |
| Delirium | 3268/350,238 | 13.82 | 12.43–15.34 | $4.06 \times 10^{-11}$ |
| Medications and Operations | | | | |
| Medication: simvastatin | 45,297/308,209 | 1.147 | 1.031–1.276 | $1.18 \times 10^{-2}$ |
| Cognitive Function | | | | |
| FI: score | 111,953 | 0.798 | 0.761–0.836 | $1.99 \times 10^{-21}$ |
| NM: number of digits to be memorized | 39,159 | 0.771 | 0.719–0.826 | $2.59 \times 10^{-13}$ |
| TM: interval in alphanumeric path (trail 2) | 79,097 | 1.00003 | 1.00001–1.00004 | $4.06 \times 10^{-3}$ |
| SDS: number of correct matches | 84,964 | 0.834 | 0.810–0.858 | $1.44 \times 10^{-35}$ |
| PM: number of incorrect matches | 353,190 | 1.084 | 1.067–1.100 | $2.42 \times 10^{-24}$ |

Only 22 of 51 factors significantly associated with incident Alzheimer's disease reported in main text are shown here. Full results of time-to-event analyses can be found in Supplemental Data Table S4 in Supplement 2. Hazard ratios were adjusted for age, gender, and *APOE* ε4 carriers. Sample sizes for binary variables are shown as case/control.

BMI, body mass index; FI, fluid intelligence; HR, hazard ratio; IPAQ, International Physical Activity Questionnaire; MET, metabolic equivalent of task; NM, numeric memory; PheWAS, phenome-wide association study; PM, pair matching; SDS, symbol digit substitution; TM, Trail Making.

metabolic rate (β = −0.0064 to −0.0061, $p_{FDR} = 4.06 \times 10^{-6}$ to $7.63 \times 10^{-3}$; HR = 0.9998, 95% CI, 0.9997–0.9999, $p = 1.04 \times 10^{-5}$), and whole body fat-free mass (β = −0.0057 to −0.0038, $p_{FDR} = 3.55 \times 10^{-5}$ to $2.72 \times 10^{-2}$; HR = 0.974, 95% CI, 0.966–0.981, $p = 1.73 \times 10^{-11}$). Significant associations were also found for multiple physical measures in different parts of the body (Figures 2 and 3).

No selective intake of 4 types of food (eggs, dairy, wheat, and sugar) was associated with lower AD-PRS and lower risk of developing AD (β = −0.0294 to −0.0246, $p_{FDR} = 8.86 \times 10^{-5}$ to $7.99 \times 10^{-3}$; HR = 0.672, 95% CI, 0.613–0.737, $p = 3.89 \times 10^{-17}$), while never eating sugar or foods/drinks containing sugar was associated with higher AD-PRS and increased risk of AD (β = 0.0151 to 0.0175, $p_{FDR} = 5.65 \times 10^{-3} - 4.54 \times 10^{-2}$; HR = 1.352, 95% CI, 1.218–1.502, $p = 1.48 \times 10^{-8}$). Interestingly, more time spent using a computer was associated with both lower AD-PRS and decreased risk of incident AD (β = −0.0251 to −0.0128, $p_{FDR} = 5.36 \times 10^{-9}$ to $2.77 \times 10^{-2}$), and more time spent outdoors in summer/winter was found to be associated with higher AD-PRS and elevated risk of AD (β = 0.0129 to 0.0244, $p_{FDR} = 2.88 \times 10^{-9}$ to $4.80 \times 10^{-2}$; HR = 1.022–1.052, 95% CI, 1.001–1.079; $p = 3.59 \times 10^{-2}$ to $8.92 \times 10^{-5}$).

Associations with both higher AD-PRS and increased risk of AD were also identified for certain health conditions and medications, including disorders of lipoprotein metabolism and other lipidemias (β = 0.0099 to 0.0688, $p_{FDR} = 4.31 \times 10^{-40}$ to $2.43 \times 10^{-5}$; HR = 1.173, 95% CI, 1.071–1.285, $p = 6.18 \times 10^{-14}$), delirium (β = 0.042 to 0.1988, $p_{FDR} = 6.12 \times 10^{-17}$ to $2.43 \times 10^{-3}$; HR = 13.81, 95% CI, 12.43–15.34, $p = 4.06 \times 10^{-11}$), and simvastatin intake (β = 0.0170 to 0.0566, $p_{FDR} = 3.27 \times 10^{-16}$ to $6.26 \times 10^{-7}$; HR = 1.15, 95% CI, 1.03–1.28, $p = 1.18 \times 10^{-12}$).

