All rights reserved. No reuse allowed without permission.

# Polygenic effects on the risk of Alzheimer's disease in the Japanese population

- 3
- 4 Masataka Kikuchi<sup>1,2</sup>, Akinori Miyashita<sup>3</sup>, Norikazu Hara<sup>3</sup>, Kensaku Kasuga<sup>3</sup>,
- 5 Yuko Saito<sup>4</sup>, Shigeo Murayama<sup>4,5</sup>, Akiyoshi Kakita<sup>6</sup>, Hiroyasu Akatsu<sup>7</sup>, Kouichi
- 6 Ozaki<sup>8,9</sup>, Shumpei Niida<sup>10</sup>, Ryozo Kuwano<sup>11</sup>, Takeshi Iwatsubo<sup>12</sup>, Akihiro
- 7 Nakaya<sup>1</sup>, Takeshi Ikeuchi<sup>3</sup>, Alzheimer's Disease Neuroimaging Initiative<sup>#</sup>, the
- 8 Japanese Alzheimer's Disease Neuroimaging Initiative\*
- 9
- <sup>1</sup> Department of Computational Biology and Medical Sciences, Graduate School
- 11 of Frontier Science, The University of Tokyo, Chiba, Japan
- 12 <sup>2</sup> Department of Medical Informatics, Graduate School of Medicine, Osaka
- 13 University, Osaka, Japan
- 14 <sup>3</sup> Department of Molecular Genetics, Brain Research Institute, Niigata
- 15 University, Niigata, Japan
- 16 <sup>4</sup> Brain Bank for Aging Research (Department of Neuropathology), Tokyo
- 17 Metropolitan Institute of Geriatrics and Gerontology, Tokyo, Japan
- <sup>5</sup> Brain Bank for Neurodevelopmental, Neurological and Psychiatric Disorders,
- 19 United Graduate School of Child Development, Osaka University, Osaka, Japan
- <sup>6</sup> Department of Pathology, Brain Research Institute, Niigata University, Niigata,
- 21 Japan
- <sup>7</sup> Department of General Medicine & General Internal Medicine, Nagoya City
- 23 University Graduate School of Medicine, Nagoya, Japan
- 24 <sup>8</sup> Medical Genome Center, Research Institute, National Center for Geriatrics
- 25 and Gerontology, Aichi, Japan
- <sup>9</sup> RIKEN Center for Integrative Medical Sciences, Kanagawa, Japan
- <sup>10</sup> Core Facility Administration, Research Institute, National Center for Geriatrics
- 28 and Gerontology, Aichi, Japan
- 29 <sup>11</sup> Asahigawaso Research Institute, Social Welfare Corporation Asahigawaso,
- 30 Okayama, Japan
- <sup>12</sup> Department of Neuropathology, Graduate School of Medicine, The University
- 32 of Tokyo, Tokyo, Japan
- 33 <sup>#</sup> The data used in the preparation of this article were obtained from the
- 34 Alzheimer's Disease Neuroimaging Initiative (ADNI) (adni.loni.usc.edu). Thus,
- 35 the investigators within the ADNI contributed to the design and implementation
- 36 of the ADNI and/or provided data but did not participate in the analysis or the NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

All rights reserved. No reuse allowed without permission.

- 1 writing of this report. A complete listing of ADNI investigators can be found at
- 2 http://adni.loni.usc.edu/wp-content/uploads/how\_to\_
- 3 apply/ADNI\_Acknowledgement\_List.pdf.
- 4 \* Data used in preparation of this article were obtained from the Japanese
- 5 Alzheimer's Disease Neuroimaging Initiative (J-ADNI) database within the
- 6 National Bioscience Database Center Human Database, Japan (Research ID:
- 7 hum0043.v1, 2016). Thus, the investigators within J-ADNI contributed to the
- 8 design and implementation of J-ADNI and/or provided data but did not
- 9 participate in the analysis or the writing of this report. A complete listing of J-
- 10 ADNI investigators can be found at
- 11 https://humandbs.biosciencedbc.jp/en/hum0043-j-adni-authors.
- 12
- 13 Correspondence:
- 14 Masataka Kikuchi, PhD, Department of Computational Biology and Medical
- 15 Sciences, Graduate School of Frontier Science, The University of Tokyo
- 16 6-2-3 Kashiwanoha, Kashiwa, Chiba 277-0882, Japan
- 17 E-mail: kikuchi@edu.k.u-tokyo.ac.jp
- 18
- 19 Takeshi Ikeuchi, MD, PhD, Department of Molecular Genetics, Brain Research
- 20 Institute, Niigata University
- 21 1-757 Asahimachi, Niigata 951-8585, Japan
- 22 E-mail: ikeuchi@bri.niigata-u.ac.jp
- 23
- 24

#### All rights reserved. No reuse allowed without permission.

#### 1 Abstract

#### 2 Background

3 Polygenic effects have been proposed to account for some disease 4 phenotypes; these effects are calculated as a polygenic risk score (PRS). This 5 score is correlated with Alzheimer's disease (AD)-related phenotypes, such as 6 biomarker abnormalities and brain atrophy, and is associated with conversion 7 from mild cognitive impairment (MCI) to AD. However, the AD PRS has been 8 examined mainly in Europeans, and owing to differences in genetic structure 9 and lifestyle, it is unclear whether the same relationships between the PRS and 10 AD-related phenotypes exist in non-European populations. In this study, we 11 calculated and evaluated the AD PRS in Japanese individuals using GWAS 12 statistics from Europeans. 13 **Methods** 

14 In this study, we calculated the AD PRS in 504 Japanese participants 15 (145 cognitively unimpaired (CU) participants, 220 participants with late mild 16 cognitive impairment (MCI), and 139 patients with mild AD dementia) enrolled in 17 the Japanese Alzheimer's Disease Neuroimaging Initiative (J-ADNI) project. In 18 order to evaluate the clinical value of this score, we (1) determined the 19 polygenic effects on AD in the J-ADNI and validated it using two independent 20 cohorts (a Japanese neuropathology (NP) cohort (n=565) and the North 21 American ADNI (NA-ADNI) cohort (n=617)), (2) examined the AD-related 22 phenotypes associated with the PRS, and (3) tested whether the PRS helps 23 predict the conversion of MCI to AD.

#### 24 Results

25 The PRS using 131 SNPs had an effect independent of APOE. The 26 PRS differentiated between CU participants and AD patients with an area under the curve (AUC) of 0.755 when combined with the APOE variants. Similar AUC 27 28 was obtained when PRS calculated by the NP and NA-ADNI cohorts was 29 applied. In MCI patients, the PRS was associated with cerebrospinal fluid 30 phosphorylated-tau levels ( $\beta$  estimate = 0.235, p value = 0.026). MCI with a 31 high PRS showed a significantly increased conversion to AD in APOE ε4 32 noncarriers with a hazard rate of 2.22.

