**Supplementary 1**

**Supplementary Index**

**Supplementary 1**

**Supplementary Methods**

**Supplementary Method 1.** Participants

**Supplementary Method 2.** Covariates

**Supplementary Method 3.** Statistical analysis: Mendelian randomization

**Supplementary References**

**Supplementary Figures**

**eFigure 1.** Sensitivity analysis of disease trajectories before ACD.

**eFigure 2.** Disease trajectories before dementia stratified by subtypes.

**eFigure 3.** Disease trajectories before ACD stratified by sex.

**e****Figure 4.** Disease trajectories before ACD stratified by median age of dementia diagnosis.

**eFigure 5.** Disease trajectory clusters before ACD among dementia patients aged 65 or above.

**eFigure 6.** Disease trajectories before ACD in *ApoE* ε4 non-carriers.

**eFigure 7.** Results of mendelian randomization on associations of lipoprotein disorders and related biomarkers with AD.

**Supplementary 2**

**eTable 1.** List of ICD-10 codes.

**eTable 2.** Baseline characteristics of the study population.

**eTable 3.** Associations between exposure diseases and ACD.

**eTable 4.** Significant disease pairs before ACD.

**eTable 5.** Significant disease pairs in disease trajectory before ACD.

**eTable 6.** Significant disease pairs before AD.

**eTable 7.** Significant disease pairs before VaD.

**eTable 8.** Primary and sensitivity analysis of significant disease pairs in trajectories before VaD.

**eTable 9.** Significant disease pairs before ACD among male.

**eTable 10.** Significant disease pairs before ACD among female.

**eTable 11.** Significant disease pairs before ACD in people among dementia patients aged 74 or above.

**e****Table 12.** Significant disease pairs before ACD among dementia patients aged under 74.

**eTable 13.** Significant disease pairs before ACD among dementia patients aged 65 or above.

**eTable 14.** Significant disease pairs before ACD among *ApoE* ε4 carrier.

**eTable 15.** Significant disease pairs before ACD among *ApoE* ε4 non-carrier.

**eTable 16.** Genetic instruments for E78 and related biochemical indexes used in mendelian randomization analysis.

**e****Table 17.** Basic statistics for mendelian randomization analysis.

**Supplementary Methods**

**Participants**

For this study, we excluded 19 participants from totally 502,389 participants who withdrew the informed consent, Non-White British ethnic participants (N = 59,854), and participants with missing Townsend deprivation index (TDI) at recruitment (N = 772), a crucial index for subsequent analysis. We further excluded dementia patients recorded before baseline (N = 2,487), with missing record dates (N = 3), or only self-reported sources (N = 2) from the analysis. Participants died within six months after enrollment were also excluded (N = 348).

**Covariates**

Several variables measured at baseline were considered as covariates for our analysis. Socio-demographic covariates included ethnic background (Field ID 21000), sex (Field ID 31), family history of dementia (Field ID 20107, 20110, and 20111), age (Field ID 21022), and TDI at recruitment (Field ID 22189). The apolipoprotein E (*ApoE*; rs7412 and rs429358) was genotyped on the UKB arrays and classified as ε4 carrier versus non-carriers.1

**Statistical analysis**

**Mendelian randomization**

Mendelian randomization (MR) is a powerful tool to suggest causality for associations found in observational studies.2 In MR analysis, Genome-wide association analysis (GWAS) summary data were extracted from the Medical Research Council Integrative Epidemiology Unit OpenGWAS database (https://gwas.mrcieu.ac.uk/). Due to the lack of widely used GWAS data on ACD or VaD, summary data from a recent GWAS meta-analysis on AD3 was used as the outcome dataset. The MR pleiotropy residual and outlier (MR-PRESSO) was employed to identify and exclude potential outlier SNPs in the instrumental variable (IV) set. Cochran's Q statistics assessed the heterogeneity, while MR-Egger devaluated the horizontal pleiotropy of IVs. We adopted several distinct MR methods with different assumptions, including inverse variance weighting (IVW), MR-Egger, weighted median, weighted mode, contamination mixture,4 along with constrained maximum likelihood and model averaging and Bayesian information criterion (cML-MA-BIC).2,5 MR analysis was conducted using *TwoSampleMR*, *MRcML* and *MendelianRandomization* packages.