Better performance in different cognitive function tests were consistently associated with both lower AD-PRS and decreased risk of AD, including higher total scores in the fluid intelligence test (β = −0.0508 to −0.0125, $p_{FDR} = 4.75 \times 10^{-11}$ to $2.05 \times 10^{-3}$; HR = 0.798, 95% CI, 0.761–0.836, $p = 1.99 \times 10^{-21}$), more numbers of digits memorized in the numeric memory test (β = −0.0301 to −0.0073, $p_{FDR} = 1.26 \times 10^{-2}$ to $4.13 \times 10^{-2}$; HR = 0.770, 95% CI, 0.719–0.826, $p = 2.59 \times 10^{-13}$), more numbers of symbol digit matches made correctly

in the symbol digit substitution test ($\beta = -0.0213$ to $-0.0053$, $p_{FDR} = 1.56 \times 10^{-4}$ to $1.63 \times 10^{-2}$; HR = 0.834, 95% CI, 0.811–0.858, $p = 1.44 \times 10^{-35}$), fewer numbers of incorrect matches per round in the pairs matching test ($\beta = 0.0024$ to 0.0098, $p_{FDR} = 1.14 \times 10^{-2}$ to $3.95 \times 10^{-2}$; HR = 1.084, 95% CI, 1.067–1.100, $p = 2.42 \times 10^{-24}$), and shorter interval in alphanumeric path (trail 2) in the Trail Making Test ($\beta = 0.0073$ to 0.02553, $p_{FDR} = 2.78 \times 10^{-6}$ to $3.07 \times 10^{-3}$; HR = 1.016, 95% CI, 1.013–1.019, $p = 1.25 \times 10^{-25}$). Other cognitive factors also showed consistent effects (Figures 2 and 3).

## MR Evidence for Associations Between 21 Factors and AD

The potentially causal effects on AD were found for 21 of 51 factors, which showed the same effect direction as those of PheWAS and time-to-event analyses. These contained the aforementioned factors, including 2 factors in education (college or university degree [$OR_{IVW} = 0.475$, $p_{IVW} = 2.01 \times 10^{-5}$] and no qualifications attained [$OR_{IVW} = 4.00$, $p_{IVW} = 1.51 \times 10^{-4}$]), 4 in physical measures (weight, standing height, basal metabolic rate, and whole body fat-free mass [$OR_{IVW}$ = 0.767–0.881, $p_{IVW}$ $3.41 \times 10^{-3}$ to 0.044]), 1 in lifestyle and environment (time spent using computer [$OR_{IVW} = 0.713$, $p_{IVW} = .041$]), and 2 in cognitive function (fluid intelligence score [$OR_{MR\text{-}PRESSO\text{–}corrected} = 0.916$, $p_{MR\text{-}PRESSO\text{–}corrected} = .040$] and interval in alphanumeric path [trail 2] in the Trail Making Test [$OR_{IVW} = 1.726$, $p_{IVW} = 7.65 \times 10^{-3}$]) (see major results in Figure 4). These associations were all significant at $p < .05$ (IVW or corrected by the MR-PRESSO method) without directional pleiotropy, and the estimates were consistent across MR methods. Full results can be found in Supplemental Data Tables S6–S8 in Supplement 2.

## Associations Between 39 Identified Factors and AD Neuroimaging Hallmarks

Out of 51 factors, 39 showed significant associations with volumes of multiple AD-related brain structures, including the hippocampus and entorhinal, inferior temporal, and middle temporal cortices (Figure S3 in Supplement 1), with the consistent effect directions as those of PheWAS and time-to-event analysis. These included all the 21 factors with MR evidence and other 18 factors such as no selective intake of 4 types of food (eggs, dairy, wheat, and sugar), never eating sugar or foods/drinks containing sugar, time spent outdoors in summer/winter, disorders of lipoprotein metabolism, simvastatin intake, and fluid intelligence score. We plotted the significant associations for those among factors with largest effects (Figure 5). Full lists of statistics ($\beta$ and $p_{FDR}$) for associations between 51 factors and 68 cortical/41 subcortical regions are presented in Supplemental Data Tables S9–S12 in Supplement 2.

## DISCUSSION

To our knowledge, this is the first study to integrate genetic, clinical, and neuroimaging information from the UKB to comprehensively evaluate the associations between a wide array of risk factors and AD. The most robust findings were found for 21 associations where higher education attainment, greater body size, greater fat-free mass, faster basal metabolic rate, more computer use, and better cognitive status were associated with a decreased risk of developing AD. Strong evidence was also found for the other 18 factors where selective food intake and more time spent outdoors were associated with an increased risk of developing AD. Collectively, fat-free mass, basal metabolic rate, computer use, selective food intake, and outdoor exposures were identified as novel risk factors for AD.

The findings on education and cognitive function were in line with common views of risk factors for AD. The longstanding evidence from the literature has validated that higher education and better cognitive function are associated with lower risks of developing AD (29,30). Recent research using PRS also found that increased education was associated with decreased odds of AD diagnosis (31), and cognitive status greatly improved prediction of AD in 3–8 years (32). These results emphasize their important roles in the risk factor profiles of AD.