#### 33 Conclusions

We showed that the AD PRS is useful in the Japanese population,
whose genetic structure is different from that of the European population. These

All rights reserved. No reuse allowed without permission.

- 1 findings suggest that the polygenicity of AD is partially common across ethnic
- 2 differences.
- 3

## 4 Keywords:

5 Polygenic Risk Score, Alzheimer's Disease, Mild Cognitive Impairment

6

# 7 1. Background

8 Alzheimer's disease (AD) is a neurodegenerative disease caused by 9 environmental and genetic factors [1, 2]. Environmental factors, which are 10 acquired and modifiable, associated with AD include smoking status, alcohol 11 consumption, diet, and physical activity [3]. On the other hand, the heritability of 12 AD is approximately 70%, and genetic factors are inborn and nonmodifiable [4, 13 5]. However, knowing one's genetic risk early in life can motivate one to 14 improve modifiable factors. Indeed, sharing genetic test results with carriers of 15 genetic risk for disease may promote behavioural changes rather than increase 16 psychological distress [6, 7]. Thus, knowledge of the individual genetic risk of 17 AD is expected to contribute to delaying the onset of AD and early therapeutic 18 intervention.

The largest genetic risk factor for AD is the ε4 allele of the apolipoprotein E
(APOE) gene, but APOE ε4 explains only approximately 10% of AD cases
based on heritability [4, 5]. In addition, even when other AD-associated genetic
variants found in previous genome-wide association studies (GWAS) are also

considered, they do not explain all the genetic variance in AD patients [8],

24 suggesting the existence of additional unknown AD-related genetic variants. To

clarify this "missing heritability", polygenic effects that aggregate the small

26 effects of many alleles have been proposed to underlie AD.

Polygenic risk score (PRS) is a measure to quantify the combined effect of
genetic variants on an individual's risk for disease. The combination of the

29 APOE ε4 allele dose and PRS has been shown to improve disease prediction

30 accuracy in the European population [9]. Moreover, the PRS is associated with

31 AD-related phenotypes, such as brain volumes [10-12], brain amyloid-beta (Aβ)

32 burden [11, 12], and plasma phosphorylated tau [13], and has been reported to

be useful in predicting conversion from mild cognitive impairment (MCI) to AD[14, 15].

However, the clinical application of the PRS must be approached with caution. One of several concerns is that the effects of the PRS are not

All rights reserved. No reuse allowed without permission.

1 consistent across races [16, 17]. This is because genetic structures, such as 2 linkage disequilibrium (LD) blocks, are different across populations and because 3 the GWAS summary statistics used as a weight for each single-nucleotide 4 polymorphism (SNP) to calculate the PRS are based primarily on people of 5 European ancestry. Taking a PRS calculation method based on GWAS summary statistics from European individuals and applying it to non-European 6 7 individuals compromises prediction accuracy since the genetic risk of that 8 population may not be reflected properly [18]. Therefore, for future clinical 9 application of the AD PRS, it is necessary to evaluate the utility of this score in 10 populations of different ancestry. In addition, harmonization of protocols such as 11 inclusion and exclusion criteria is critical for rigorous comparisons between 12 different cohorts. 13 Therefore, in this study, we calculated the AD PRS in 504 Japanese 14 participants (145 cognitively unimpaired participants, 220 participants with late 15 MCI, and 139 patients with mild AD dementia) enrolled in the Japanese 16 Alzheimer's Disease Neuroimaging Initiative (J-ADNI) project and evaluated its 17 effectiveness in the North American ADNI (NA-ADNI) cohort including North 18 American 1,070 participants. The J-ADNI study used a harmonized protocol to 19 the NA-ADNI study. The previous comparative study of AD dementia between 20 the US and Japan in the ADNI projects reported that MCI in the Japanese 21 population shows similar progression profile as MCI in North America in terms 22 of cognitive function [19]. We moreover validated the AD PRS using 23 independent genomic data from 565 Japanese individuals with a 24 neuropathological diagnosis by autopsy. Furthermore, we also examined the 25 AD endophenotypes in association with PRS, and tested whether the PRS is 26 useful for predicting conversion from MCI to AD. 27 28 2. Materials and methods 29 2.1 Japanese participants from the J-ADNI cohort 30 Data used in the preparation of this article were obtained from the J-31 ADNI database deposited in the National Bioscience Database Center Human

- 32 Database, Japan (Research ID: hum0043.v1, 2016) [19]. This database
- 33 enrolled cognitively unimpaired (CU) participants, participants with late MCI,
- 34 and patients with mild AD dementia (ADD) using criteria consistent with those of
- 35 the North American ADNI (NA-ADNI) [20]. The J-ADNI was launched in 2007 as
- 36 a public-private partnership led by Principal Investigator Takeshi Iwatsubo, MD.

All rights reserved. No reuse allowed without permission.

- 1 The J-ADNI was aimed to test whether serial magnetic resonance imaging
- 2 (MRI), positron emission tomography (PET), other biological markers, and
- 3 clinical and neuropsychological assessment can be combined to measure the
- 4 progression of late MCI and mild ADD in the Japanese population. The J-ADNI
- 5 did not recruit participants with early MCI. The ethics committees of the
- 6 University of Tokyo, Osaka University and Niigata University approved the study.
- 7

8 A total of 715 volunteer participants between the ages of 60 and 84 9 years were diagnosed with late MCI or mild ADD or were CU and considered for 10 inclusion in the J-ADNI. Of the 715 participants assessed for study eligibility, 11 537 met the criteria and were enrolled. Of these 537 participants, 508 (CU, 147; 12 MCI, 221; ADD, 140) underwent genotyping analysis. Participants were evaluated every 6 or 12 months over a period of 36 months for CU and MCI 13 14 participants and over a period of 24 months for participants with ADD, as in the 15 NA-ADNI. As detailed below, the J-ADNI collected various imaging, clinical and 16 neuropsychological data from these participants in addition to the genomic data. 17 These data were obtained from the database described above.

18

#### 19 2.2 Japanese neuropathological cohort

20 An independent neuropathological (NP) cohort composed of 577 brain 21 donors was used for PRS validation [21]. Of these donors, 365 control donors 22 had little pathological findings associated with AD and 212 case donors had 23 those consistent with AD. All ADD patients were neuropathologically diagnosed 24 by senile plaque and neurofibrillary tangle. No neuropathological features of 25 other neurodegenerative disorders such as dementia with Lewy body disease, 26 frontotemporal lobal degeneration, and Parkinson's disease, were observed. 27 Control individuals did not show the typical neuropathological hallmarks of AD. 28 As no clinical diagnosis is provided in this cohort, the term case or control is 29 used in this study. As shown below, 565 brain donors (358 controls and 207 30 cases) passed QC. The demographic data of all the participants from the NP 31 cohort are shown in Table S1.