**Supplementary References**

1. Lyall DM, Ward J, Ritchie SJ, et al. Alzheimer disease genetic risk factor APOE e4 and cognitive abilities in 111,739 UK Biobank participants. *Age Ageing* 2016; **45**(4): 511-7.

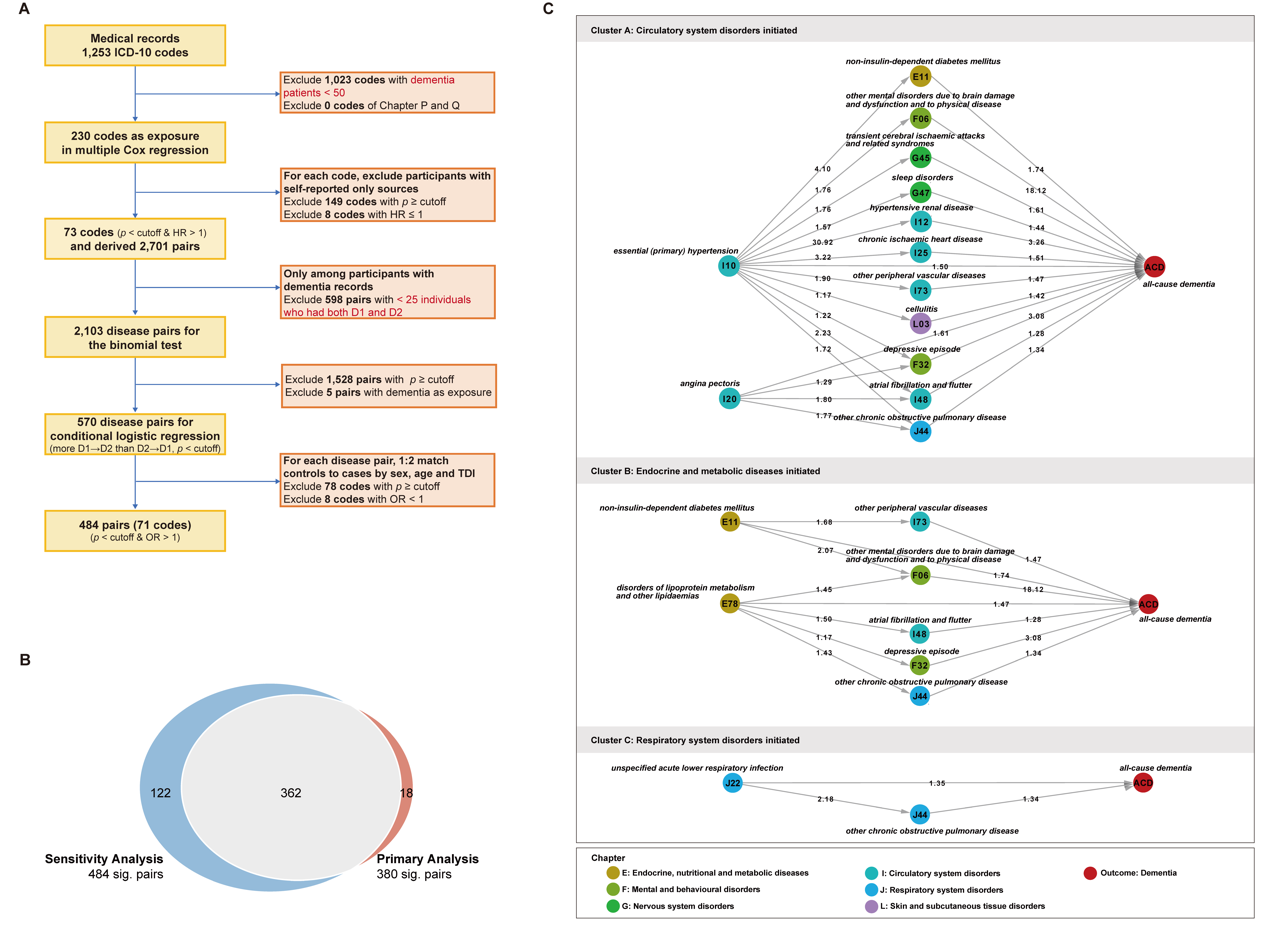
2. Burgess S, Davey Smith G, Davies NM, et al. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res* 2019; **4**: 186.

3. Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet* 2019; **51**(3): 414-30.

4. Slob EAW, Burgess S. A comparison of robust Mendelian randomization methods using summary data. *Genet Epidemiol* 2020; **44**(4): 313-29.

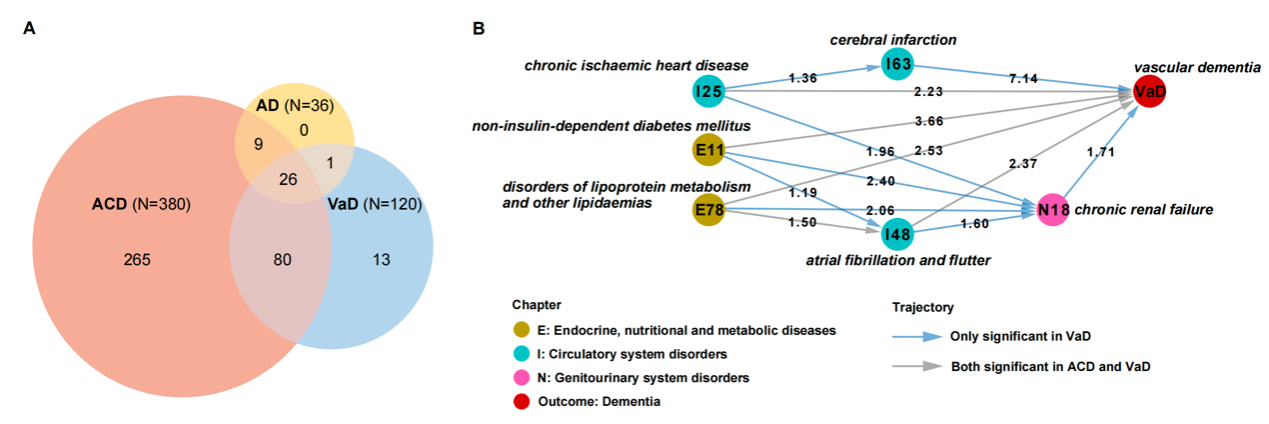
5. Xue H, Shen X, Pan W. Constrained maximum likelihood-based Mendelian randomization robust to both correlated and uncorrelated pleiotropic effects. *Am J Hum Genet* 2021; **108**(7): 1251-69.