Fat-free mass and basal metabolic rate were new risk factors for AD. The direct evidence on the associations between these 2 factors and AD is limited. One case-control study observed that loss of lean mass was associated with brain atrophy and cognitive performance (33), and a recent cohort study found that decreased lean mass was an indicator of increased all-cause dementia risk in older adults (34). Few studies on AD has involved basal metabolic rate, partially because the variable was not commonly measured in research. In our study, significant associations identified in PheWAS indicated that they shared strong genetic background with AD and time-to-event analysis confirmed their clinical association with disease occurrence. We also ran sensitivity analyses, and the significant results remained after adjusting for covariates (Supplemental Data Tables S15 and S17 in Supplement 2). Evidence from MR further strengthened their associations with AD, and imaging association analyses suggested that they influenced the risk of AD by altering AD-specific brain areas. Future research should investigate whether preventing fat-free mass loss and maintaining the basal metabolic rate in older adults reduces AD risk.

Our study also adds to the body of evidence on the link between computer use and AD, which is lacking in the literature. We thought the association was likely driven by the association between education and AD. However, significant results remained after we adjusted for covariates such as education level (Supplemental Data Tables S15 and S17 in Supplement 2). One study previously observed less computer use in cognitively impaired individuals compared with the cognitively unimpaired individuals (35), and less computer use was related to smaller hippocampal volumes and worse cognitive performance (35,36), which partly supported our findings. One explanation is that computer use may act as a form of cognitive training to maintain cognitive function and prevent the onset of AD (37). Because our findings were consistent across all 4 analytic steps, further investigations are warranted to assess its role in AD prevention as a potentially modifiable risk factor.

Novel findings on associations with other lifestyles were also supported by strong evidence from our research. Significant results from PheWAS, time-to-event, and imaging analysis consistently showed that people who had no selective