32 33

#### 34 2.3 Genotyping, quality control, and imputation

35 Whole blood samples from 508 participants in the J-ADNI cohort and 36 post-mortem frontal cortices from 577 donors in the NP cohort were genotyped

All rights reserved. No reuse allowed without permission.

1 using the Infinium Asian Screening Array (Illumina), containing 657,490 SNPs.

- 2 *APOE* genotypes in each participant were determined by haplotypes derived
- 3 from rs7412 and rs429358, which were genotyped using TagMan Assays
- 4 (Applied Biosystems). We excluded SNPs that (i) had duplicated genomic
- 5 positions, (ii) had low call rates (<5%), (iii) deviated from Hardy-Weinberg
- 6 equilibrium compared to controls ( $p < 1 \times 10^{-5}$ ), or (iv) had low minor allele
- 7 frequency (<0.01). For QC purposes, we excluded participants who (i) had sex
- 8 inconsistencies, (ii) had autosomal heterozygosity deviation ( $|F_{het}| \ge 0.2$ ), (iii)
- 9 had <99% of their genotypes called, or (iv) were in the same family according to
- 10 pi-hat (>0.2). Furthermore, we used principal component analysis to remove
- outliers based on the 1000 Genomes Project samples [22]. Finally, 451,713
- 12 autosomal SNPs and the samples, including 504 participants from the J-ADNI
- 13 cohort and 565 brain donors from the NP cohort passed the QC procedures.
- Next, we performed phasing with Eagle v2.4.1 [23] and imputation with Minimac4 [24] using the whole-genome sequencing data of 3,541 participants obtained from the BioBank Japan Project [25] and the 1000 Genomes Project [22] as reference genome data. After repeating the above QC procedure for the imputed SNP markers, we excluded SNPs with poor imputation quality ( $r^2 \le$ 0.3). Finally, we obtained 7,633,670 SNPs and the samples, including the 504 participants from the J-ADNI (CU, 145; MCI, 220; and ADD, 139) and 565 brain
- 21 donors from the NP cohort (control, 358; case, 207).
- 22

# 23 2.4 The NA-ADNI genetic data

The independent cohort data used in this study were obtained from the NA-ADNI [26]. The NA-ADNI was launched in 2003 as a public–private partnership led by Principal Investigator Michael W. Weiner, MD. The NA-ADNI was aimed to test whether serial MRI and PET data and the analysis of other biological markers and clinical and neuropsychological assessments can be combined to characterize the progression of MCI and early ADD.

SNP data from the NA-ADNI project were available for 1,674
participants across ADNI 1 and ADNI GO/2. Genotyping was conducted using
three different platforms: Human610-Quad, HumanOmniExpress and Omni
2.5 M (Illumina) [27]. The SNP data were imputed using the TOPMeD
imputation server after identical marker QC and sample QC as was used for the
J-ADNI was performed. The SNP data analysed on each of the three platforms
were imputed separately. After repeating the QC for the imputed SNP markers,

1 we excluded SNPs with poor imputation quality ( $r^2 \le 0.3$ ). If a participant was 2 genotyped on more than one genotyping array, the dataset with the fewest

3 missing values was selected.

4 According to the following procedures, we selected participants with 5 predicted central European ancestry and self-reported white non-Hispanic 6 ethnicity. For predicted ancestry, we used SNPweights software to infer genetic 7 ancestry from genotyped SNPs [28]. The reference panel comprised European, 8 West African, East Asian and Native American ancestral populations. 9 Participants with predicted central European ancestry of 80% or more were 10 retained. We obtained self-reported ethnicity information from the NA-ADNI 11 database. The clinical diagnosis at the final visit was used to categorize the 12 data. Furthermore, four participants who had significant memory concerns but no cognitive impairment were excluded. Finally, 1,482 participants (CU, 377; 13 14 MCI, 481; and ADD, 624) remained. 15 Of the 1,482 participants, 412 participants were participants in the Alzheimer's Disease Genetics Consortium (ADGC) and were included in the

Alzheimer's Disease Genetics Consortium (ADGC) and were included in the
meta-analysis of AD GWAS used as SNP weights in the PRS calculation
described below. We analysed a set of 1,070 participants (CU, 257; MCI, 453;
and ADD, 360), excluding the 412 participants to avoid overfitting. The
demographic data of all the participants from the NA-ADNI cohort are shown in
Table S2.

22

# 23 **2.5 Calculation of the PRS and prediction accuracy**

The PRS was calculated for each individual and is expressed as thefollowing weighted sum:

26  $PRS_i = \sum_{i=1}^{M} \beta_j x_{i,j} / M,$ 

where 
$$PRS_i$$
 is the PRS for individual *i*; *M* is the total number of SNPs used in  
the calculation;  $\beta_j$  is the weight of  $SNP_j$ , defined according to the effect size  
calculated by an independent GWAS; and  $x_{i,j}$  is the number of minor alleles of  
 $SNP_j$  that individual *i* has, thus has a value of 0, 1, or 2. In other words, the  
more minor alleles that are strongly associated with the disease, the higher the  
PRS.

SNPs included in the PRS were determined by the clumping and
 thresholding (C+T) method, the most common and supported method in AD

All rights reserved. No reuse allowed without permission.

1 studies [29, 30]. We used PRSice software implementing the C+T method to

- 2 calculate the PRS [31]. The clumping method preferentially retains markers
- 3 most strongly associated with disease from correlated markers in the same LD
- 4 block. The thresholding method removes variants with GWAS p values greater
- 5 than the selected p value threshold  $(p_T)$   $(p > p_T)$ . To determine the optimal  $p_T$ ,
- 6 we tested  $p_T$  values of 1×10<sup>-6</sup>, 1×10<sup>-5</sup>, 1×10<sup>-4</sup>, 1×10<sup>-3</sup>, 1×10<sup>-2</sup>, 0.05, 0.5, and 1.0.