**Supplementary Figures**

****

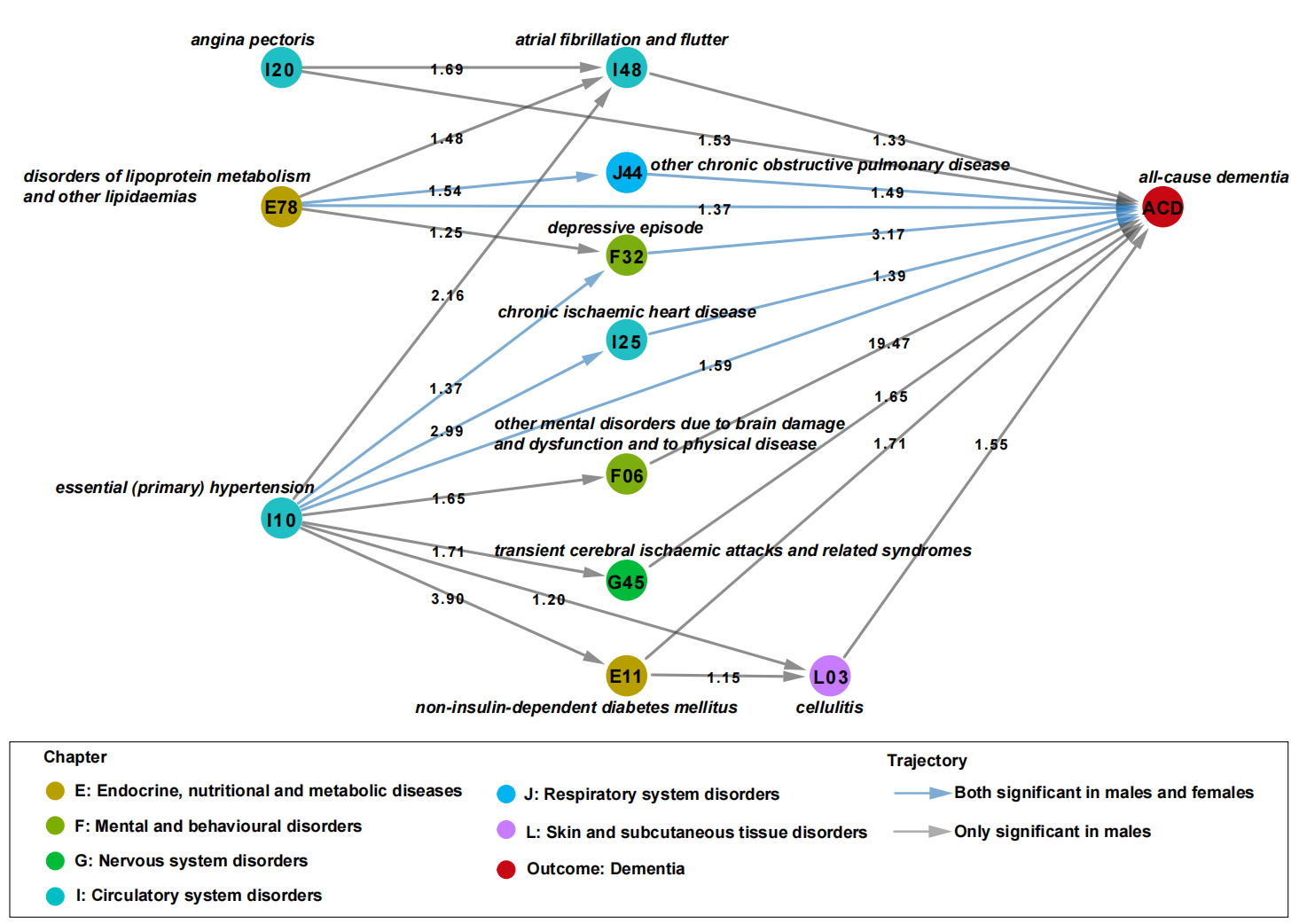
**eFigure 1. Sensitivity analysis of disease trajectories before ACD. (A)** Flowchart of diseases selection; **(B)** Comparison of the number of significant disease pairs in the primary and sensitivity analyses; **(C)** Disease trajectories before ACD diagnosis in the sensitivity analysis. The numbers on the arrows indicate the odds ratios of the association between disease exposure and ACD.

Abbreviation: ACD, all-cause dementia; ICD, International Classification of Diseases.



**eFigure 2. Disease trajectories before dementia stratified by subtypes.** **(A)** the number of diseases significantly associated with an increased risk of ACD, AD, and VaD, respectively; **(B)** the disease trajectories before VaD diagnosis. The numbers on the arrows indicate the odds ratios of the association between disease exposure and VaD.