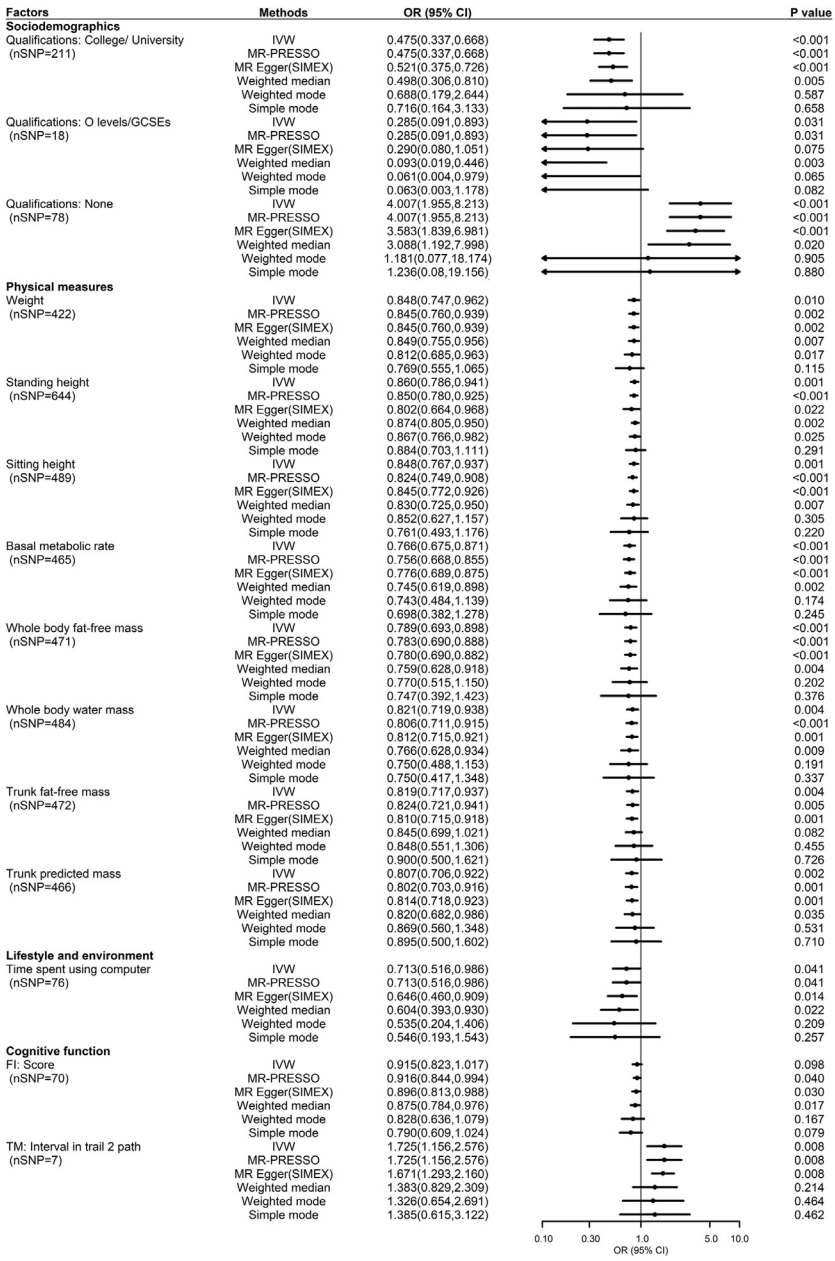



**Figure 4.** MR estimates of factors in relation to AD risk. Major factors (14 of 21) found to be associated with AD in MR analysis are presented here. Full results can be found in Supplemental Data Table S6 in Supplement 2, and details can be found in Figures S6–S42 in Supplement 1. AD, Alzheimer's disease; FI, fluid intelligence; IVW, inverse variance weighted; MR, Mendelian randomization; nSNP, number of single nucleotide polymorphisms; O levels/GCSEs, General Certificate of Secondary Education Ordinary levels; OR, odds ratio; PRESSO, pleiotropy residual sum and outlier; TM, Trail Making.

intake of eggs, dairy, wheat, sugar, or products containing them had lower risks of AD. In contrast, people who never ate sugar had higher risks of AD. Both results supplement the current dietary research on AD (2,38) that a balanced diet cannot be neglected. Three analytic steps also consistently found that time spent outdoors was associated with increased risk of AD. This was indirectly supported by a recent work also conducted by our team wherein time spent outdoors was associated with elevated risks of developing all-cause dementia after adjusting for a series of covariates (37). We noted that associations between these factors and AD were not bolstered by MR. As estimates of MR represent lifelong average effects of genetic variants, we think that MR may not be interpreted in the same way as those from a relatively briefer life period and that absence of MR support does not refute the potential importance of the factors (39). Considering that those factors were not previously identified and they have mild-to-moderate effects, it is worthwhile to thoroughly investigate their roles in AD in separate studies.

One may find that the factors' effects on brain structures were relatively small in our imaging analysis. This was consistent with recent findings in brain-wide association studies that the effects of brain–phenotype associations were smaller than previously thought. It was estimated that the top