7 SNPs were weighted by their effect sizes (beta coefficient) from the AD GWAS8 in the European population [32].

- 9 The ability of the PRS to accurately classify CU participants and ADD 10 patients was estimated in terms of (1) Nagelkerke's  $R^2$ , the proportion of the 11 variance explained by the regression model, and (2) the area under the receiver 12 operator characteristic curve (AUC). To calculate Nagelkerke's  $R^2$ , we 13 constructed a logistic regression model, including the PRS and the first two 14 components from the multidimensional scaling (MDS) analysis (full model), and 15 compared it to a model with only the first two MDS components (null model). We assessed the difference in Nagelkerke's  $R^2$  between the full and null models 16 17  $(R^2 = R^2_{Full} - R^2_{Null})$  and used the  $p_T$  corresponding to the highest value of Nagelkerke's  $R^2$ . The Nagelkerke's  $R^2$  was calculated by PRSice software [31]. 18 19 The AUC was calculated based on the prediction results of the logistic 20 regression model using the J-ADNI cohort as a test cohort. We also performed 21 5-fold cross validation (CV) to evaluate a predictive performance in a test 22 cohort. We estimated the 95% credible intervals by using the ci.auc function 23 from the R package "pROC". DeLong's test was conducted to assess potential 24 significant differences between curves using the roc.test function from the R 25 package "pROC".
- 26

#### 27 2.6 CSF biomarkers

In the J-ADNI cohort, cerebrospinal fluid (CSF) samples were assayed
for Aβ(1–42), total tau (tTau), and phosphorylated tau (pTau) by using a
multiplex xMAP Luminex platform (Luminex Corp, Austin, TX) with an
Innogenetics (INNO-BIA AlzBio3; Ghent, Belgium) immunoassay kit-based
reagent [33]. Of the 504 participants who underwent genotyping, 192
participants (CU, 52; MCI, 85; ADD, 55) also underwent CSF biomarker

- 34 measurements at baseline.
- 35

#### 36 2.7 Structural MRI and PET imaging

All rights reserved. No reuse allowed without permission.

1 All participants in the J-ADNI cohort underwent a structural MRI scan at 2 a signal strength of 1.5 tesla using a three-dimensional magnetization-prepared 3 rapid-acquisition gradient-echo sequence according to a standardized protocol 4 [34]. Cross-sectional and longitudinal processing streams in FreeSurfer, version 5 5.3, were used to estimate the atrophic changes in specific regions; we also evaluated the cortical thickness extracted in the longitudinal analysis. Of the 6 7 504 participants who underwent genotyping, the entorhinal cortex and 8 hippocampus of 443 participants (CU, 133; MCI, 196; ADD, 114) was assessed 9 by the FreeSurfer longitudinal stream. Each cortical thickness value was 10 adjusted by the total intracranial volume.

11

12 Of the 504 participants, 315 and 162 individuals underwent a positron emission tomography (PET) scan using <sup>18</sup>F-2-fluoro-2-deoxy-D-glucose (FDG) 13 14 and <sup>11</sup>C-Pittsburgh compound B (PiB), respectively. The PET scanning protocol 15 was standardized to minimize the inter-site and inter-scanner variability [35]. All 16 PET images went through the J-ADNI PET QC process [35]. The FDG PET 17 images were classified into seven categories based on the criteria of Silverman 18 et al. [36]. We analysed only PET images of 110 participants classified as 19 having a normal pattern (N1 pattern) and 161 participants classified as having 20 an AD pattern (P1 pattern). For PiB PET, the visual interpretation of four cortical 21 areas on each side (frontal lobe, lateral temporal lobe, lateral parietal lobe, and 22 precuneus/posterior cingulate gyrus) was evaluated by classifying PiB uptake in 23 each cortical region as positive, equivocal, or negative. Cases with one or more 24 positive cortical areas were defined as amyloid scan positive, and those with 25 negative results in all four cortical regions were defined as amyloid scan 26 negative. Other cases were considered equivocal. We analysed 65 negative 27 and 87 positive amyloid scans, excluding 10 participants who were judged to be 28 equivocal.

29

#### 30 2.8 Neuropsychological tests

All participants in the J-ADNI cohort underwent the following
 neuropsychological tests: Mini–Mental State Examination (MMSE), Functional
 Assessment Questionnaire (FAQ), Clinical Dementia Rating Scale Sum of

34 Boxes (CDR-SB), and AD Assessment Scale–Cognitive Subscale (ADAS-Cog).

35

#### 36 **2.9 Statistical analyses**

All rights reserved. No reuse allowed without permission.

1 Gene functional enrichment analysis of the closest genes around SNPs

- 2 included in the PRS was performed using the Metascape database
- 3 (http://metascape.org/) [37].
- For the association analyses between the PRS and endophenotypes,
  we compared slopes with zero by linear regression model analyses. The
  covariates included age at baseline examination, sex, years of education, the
  first two principal components (PCs), and doses of *APOE* ε4 and ε2 alleles. P
  values were adjusted by false discovery rate (FDR) to avoid type I error.
  Cox proportional hazards models using months of follow-up as a time
  scale were used to analyse the effects of PRSs on incident AD, presented as
- 11 hazard ratios (HRs) and 95% confidence intervals (CIs) derived from a model
- 12 with the following covariates: age at baseline examination, sex, years of
- 13 education, the first two PCs, and dose of APOE  $\epsilon$ 4 and  $\epsilon$ 2 alleles. We analysed
- 14 208 MCI participants over a follow-up period of  $\geq$  12 months. Nonconverters
- 15 were censored at the end of follow-up. Log-rank test was performed to examine
- 16 the difference in conversion to AD between two PRS groups. This test was
- 17 performed using only the PRS without covariates because the covariates other
- 18 than PRS could affect the differences between the groups. Cox proportional
- 19 hazard model analyses and log-rank tests were performed using the coxph and
- 20 survdiff functions from the R package "survival", respectively.
- 21

# 22 **3. Results**

# 3.1 The PRS successfully distinguish ADD patients and CU individuals in the J-ADNI cohort

After quality control of the genotyping data, the J-ADNI cohort included the 504 participants. The group with ADD had a higher mean age (p value <.001), a lower mean length of education (p value <.001), and a higher frequency of *APOE*  $\epsilon$ 4 carriers (p value <.001) than the CU group, whereas no differences were found in sex (p value = 0.429) or the frequency of *APOE*  $\epsilon$ 2 carriers (p value = 0.292) (**Table 1**).

We investigated whether the PRSs that were calculated using the statistics from the AD GWAS in the European population [32] are useful for discriminating between patients with ADD and CU individuals in the Japanese population. We calculated PRSs for 145 CU participants and 139 patients with ADD from the J-ADNI cohort. Our model using 173 SNPs showed the highest predictive power at  $p_T < 1 \times 10^{-5}$  and had a Nagelkerke's  $R^2$  of 0.167 (left side of

All rights reserved. No reuse allowed without permission.

1 **Table 2**), indicating that it explained more than 15% of the variance between

2 the CU and ADD groups.

3 Given the known predictive power of SNPs in the APOE region for AD, 4 we next removed this region from our PRS calculation to evaluate the predictive 5 power of other loci. To exclude the effect of APOE, we excluded ±500 kb 6 around APOE (Figure S1). This PRS, referred to as the PRS.noAPOE, was used in subsequent analyses. The predictive power of the PRS.noAPOE was 7 the highest for  $p_T < 1 \times 10^{-5}$ , with a Nagelkerke's  $R^2$  of 0.085 (right side of Table 8 9 2). The normalized values of the PRS.noAPOE of the ADD patients were 10 significantly higher than those of the CU and MCI participants (p value < 0.05, 11 Tukey's honestly significant difference (HSD) test; Figure 1), while there was no 12 significant difference between the CU and MCI participants (p value = 0.180, 13 Tukey's HSD test; Figure 1). These results suggest that the PRS contribute to 14 distinguish between ADD patients and CU individuals in J-ADNI cohort even 15 when the APOE region is excluded.