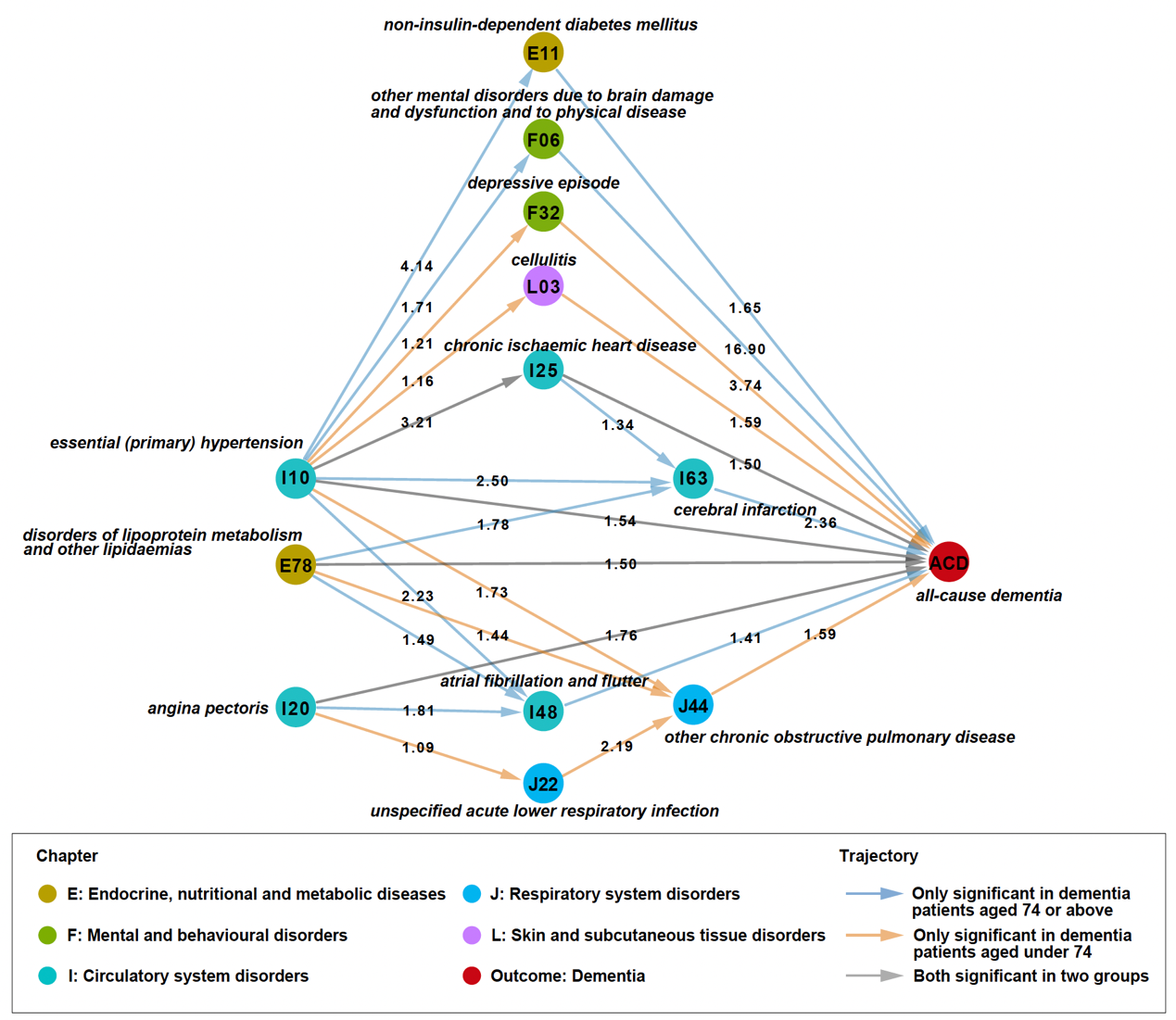
Abbreviation: ACD, all-cause dementia; AD, Alzheimer's disease; VaD, vascular dementia.



**eFigure 3**. **Disease trajectories before ACD stratified by sex.**

The numbers on the arrows indicate the odds ratios of the disease exposure for ACD in males.