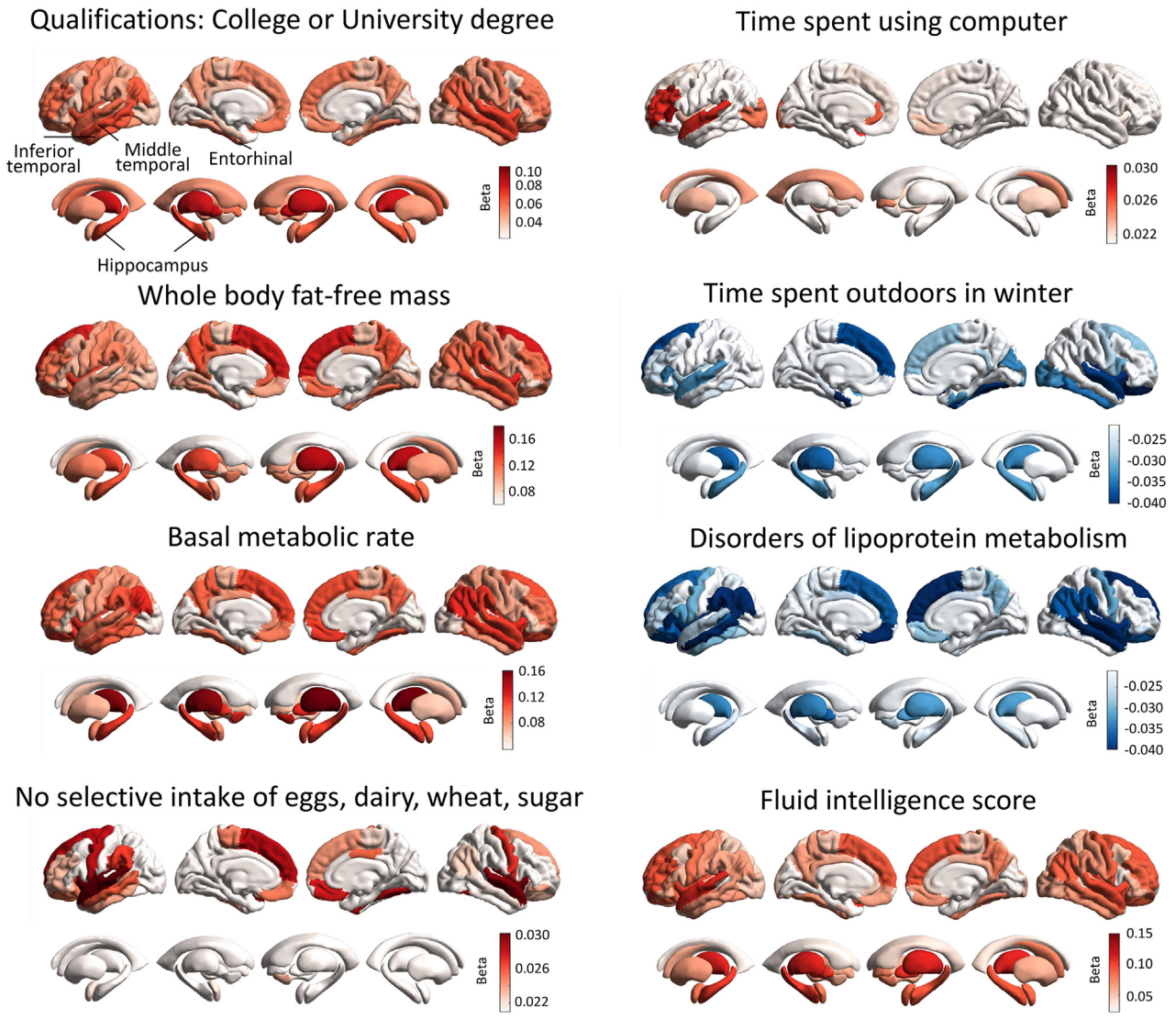



**Figure 5.** Significant associations between factors and AD-related regions. Of 51 factors, 39 showed consistent associations between cortical and subcortical regions, of which the 8 with largest effect are presented. AD-related regions are marked on the left hemisphere. The color bars indicate the effect size of factors on brain regions. Associations between 51 factors and AD-related regions are shown in the heatmap in Figure S3 in Supplement 1. Statistic details can be found in Supplemental Data Tables S9–S12 in Supplement 2. AD, Alzheimer's disease.

1% largest of all possible brain-wide associations only reached a standard correlation (bivariate linear $|r|$) value $> 0.06$ (40). Thus, small effect sizes did not imply null findings, and factors may be associated with AD via mechanisms other than altering the morphology of brain structures.

In PheWAS, we noticed a differential association between multiple categories of risk factors. More than 10% of cognitive phenotypes and physical measures were significantly associated with AD, likely reflecting their overlapping genetic architecture with AD. In contrast, phenotypes of other categories such as lifestyle and environment and medications and operations showed much lower proportions of significance, suggesting that AD shares its genetic architecture with only a small proportion of them. One potential reason for this discrepancy is a relative lack of signal in GWASs of behavioral variables such as lifestyle, which contrasts with endophenotypes such as cognitive and physical conditions. When putting phenotypes of various categories together, the results for endophenotypes will most likely come out on top.

A novel concept provides new insights into PheWAS. Desikan *et al.* (41) recently developed and validated a novel PRS, the polygenic hazard score, which was found to predict the onset age of AD dementia. A higher polygenic hazard score was linked to faster clinical decline, faster atrophy of AD-specific brain regions, and more amyloid/tau deposition (42). As a result, we believe this score represents the genetic architecture of AD progression rather than the PRS, which only represents the risk of AD. Future research could use polygenic hazard score to conduct a phenome-wide scan and identify factors that share genetic backgrounds with AD progression.

This may reveal different factor profiles from this study, providing more clues on disease-modifying therapies and effective intervention.

Several limitations should be considered when interpreting our findings. First, our PheWAS was restricted by the available variables from the UKB database. Variables such as air pollution and blood glucose measures were not included, although they have been previously found to be associated with AD (43–46). Second, because PheWAS looks for phenotypes that are associated with genetic architecture of AD, our study was less likely to uncover AD risk factors that have no or a weak genetic link to the disease. This limitation was highlighted by the current PheWAS's stringent filtering criteria. They may obscure certain associations that are potentially noteworthy if studied individually. In our PheWAS, for example, sleep duration, daytime naps, and time spent watching television were all significantly associated with AD-PRS in at least 2 *p* thresholds, but due to conservative thresholds, they were excluded from further analysis. Thus, more nuanced analyses are required. Third, although we performed additional analysis adjusting for covariates such as education and body mass index, we were unable to triangulate the relationships between all the identified factors. Future studies may seek to distinguish between these factors' independent and mediation effects to decipher potential pathways that underpin the associations.

In conclusion, leveraging phenotypic and genomic data from more than 500,000 individuals in UKB, we used a novel four-step approach to systematically screen and rigorously assess associations of a broad array of risk factors with AD and found strong evidence for fat-free body mass, basal metabolic rate, computer use, selective food intake, and outdoor exposures as new risk factors for AD. Integration of genetic, clinical, and neuroimaging information may help prioritize risk factors and prevention targets for AD.

## ACKNOWLEDGMENTS AND DISCLOSURES

This study was supported by grants from the Science and Technology Innovation 2030 Major Projects (Grant No. 2022ZD0211600 [to J-TY]), National Natural Science Foundation of China (Grant Nos. 82071201 [to J-TY], 81971032 [to J-TY], and 82071997 [to WC]), Shanghai Municipal Science and Technology Major Project (Grant No. 2018SHZDZX01 [to J-TY and J-FF]), Shanghai Rising-Star Program (Grant No. 21QA1408700 [to WC]), Research Startup Fund of Huashan Hospital (Grant No. 2022QD002 [to J-TY]), Excellence 2025 Talent Cultivation Program at Fudan University (Grant No. 3030277001 [to J-TY]), Shanghai Center for Brain Science and Brain-Inspired Technology (to J-FF), and Zhangjiang Laboratory, Tianqiao and Chrissy Chen Institute, and the State Key Laboratory of Neurobiology and Frontiers Center for Brain Science of Ministry of Education, Fudan University (to J-TY).