16

# 3.2 The PRS in combination with the APOE alleles improves predictive power

19 Next, we examined whether the PRS.noAPOE and the characteristics of 20 the participants independently influence the predictive power in J-ADNI cohort. 21 The PRS.noAPOE was not correlated with sex, years of education, age at 22 baseline examination, or the dose of the APOE ε4 or ε2 allele, even when 23 participants were stratified into CU, MCI, and ADD groups (p value > 0.05; 24 Figure S2). These results suggest that these factors contribute independently 25 to the discrimination of AD and that combinations of these factors improve 26 discrimination accuracy. We constructed a model including only the PRS.noAPOE and doses of APOE £4 and £2 alleles. This model showed 27 28 predictive performance of AUC = 0.755 (95% CI = 0.695-0.807) (Table 3). The 29 predictive performance of a monogenic model of only APOE alleles without the 30 PRS.noAPOE was AUC = 0.696 (95% CI = 0.640-0.751) (Table 3). The 31 addition of polygenic effects significantly improved the predictive accuracy of the monogenic model using only APOE (p value =  $9.36 \times 10^{-4}$ , DeLong test). 32 33 Additionally, the PRS model incorporating APOE alleles independently 34 (PRS.noAPOE + APOE doses) has higher accuracy than the PRS model that 35 includes SNPs in the APOE region (PRS.incAPOE) (AUC = 0.706; 95% CI =

36 0.643-0.764; p value = 0.049, DeLong test). Therefore, we constructed a

All rights reserved. No reuse allowed without permission.

1 predictive model including the PRS.noAPOE, sex, years of education, age at

2 baseline examination, and doses of APOE  $\epsilon$ 4 and  $\epsilon$ 2 alleles. This model

3 showed discriminative performance of AUC = 0.855 in distinguishing between

4 the ADD patients and CU individuals in the J-ADNI cohort (95% CI = 0.808-

5 0.898) (Table 3). These predictive performances showed the similar tendencies

6 when evaluated by 5-fold CV (Table S3). Taken together, these results showed

7 that the PRS based on European GWAS statistics was useful in discriminating

8 between patients with ADD and CU participants in the Japanese population.

9 Furthermore, the PRS had an effect independent of *APOE* alleles, and their

- 10 combination improved predictive accuracy.
- 11

12

#### 3.3 The effect of our PRS model is replicated in the independent cohorts

13 To examine the predictive accuracy of PRS.noAPOE in independent 14 cohorts, we calculated the PRS values for 565 brain donors in the NP cohort 15 (control, 358; case, 207) and 617 participants (CU, 257; ADD, 360) in the NA-16 ADNI using our PRS.noAPOE model. We note that the samples from the NP 17 cohort received a definitive diagnosis based on the typical neuropathological 18 hallmarks of AD using autopsy brains. The logistic regression model 19 constructed in the J-ADNI cohort was applied to each cohort to assess 20 discrimination accuracy. The predictive performance of PRS.noAPOE for the 21 NP cohort was lower than that for the J-ADNI cohort (AUC = 0.550 (95% CI = 22 0.500-0.599)), but when APOE alleles were added, the predictive performance 23 was replicated (AUC = 0.731 (95% CI = 0.686-0.773)) (Table 3).

In the NA-ADNI cohort, the imputed genotyping data included 130 of the 131 SNPs used in the PRS.noAPOE. A similar analysis in the NA-ADNI cohort also showed that the predictive performance of PRS.noAPOE in combination with *APOE* alleles (AUC = 0.730 (95% CI = 0.692-0.767)) was similar to that of the NP cohort. These analyses showed the reproducibility of our PRS model in independent cohorts.

30

```
31
```

### 3.4 ADD in the J-ADNI shows the polygenicity related to immune pathway

In order to examine the polygenicity of our PRS, we compared a model including only the PRS.noAPOE with a single-variable model for each of the 131 SNPs comprising the PRS.noAPOE. The single models with individual SNPs showed AUCs of 0.499 to 0.605 (median AUC = 0.515), while the model including only the PRS.noAPOE showed an AUC of 0.640 (95% CI = 0.576-

All rights reserved. No reuse allowed without permission.

1 0.704) (Tables 3 and S4), suggesting that the PRS.noAPOE reflects a

2 polygenic effect. Here, SNPs with AUCs of less than 0.5 indicate protection

3 rather than risk in our data.

4 We examined the genes closest to 131 SNPs included in the 5 PRS.noAPOE. We found the 97 closest genes located within ±100 kb around the SNPs. These genes were associated with leukocyte-mediated immunity 6 7 (FDR = 3.78×10<sup>-5</sup>), haematopoietic cell lineage (FDR = 4.45×10<sup>-5</sup>), the amyloid precursor protein (APP) catabolic process (FDR = 5.16×10<sup>-5</sup>), regulation of 8 9 transferase activity (FDR =  $3.57 \times 10^{-4}$ ), and glial cell proliferation (FDR = 10 5.60×10<sup>-3</sup>) (Table S5). Overall, we found that the integrated scores of multiple 11 SNPs around genes mainly associated with immune pathways may explain the 12 Japanese AD traits.

13

#### 14 3.5 The PRS associates with AD-related phenotypes

15 To examine whether our PRS associates with clinical characteristics, 16 we next investigated the correlation between the PRS.noAPOE and AD-related 17 phenotypes, namely CSF biomarker data and FDG and PiB PET brain imaging 18 data. We performed linear regression model analyses based on three models 19 controlling for seven covariates: age at baseline examination, sex, years of 20 education, the first two PCs, and the doses of APOE £4 and £2 alleles. Model 1 21 controlled only age at baseline examination, sex, years of education, and the 22 first two PCs. Models 2 and 3 took into the dose of APOE £4 allele in addition to 23 Model 1. Model 3 also added the dose of APOE  $\varepsilon^2$  allele as a full model.

The CSF tTau/Aβ42 and pTau/Aβ42 ratios were significantly
associated with the PRS.noAPOE values. These associations were basically
maintained in all models (FDR < 0.05, Wald test; Table 4a and Figure 2) and</li>
reflected the influences of tTau and pTau levels but not Aβ42 levels (Table S6).

To investigate the PRS effects to brain atrophy, we first tested the associations between the PRS and the volumes of the entorhinal cortex and hippocampus. Hippocampal volume showed a significant association with the PRS.noAPOE in Model 1 that did not include *APOE* alleles, but this association

- 32 did not remain significance after FDR correction (p value = 0.042, Wald test;
- 33 Table 4b). We investigated whether the PRS.noAPOE contributes to the
- 34 discrimination between the normal pattern (N1 pattern) and the AD pattern (P1
- 35 pattern) in FDG PET imaging and between negative and positive amyloid scans

1 in PiB PET imaging. As a result, the PRS was associated only with PiB PET

2 imaging (FDR < 0.05, Wald test; **Table 4c**).