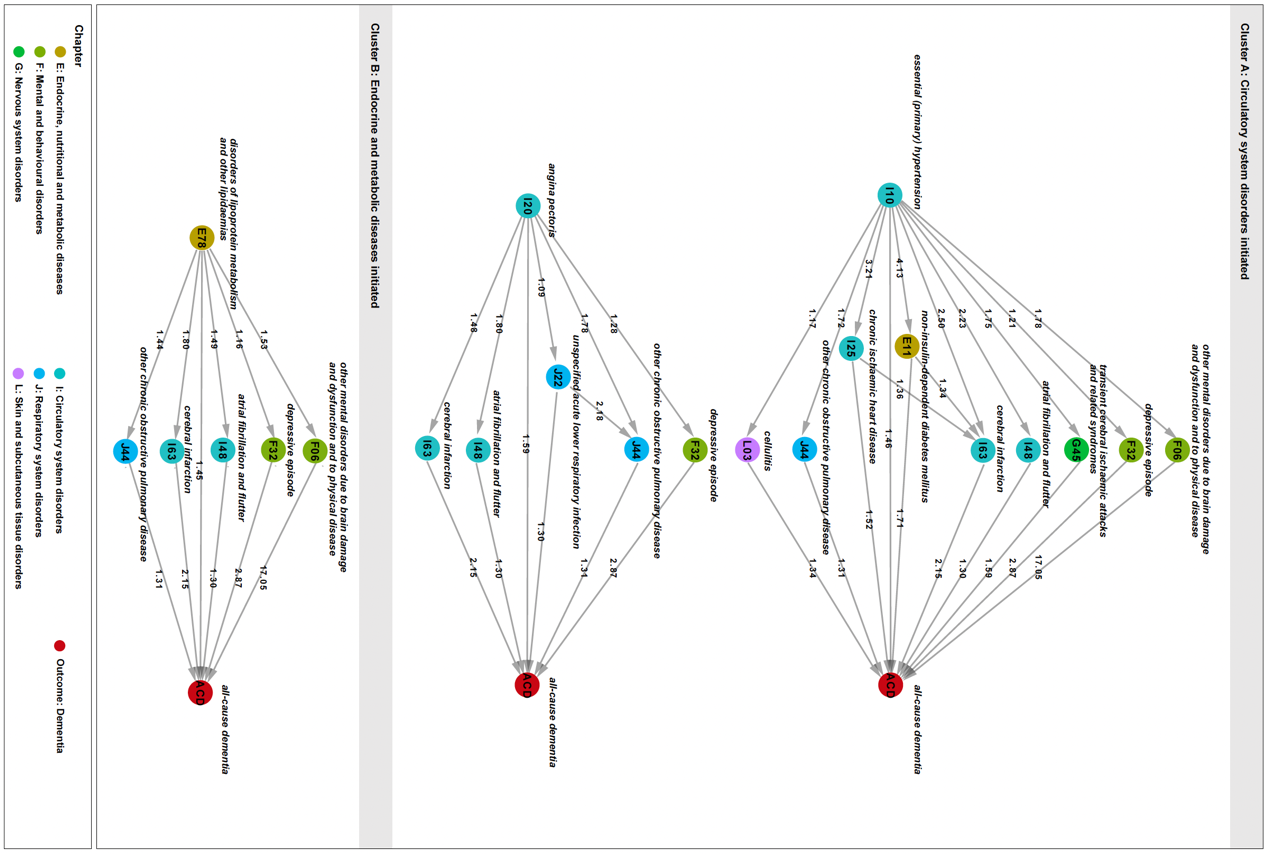
Abbreviation: ACD, all-cause dementia.



**eFigure 4**. **Disease trajectories before ACD stratified by median age of dementia diagnosis.**

The numbers on the arrows indicate the odds ratios of the disease exposure for ACD. For both significant pairs in two groups, the odds ratio in dementia patients aged 74 or above was showed.