The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

J-TY had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. J-TY, WC, and QD contributed to concept and design; S-DC, WZ, and Y-ZL contributed to drafting of the manuscript; S-DC, WZ, Y-ZL, and WC performed statistical analysis; J-TY, WC, and J-FF obtained funding; WC, J-FF, QD, and J-TY provided administrative, technical, or material support; J-FF, WC, J-TY, and QD supervised the study; all authors contributed to acquisition, analysis, or interpretation of data, and to critical revision of the manuscript for important intellectual content; and all authors read and approved the final manuscript.

This study is based on publicly available data with different levels of accessibility. The UK Biobank (UKB) and individual studies within each GWAS received approval from an appropriate institutional review board, and informed consent was obtained from participants or from a caregiver, legal guardian, or other proxy. The data used in our study are from the UKB with restrictions applied. Data are used under license and thus are not publicly available. Access to the UKB data can be requested through a standard protocol (https://www.ukbiobank.ac.uk/register-apply/). The summary statistics of AD-GWASs for polygenic risk score calculation is available at http://www.pasteur-lille.fr/en/recherche/u744/Igap_stage1.zip. GWAS summary statistics used in MR are detailed in the Supplement.

Scripts used to perform the analyses are available at https://github.com/HaloForest/UKB_PWAS.

Data acquisition and analyses were conducted using the UKB Resource under approved project No. 19542. Participation of the UKB participants is gratefully appreciated. We also thank the UKB team for collecting and preparing data for analyses. This work was made possible by the generous sharing of GWAS summary statistics. We thank the Integrative Epidemiology Unit GWAS database project as well as the individuals and institutes joining this project for offering the summary statistics of the GWAS data. We also thank the Alzheimer's Disease Genetics Consortium (ADGC); the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium (CHARGE); the European Alzheimer's Disease Initiative (EADI); the Genetic and Environmental Risk in AD/Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium (GERAD/PERADES); the Alzheimer's Disease Sequencing Project (ADSP); the Alzheimer's disease working group of the Psychiatric Genomics Consortium (PGC-ALZ); and the International Genomics of Alzheimer's Project (IGAP) for providing summary statistics for these analyses. The investigators within IGAP contributed to the design and implementation of IGAP and/or provided data but did not participate in analysis or writing of this report. IGAP was made possible by the generous participation of the control participants, 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 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 was supported by the Medical Research Council (Grant No. 503480), Alzheimer's Research UK (Grant No. 503176), the Wellcome Trust (Grant No. 082604/2/07/Z), and the German Federal Ministry of Education and Research: Competence Network Dementia grant (Grant Nos. 01GI0102, 01GI0711, and 01GI0420). CHARGE was partly supported by the National Institutes of Health/National Institute on Aging (NIA) grant (Grant No. R01 AG033193) and the NIA AG081220 and AGES contract N01–AG–12100, the National Heart, Lung, and Blood Institute (NHLBI) grant (Grant No. R01 HL105756), the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA grants (Grant Nos. U01 AG032984, U24 AG021886, and U01 AG016976), and the Alzheimer's Association grant (Grant No. ADGC–10–196728). The ADSP comprises 2 AD genetics consortia and 3 National Human Genome Research Institute–funded Large Scale Sequencing and Analysis Centers (LSAC). The 2 AD genetics consortia are the ADGC funded by NIA (Grant No. U01 AG032984) and the CHARGE funded by NIA (Grant No. R01 AG033193), the NHLBI, other NIH institutes, and other foreign governmental and nongovernmental organizations.

J-TY serves as Editor-in-Chief for *Brain Disorders*, Associate Editor-in-Chief for *Annals of Translational Medicine*, and Senior Editor for *Journal of Alzheimer's Disease*. All other authors report no biomedical financial interests or potential conflicts of interest.

## ARTICLE INFORMATION

From the 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, National Center for Neurological Disorders, Shanghai, China (S-DC, LY, Y-YH, Y-TD, B-SW, WC, QD, J-TY); Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China (WZ, Y-ZL, ETR, J-FF, WC, J-TY); Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence, Fudan University,

Ministry of Education, Shanghai, China (WZ, Y-ZL, J-FF, WC); Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom (JS); Oxford Centre for Computational Neuroscience, Oxford, United Kingdom (ETR); the Department of Computer Science, University of Warwick, Coventry, United Kingdom (ETR); Fudan ISTBI—ZJNU Algorithm Centre for Brain-Inspired Intelligence, Zhejiang Normal University, Jinhua, China (J-FF, WC); MOE Frontiers Center for Brain Science, Fudan University, Shanghai, China (J-FF); and Zhangjiang Fudan International Innovation Center, Shanghai, China (J-FF).