We also investigated the correlations between the PRS and cognitive
functions. The neuropsychological tests, including the ADAS-Cog, CDR-SB,
FAQ, and MMSE, were significantly associated in all models (FDR < 0.01, Wald</li>
test; Table 4d).

7 We next stratified the participants into the CU, MCI and ADD groups 8 and examined the association between the PRS.noAPOE and each phenotype. 9 Significant positive correlations between the PRS.noAPOE and CSF tTau/Aß and between the PRS.noAPOE and pTau/Aβ42 ratios were observed in only 10 11 the MCI participants (FDR < 0.05, Wald test; Table 4a; Figure 2). In contrast, 12 these ratios remained stable or reached a plateau relative to the PRS.noAPOE in the CU and ADD participants (Figure 2), suggesting that the polygenic 13 14 burden beyond APOE explains some of the heterogeneity in MCI, especially in

- 15 terms of tau-related biomarker.
- 16

17 **3.6** APOE ε4 non-carriers with high PRS are at high risk of AD conversion

18 Finally, we examined difference in conversion to AD in the participants with 19 MCI stratified by PRS.noAPOE. We divided MCI participants into three groups 20 based on the PRS.noAPOE distribution of all participants. We compared the 21 conversion to AD of MCI participants in the 1st tertile, referred to as the low-22 PRS group, and of MCI participants in the 3rd tertile, noted as the high-PRS 23 group. We performed Cox proportional hazard model analysis controlling seven 24 covariates: age at baseline examination, sex, years of education, the first two 25 PCs, and the doses of APOE  $\varepsilon$ 4 and  $\varepsilon$ 2 alleles. We did not find significantly 26 different conversion patterns between the high- and low-PRS groups (p value = 27 0.202, log-rank test; Table 5 and Figure 3).

28 When we examined the contribution of each variable, we found that the 29 dose of the APOE ε4 allele significantly affected the conversion to AD (HR = 30 1.604; 95% CI = 1.153-2.230; p value = 0.005, Wald test; Table 5), suggesting 31 that this difference in conversion between the two PRS groups was influenced 32 by the APOE £4 allele dose. Therefore, we stratified MCI participants into those 33 with and without APOE £4. In that analysis, we found that among MCI 34 participants without APOE £4, the high-PRS group showed a significantly higher 35 convertion to AD than the low-PRS group (p value = 0.031, log-rank test; Table 36 5 and Figure 3). Moreover, the PRS.noAPOE significantly contributed to the

All rights reserved. No reuse allowed without permission.

- 1 difference in AD conversion between the two groups (HR = 2.216; 95% CI =
- 2 1.058-4.643; p value = 0.035, Wald test; Table 5). We also found no difference
- 3 in AD conversion among MCI participants with APOE ε4 (p value = 0.292, log-
- 4 rank test; Table 5 and Figure 3). These results suggested that polygenic
- 5 effects increase the risk of AD conversion, particularly in MCI subjects without
- 6 *APOE* ε4.
- 7 On the other hand, in APOE  $\varepsilon$ 4 carriers, a single factor, namely, APOE  $\varepsilon$ 4,
- 8 may explain much of the AD conversion risk. As expected, there was no
- 9 significant difference between the *APOE* ε4 noncarrier group with high-PRS and
- 10 the APOE  $\epsilon$ 4 carrier group (p value = 0.595, log-rank test; **Figure S3**). Although
- 11 age differences between the groups compared in the above analysis could have
- 12 affected the results, there were no differences in age at baseline examination
- 13 between the low- and high-PRS groups or between the converted and
- 14 nonconverted participants (p value > 0.05, Wilcoxon rank-sum test; Figure S4).
- These results suggest that the PRS contributes to the conversion to AD in
  participants without *APOE* ε4.
- 17

#### 18 4. Discussion

19 In this study, we evaluated the utility of the PRS for AD in a Japanese 20 cohort. The results showed that the PRS had an effect independent of APOE 21 and showed relatively high predictive accuracy when combined with APOE  $\varepsilon 4$ . 22 In addition, this effect was replicated in the cohort with a neuropathological 23 diagnosis and the protocol-harmonized independent NA-ADNI cohort. The PRS 24 was significantly associated with CSF tau levels in MCI participants, and MCI 25 with a high PRS was associated with an elevated risk of AD conversion in 26 APOE ε4 noncarriers.

27

28 Despite the difference in genetic structure between the European and 29 Japanese populations [38], the PRS developed in this study, PRS.noAPOE, 30 showed meaningful predictive accuracy. Such predictive accuracy may be 31 achieved because all participants were diagnosed according to unified inclusion 32 and exclusion criteria and harmonized standardized diagnostic criteria using the 33 same neuropsychological tests (MMSE, CDR-SB, and Wechsler Memory Scale 34 Logical Memory II). The optimal p value threshold for the PRS excluding the 35 APOE region was also similar to that reported in previous studies,  $p_T < 1 \times 10^{-5}$ 36 [5, 10, 39]. Moreover, while dozens of SNPs were incorporated into these

All rights reserved. No reuse allowed without permission.

1 previous PRSs, 131 SNPs were included to calculate the PRS in our study. This

- 2 difference in the number of SNPs is likely due to differences in genetic structure
  3 such as LD blocks. Hence, even if there are racial differences, adding a few
- 4 dozen SNPs may preserve accuracy.

5 There is no consensus on the number of SNPs that should be included in 6 the AD PRS. According to a systematic review of PRS studies in AD, PRSs of 7 AD can be organized into two groups: PRSs containing relatively large numbers 8 of SNPs, ranging from 4,431 to 359,500, and PRSs containing relatively small 9 numbers, ranging from 5 to 31 [40]. The latter group is referred to as the 10 oligogenic effect, in contrast to the polygenic effect [41]. From this perspective, 11 our PRS apparently represents an oligogenic effect. Notably, a relatively small 12 number of SNPs has the advantage of providing an inexpensive gene panel. In 13 addition, a PRS composed of many SNPs may be sensitive to geographic 14 differences in genetic structure, whereas a PRS composed of a few dozen 15 SNPs is robust to population bias [42, 43]. However, we should note that our 16 PRS may reflect ancestral differences due to the use of European GWAS 17 statistics. In the future, more robust polygenic effects could be verified by using 18 GWAS statistics for large groups of East Asians, including Japanese 19 individuals.