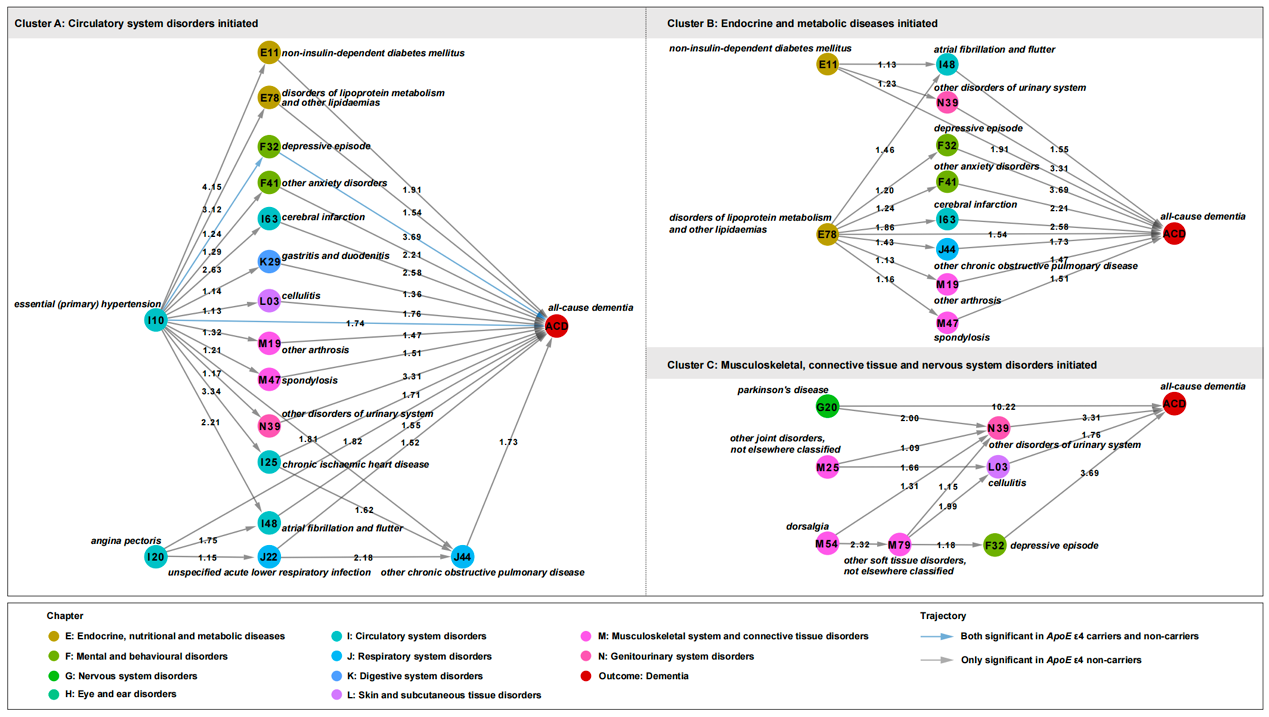
Abbreviation: ACD, all-cause dementia.