S-DC, WZ, and Y-ZL contributed equally to this work.

Address correspondence to Jin-Tai Yu, M.D., Ph.D., at jintai_yu@fudan.edu.cn, or Qiang Dong, M.D., Ph.D., at dong_qiang@fudan.edu.cn.

Received Mar 25, 2022; revised Jun 15, 2022; accepted Aug 5, 2022.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.biopsych.2022.08.002.

## REFERENCES

1. GBD 2016 Dementia Collaborators (2019): Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 18:88–106.
2. Stevenson JS, Clifton L, Kuźma E, Littlejohns TJ (2022): Speech-in-noise hearing impairment is associated with an increased risk of incident dementia in 82,039 UK Biobank participants. Alzheimers Dement 18:445–456.
3. Malik R, Georgakis MK, Neitzel J, Rannikmäe K, Ewers M, Seshadri S, *et al.* (2021): Midlife vascular risk factors and risk of incident dementia: Longitudinal cohort and Mendelian randomization analyses in the UK Biobank. Alzheimers Dement 17:1422–1431.
4. Scheltens P, De Strooper B, Kivipelto M, Holstege H, Chételat G, Teunissen CE, *et al.* (2021): Alzheimer's disease. Lancet 397:1577–1590.
5. Schwartzentruber J, Cooper S, Liu JZ, Barrio-Hernandez I, Bello E, Kumasaka N, *et al.* (2021): Genome-wide meta-analysis, fine-mapping and integrative prioritization implicate new Alzheimer's disease risk genes. Nat Genet 53:392–402.
6. Bellenguez C, Küçükali F, Jansen IE, Kleineidam L, Moreno-Grau S, Amin N, *et al.* (2022): New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat Genet 54:412–436.
7. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, *et al.* (2019): 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.
8. Mukherjee S, Walter S, Kauwe JSK, Saykin AJ, Bennett DA, Larson EB, *et al.* (2015): Genetically predicted body mass index and Alzheimer's disease-related phenotypes in three large samples: Mendelian randomization analyses. Alzheimers Dement 11:1439–1451.
9. Andrews SJ, Goate A, Anstey KJ (2020): Association between alcohol consumption and Alzheimer's disease: A Mendelian randomization study. Alzheimers Dement 16:345–353.
10. Axelrud LK, Santoro ML, Pine DS, Talarico F, Gadelha A, Manfro GG, *et al.* (2018): Polygenic risk score for Alzheimer's disease: Implications for memory performance and hippocampal volumes in early life. Am J Psychiatry 175:555–563.
11. Richardson TG, Harrison S, Hemani G, Davey Smith G (2019): An atlas of polygenic risk score associations to highlight putative causal relationships across the human phenome. eLife 8.
12. Milaneschi Y, Lamers F, Peyrot WJ, Abdellaoui A, Willemsen G, Hottenga JJ, *et al.* (2016): Polygenic dissection of major depression clinical heterogeneity. Mol Psychiatry 21:516–522.
13. Lawlor DA, Tilling K, Davey Smith G (2016): Triangulation in aetiological epidemiology. Int J Epidemiol 45:1866–1886.
14. Miller KL, Alfaro-Almagro F, Bangerter NK, Thomas DL, Yacoub E, Xu J, *et al.* (2016): Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat Neurosci 19:1523–1536.
15. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, *et al.* (2018): The UK biobank resource with deep phenotyping and genomic data. Nature 562:203–209.
16. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, *et al.* (2013): Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 45:1452–1458.
17. Vilhjálmsson BJ, Yang J, Finucane HK, Gusev A, Lindström S, Ripke S, *et al.* (2015): Modeling linkage disequilibrium increases accuracy of polygenic risk scores. Am J Hum Genet 97:576–592.
18. Shen X, Howard DM, Adams MJ, Hill WD, Clarke TK, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, et al (2020): A phenome-wide association and Mendelian Randomisation study of polygenic risk for depression in UK Biobank. Nat Commun 11:2301.
19. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, *et al.* (2006): An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31:968–980.
20. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, *et al.* (2002): Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron 33:341–355.
21. Millard LAC, Davies NM, Gaunt TR, Davey Smith G, Tilling K (2018): Software Application Profile: PHESANT: A tool for performing automated phenome scans in UK Biobank. Int J Epidemiol 47:29–35.
22. Halliday G (2017): Pathology and hippocampal atrophy in Alzheimer's disease. Lancet Neurol 16:862–864.
23. Mu Y, Gage FH (2011): Adult hippocampal neurogenesis and its role in Alzheimer's disease. Mol Neurodegener 6:85.
24. Bobinski M, de Leon MJ, Convit A, De Santi S, Wegiel J, Tarshish CY, *et al.* (1999): MRI of entorhinal cortex in mild Alzheimer's disease. Lancet 353:38–40.
25. Gómez-Isla T, Price JL, McKeel DW Jr, Morris JC, Growdon JH, Hyman BT (1996): Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. J Neurosci 16:4491–4500.
26. Convit A, de Asis J, de Leon MJ, Tarshish CY, De Santi S, Rusinek H (2000): Atrophy of the medial occipitotemporal, inferior, and middle temporal gyri in non-demented elderly predict decline to Alzheimer's disease. Neurobiol Aging 21:19–26.
27. Scott MR, Hampton OL, Buckley RF, Chhatwal JP, Hanseeuw BJ, Jacobs HI, *et al.* (2020): Inferior temporal tau is associated with accelerated prospective cortical thinning in clinically normal older adults. Neuroimage 220:116991.
28. Scheff SW, Price DA, Schmitt FA, Scheff MA, Mufson EJ (2011): Synaptic loss in the inferior temporal gyrus in mild cognitive impairment and Alzheimer's disease. J Alzheimers Dis 24:547–557.
29. Norton S, Matthews FE, Barnes DE, Yaffe K, Brayne C (2014): Potential for primary prevention of Alzheimer's disease: An analysis of population-based data. Lancet Neurol 13:788–794.
30. Stern Y (2012): Cognitive reserve in ageing and Alzheimer's disease. Lancet Neurol 11:1006–1012.
31. Andrews SJ, Fulton-Howard B, O'Reilly P, Marcora E, Goate AM, collaborators of the Alzheimer's Disease Genetics Consortium (2021): Causal Associations Between Modifiable Risk Factors and the Alzheimer's Phenome. Ann Neurol 89:54–65.
32. Calvin CM, Wilkinson T, Starr JM, Sudlow C, Hagenaars SP, Harris SE, *et al.* (2019): Predicting incident dementia 3–8 years after brief cognitive tests in the UK Biobank prospective study of 500,000 people. Alzheimers Dement 15:1546–1557.
33. Burns JM, Johnson DK, Watts A, Swerdlow RH, Brooks WM (2010): Reduced lean mass in early Alzheimer disease and its association with brain atrophy. Arch Neurol 67:428–433.
34. Cui C, Mackey RH, Shaaban CE, Kuller LH, Lopez OL, Sekikawa A (2020): Associations of body composition with incident dementia in older adults: Cardiovascular Health Study-Cognition Study. Alzheimers Dement 16:1402–1411.
35. Silbert LC, Dodge HH, Lahna D, Promjunyakul NO, Austin D, Mattek N, *et al.* (2016): Less daily computer use is related to smaller hippocampal volumes in cognitively intact elderly. J Alzheimers Dis 52:713–717.
36. Bernstein JPK, Dorociak KE, Mattek N, Leese M, Beattie ZT, Kaye JA, Hughes A (2021): Passively measured routine home computer activity and application use can detect mild cognitive impairment and