20

21 In our study, 97 genes contributing to the PRS.noAPOE were associated 22 with APP degradation, immunity, and glial cell proliferation. Genetic variants 23 found in a recent AD GWAS were associated with the APP catabolic process 24 and tau protein binding [44]. In addition, many of the genes affected by their 25 genetic variants are expressed in microglia [44]. An analysis of cognitively 26 healthy centenarians in addition to ADD patients and healthy controls revealed 27 that the PRS associated with the immune system was lower in the centenarian 28 group independent of APOE  $\varepsilon 4$ , indicating that immune system function is 29 involved in AD resistance [45]. Therefore, our results suggest that common 30 factors related to AD may be shared in the vulnerability of clearance 31 mechanisms and neuroimmune surveillance in the brain among different 32 population.

33

34 In our study, the PRS.noAPOE showed significant correlations with CSF

- 15 tTau/A $\beta$ 42 and pTau/A $\beta$ 42 ratios only in individuals with MCI. Tau but not A $\beta$ 42
- 36 strongly influenced this result even controlling APOE effect. This correlation

All rights reserved. No reuse allowed without permission.

may have been observed on in the MCI group because individuals with MCI can have a broad spectrum of clinical characteristics, including CSF tau values, as observed in this study. Interestingly, NA-ADNI studies have shown that the PRS is associated beyond *APOE* with CSF tau but not CSF A $\beta$ 42 [43, 46]. From the above, independent studies in different ancestry groups have confirmed that polygenic effects are associated with tau-related biomarkers, especially in individuals with MCI.

8

9 Although our results are noteworthy, we must approach the clinical 10 application of our PRS with caution at this stage because the predictive 11 accuracy of our PRS alone is not very high. Similar to currently available PRSs, 12 few biomarkers can perfectly distinguish disease or not; most markers bear 13 some uncertainty. AD and MCI are explained not only by genetic aspects such 14 as PRS, but also by anatomic aspects such as MRI and PET imaging and 15 biological aspects such as CSF biomarkers [47], suggesting that combining 16 multiple biomarkers could compensate for each other's weaknesses in 17 predictive performance. PRS will allow individuals' disease risk to be assessed 18 at a relatively early stage, leading to future lifestyle modification and disease 19 prevention.

20

21 There were several limitations to this study. First, the CU participants 22 included in the J-ADNI were relatively young. We acknowledge that these CU 23 participants include potential patients who will develop AD in the future. 24 Considering the average age of onset of AD and the allele frequency of APOE 25 ε4 in the Japanese population, future work should ideally include CU 26 participants that are over 70 years old [48]. Second, because the number of 27 participants available for the study was small, there was limited power to identify 28 relationships between the PRS and some phenotypes. Larger studies are 29 needed to validate the results of this study. Therefore, combining samples from 30 multiple East Asian cohorts, including cohorts from Japan, is necessary for 31 analysis.

32

#### 33 5. Conclusion

This study demonstrated that the AD PRS showed a relatively high
performance in the Japanese population, despite differences in genetic
structure from the European population. Furthermore, this PRS was replicated

All rights reserved. No reuse allowed without permission.

1 in the independent Japanese and European cohorts. The AD PRS correlated

- 2 with phenotypes such as CSF tau levels in MCI. The AD PRS predicted the
- 3 development of AD in MCI participants without APOE ε4. The application of the
- 4 PRS will allow us to know an individuals' disease risk at a relatively early life
- 5 stage, which may lead to future lifestyle modification and disease prevention.
- 6

# 7 Abbreviations

8 Polygenic risk score (PRS); Alzheimer's disease (AD); Japanese

9 Alzheimer's Disease Neuroimaging Initiative (J-ADNI); apolipoprotein E

- 10 (*APOE*); amyloid-beta (Aβ); mild cognitive impairment (MCI); magnetic
- 11 resonance imaging (MRI); cognitively unimpaired (CU); Alzheimer's disease
- 12 dementia (ADD); North American Alzheimer's Disease Neuroimaging Initiative
- 13 (NA-ADNI); clumping and thresholding (C+T); area under the receiver operator
- 14 characteristic curve (AUC); multidimensional scaling (MDS); cerebrospinal fluid
- 15 (CSF); total tau (tTau); phosphorylated tau (pTau); positron emission
- 16 tomography (PET); <sup>18</sup>F-2-fluoro-2-deoxy-D-glucose (FDG); <sup>11</sup>C-Pittsburgh
- 17 compound B (PiB); Mini–Mental State Examination (MMSE); Functional
- 18 Assessment Questionnaire (FAQ); Clinical Dementia Rating (CDR); CDR–Sum
- 19 of Boxes (CDR-SB); AD Assessment Scale–Cognitive Subscale (ADAS-Cog);

20 false discovery rate (FDR); hazard ratio (HR); confidence interval (CI)

21

# 22 Acknowledgements

- 23 We thank all the participants and staff of the J-ADNI and NA-ADNI, and
- 24 the donors and facility staff for providing autopsy brains. The J-ADNI was
- 25 supported by the following funding sources: the Translational Research
- 26 Promotion Project from the New Energy and Industrial Technology
- 27 Development Organization of Japan; Research on Dementia, Health Labor
- 28 Sciences Research Grant; the Life Science Database Integration Project of
- 29 Japan Science and Technology Agency; the Research Association of
- 30 Biotechnology (Astellas Pharma Inc., Bristol-Myers Squibb, Daiichi-Sankyo,
- 31 Eisai, Eli Lilly and Company, Merck-Banyu, Mitsubishi Tanabe Pharma, Pfizer
- 32 Inc., Shionogi & Co., Ltd., Sumitomo Dainippon, and Takeda Pharmaceutical
- 33 Company), Japan; and a grant from an anonymous foundation. The reference
- 34 genome data used for this research were originally obtained by participants in
- 35 the Tailor-made Medical Treatment Program (BioBank Japan: BBJ), led by Prof.

All rights reserved. No reuse allowed without permission.

- 1 Michiaki Kubo; these data are available at the website of the NBDC Human
- 2 Database/the Japan Science and Technology Agency (JST).
- 3

## 4 Authors' contributions

5 MK: Study design, analysis and interpretation of data, and manuscript 6 draft. AM and NH: Genotyping analysis, interpretation of data, and manuscript 7 revision. YS, SM, AK, HA: Provision of autopsy brains and manuscript revision. 8 KK, KO, SN, RK, Tlwatsubo, and AN: Interpretation of data and manuscript 9 revision. Tlkeuchi: Study design, interpretation of data, and manuscript draft. All 10 authors read and approved the final manuscript.

11

# 12 Funding

13 This work was supported by a Grant-in-Aid for Scientific Research (grant 14 numbers 20K15778 to MK, 21K07271 and 21H03537 to AM, and 22H04923 to 15 YS) from the Ministry of Education, Culture, Sports, Science and Technology 16 (MEXT), by grants from the Japan Agency for Medical Research and 17 Development (AMED) (grant numbers JP21dk0207045 and JP23dk0207060 to 18 MK, AM, KO, SN, and TI, JP23wm0525019 to MK and TI, and JP21wm0425019 19 to YS), by Grants-in Aid from the Research Committee of CNS Degenerative 20 Diseases, Research on Policy Planning and Evaluation for Rare and Intractable 21 Diseases, Health, Labour and Welfare Sciences Research Grants, the Ministry 22 of Health, Labour and Welfare, Japan (grant number 20FC1049 to YS). The funders had no role in the study design, data collection, decision to publish, or 23 24 preparation of the manuscript.