****

**eFigure 5. Disease trajectory clusters before ACD among dementia patients aged 65 or above.**

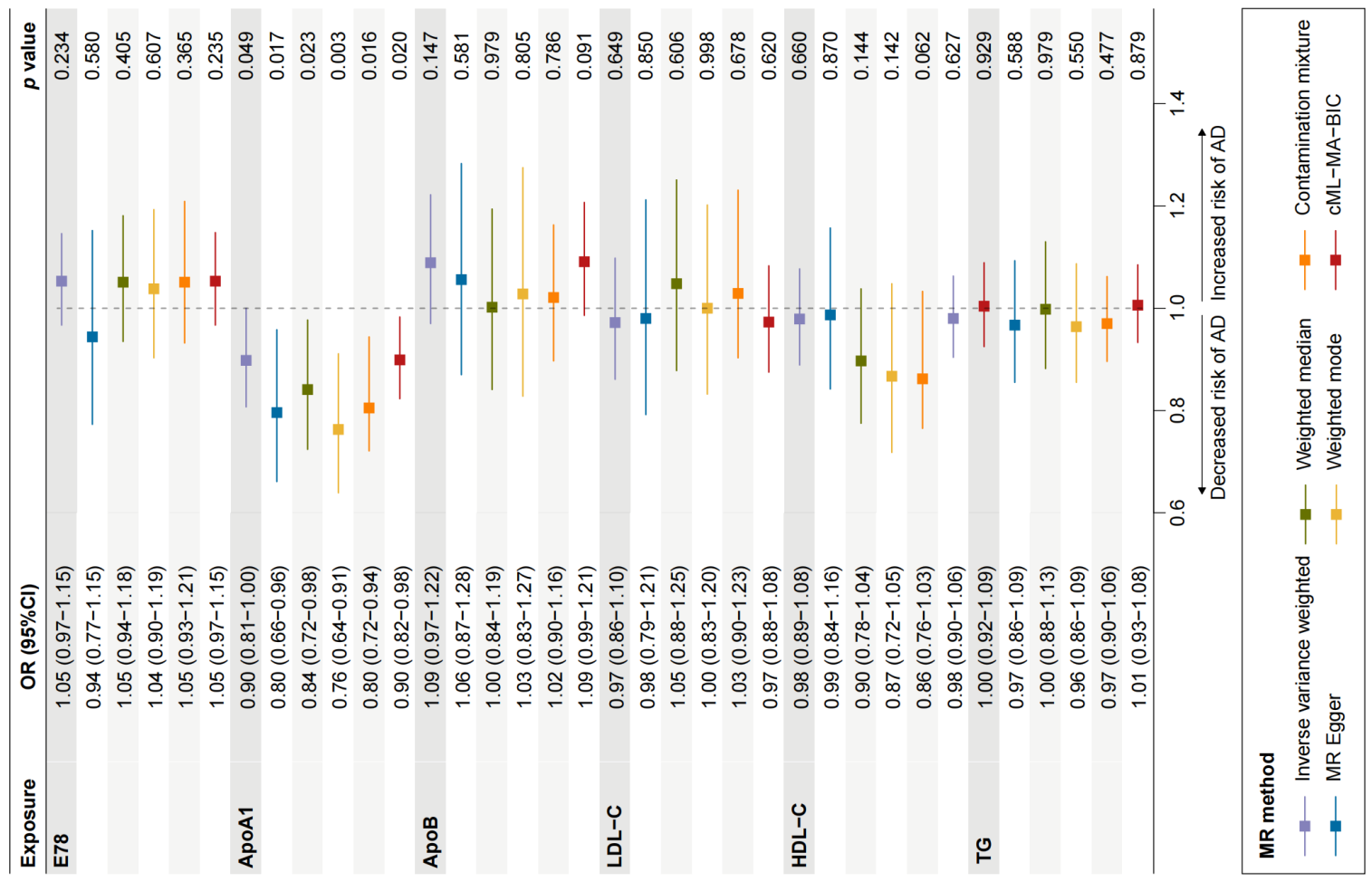
The numbers on the arrows indicate the odds ratios of the association between disease exposure and ACD.

Abbreviation: ACD, all-cause dementia.



**eFigure 6. Disease trajectories before ACD diagnosis in** ***ApoE* ε4 non-carriers.**

The numbers on the arrows indicate the odds ratios of the association between disease exposure and ACD in *ApoE* ε4 non-carriers. Abbreviation: ACD, all-cause dementia, *ApoE*, apolipoprotein E. Abbreviation: ACD, all-cause dementia.



**eFigure 7. Results of mendelian randomization on associations of lipoprotein disorders and related biomarkers with AD.**

Abbreviation: AD, Alzheimer's disease; Apo, apolipoprotein; cML-MA-BIC, constrained maximum likelihood and model averaging and Bayesian information criterion; E78: ICD-10 code of disorders of lipoprotein metabolism and other lipidaemias; HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol; MR, mendelian randomization; OR, odds ratio; TG, triglyceride.