correlate with important cognitive functions in older adulthood. J Alzheimers Dis 81:1053–1064.

37. Ma LZ, Ma YH, Ou YN, Chen SD, Yang L, Dong Q, *et al.* (2022): Time spent in outdoor light is associated with the risk of dementia: A prospective cohort study of 362094 participants. BMC Med 20:132.
38. McEvoy CT, Hoang T, Sidney S, Steffen LM, Jacobs DR Jr, Shikany JM, *et al.* (2019): Dietary patterns during adulthood and cognitive performance in midlife: The CARDIA study. Neurology 92:e1589–e1599.
39. Choi KW, Stein MB, Nishimi KM, Ge T, Coleman JRI, Chen CY, *et al.* (2020): An exposure-wide and Mendelian randomization approach to identifying modifiable factors for the prevention of depression. Am J Psychiatry 177:944–954.
40. Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, *et al.* (2022): Reproducible brain-wide association studies require thousands of individuals. Nature 603:654–660.
41. Desikan RS, Fan CC, Wang Y, Schork AJ, Cabral HJ, Cupples LA, *et al.* (2017): Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score. PLoS Med 14:e1002258.
42. Tan CH, Bonham LW, Fan CC, Mormino EC, Sugrue LP, Broce IJ, *et al.* (2019): Polygenic hazard score, amyloid deposition and Alzheimer's neurodegeneration. Brain 142:460–470.
43. Calderón-Garcidueñas L, Villarreal-Ríos R (2017): Living close to heavy traffic roads, air pollution, and dementia. Lancet 389:675–677.
44. Zhang X, Chen X, Zhang X (2018): The impact of exposure to air pollution on cognitive performance. Proc Natl Acad Sci U S A 115:9193–9197.
45. Crane PK, Walker R, Hubbard RA, Li G, Nathan DM, Zheng H, *et al.* (2013): Glucose levels and risk of dementia. N Engl J Med 369:540–548.
46. de Matos AM, de Macedo MP, Rauter AP (2018): Bridging Type 2 diabetes and Alzheimer's disease: Assembling the puzzle pieces in the quest for the molecules with therapeutic and preventive potential. Med Res Rev 38:261–324.