25

# 26 Availability of data and materials

All the J-ADNI data except for the genome data and the reference genome
data were obtained from the NBDC Human Database/the Japan Science and
Technology Agency (JST) (<u>https://humandbs.biosciencedbc.jp/en/hum0043-v1</u>),
(<u>https://humandbs.biosciencedbc.jp/en/hum0014-latest#JGAS000114rp</u>).
GWAS statistics were obtained from the Center for Neurogenomics and
Cognitive Research (<u>https://ctg.cncr.nl/software/summary\_statistics</u>). The J-

- 33 ADNI genome data are available on request.
- 34

# 35 Declarations

1	Ethics approval and consent to participate
2	This study was approved by the ethics committees of the University of
3	Tokyo, Osaka University and Niigata University.
4	
5	Consent for publication
6	Consent for publication has been granted by J-ADNI administrators.
7	
8	Competing interests
9	All authors confirm that they have no competing interests to declare.
10	
11	References
12	1. Lourida I, Hannon E, Littlejohns TJ, Langa KM, Hypponen E, Kuzma E, et al.
13	Association of Lifestyle and Genetic Risk With Incidence of Dementia. JAMA.
14	2019;322(5):430-7.
15	2. Licher S, Ahmad S, Karamujic-Comic H, Voortman T, Leening MJG, Ikram MA,
16	et al. Genetic predisposition, modifiable-risk-factor profile and long-term dementia risk
17	in the general population. Nat Med. 2019;25(9):1364-9.
18	3. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al.
19	Dementia prevention, intervention, and care: 2020 report of the Lancet Commission.
20	Lancet. 2020;396(10248):413-46.
21	4. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, et al.
22	Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry.
23	2006;63(2):168-74.
24	5. Karlsson IK, Escott-Price V, Gatz M, Hardy J, Pedersen NL, Shoai M, et al.
25	Measuring heritable contributions to Alzheimer's disease: polygenic risk score analysis
26	with twins. Brain Commun. 2022;4(1):fcab308.
27	6. Frieser MJ, Wilson S, Vrieze S. Behavioral impact of return of genetic test results
28	for complex disease: Systematic review and meta-analysis. Health Psychol.
29	2018;37(12):1134-44.
30	7. Smit AK, Allen M, Beswick B, Butow P, Dawkins H, Dobbinson SJ, et al. Impact
31	of personal genomic risk information on melanoma prevention behaviors and
32	psychological outcomes: a randomized controlled trial. Genet Med. 2021;23(12):2394-
33	403.
34	8. Ridge PG, Hoyt KB, Boehme K, Mukherjee S, Crane PK, Haines JL, et al.
35	Assessment of the genetic variance of late-onset Alzheimer's disease. Neurobiol Aging.
36	2016;41:200 e13- e20.

All rights reserved. No reuse allowed without permission.

1 9. Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, et al. 2 Common polygenic variation enhances risk prediction for Alzheimer's disease. Brain. 3 2015;138(Pt 12):3673-84. 4 10. Sabuncu MR, Buckner RL, Smoller JW, Lee PH, Fischl B, Sperling RA, et al. 5 The association between a polygenic Alzheimer score and cortical thickness in clinically 6 normal subjects. Cereb Cortex. 2012;22(11):2653-61. 7 Mormino EC, Sperling RA, Holmes AJ, Buckner RL, De Jager PL, Smoller JW, 11. 8 et al. Polygenic risk of Alzheimer disease is associated with early- and late-life processes. 9 Neurology. 2016;87(5):481-8. 10 12. Ge T, Sabuncu MR, Smoller JW, Sperling RA, Mormino EC, Alzheimer's 11 Disease Neuroimaging I. Dissociable influences of APOE epsilon4 and polygenic risk of 12 AD dementia on amyloid and cognition. Neurology. 2018;90(18):e1605-e12. 13 13. Zettergren A, Lord J, Ashton NJ, Benedet AL, Karikari TK, Lantero Rodriguez 14 J, et al. Association between polygenic risk score of Alzheimer's disease and plasma 15 phosphorylated tau in individuals from the Alzheimer's Disease Neuroimaging Initiative. 16 Alzheimers Res Ther. 2021;13(1):17. 17 Daunt P, Ballard CG, Creese B, Davidson G, Hardy J, Oshota O, et al. Polygenic 14. 18 Risk Scoring is an Effective Approach to Predict Those Individuals Most Likely to 19 Decline Cognitively Due to Alzheimer's Disease. J Prev Alzheimers Dis. 2021;8(1):78-20 83. 21 15. Pyun JM, Park YH, Lee KJ, Kim S, Saykin AJ, Nho K, et al. Predictability of 22 polygenic risk score for progression to dementia and its interaction with APOE epsilon4 23 in mild cognitive impairment. Transl Neurodegener. 2021;10(1):32. 24 16. Wand H, Lambert SA, Tamburro C, Iacocca MA, O'Sullivan JW, Sillari C, et al. 25 Improving reporting standards for polygenic scores in risk prediction studies. Nature. 2021;591(7849):211-9. 26 27 17. Polygenic Risk Score Task Force of the International Common Disease A. 28 Responsible use of polygenic risk scores in the clinic: potential benefits, risks and gaps. 29 Nat Med. 2021;27(11):1876-84. 30 18. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use 31 of current polygenic risk scores may exacerbate health disparities. Nat Genet. 32 2019;51(4):584-91. 33 19. Iwatsubo T, Iwata A, Suzuki K, Ihara R, Arai H, Ishii K, et al. Japanese and North 34 American Alzheimer's Disease Neuroimaging Initiative studies: Harmonization for 35 international trials. Alzheimers Dement. 2018;14(8):1077-87. 36 20. Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, et al.

All rights reserved. No reuse allowed without permission.

1 Recent publications from the Alzheimer's Disease Neuroimaging Initiative: Reviewing 2 progress toward improved AD clinical trials. Alzheimers Dement. 2017;13(4):e1-e85. 3 21. Wen Y, Miyashita A, Kitamura N, Tsukie T, Saito Y, Hatsuta H, et al. SORL1 is 4 genetically associated with neuropathologically characterized late-onset Alzheimer's 5 disease. J Alzheimers Dis. 2013;35(2):387-94. 6 22. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, 7 et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74. 8 23. Loh PR, Danecek P, Palamara PF, Fuchsberger C, YAR, HKF, et al. Reference-9 based phasing using the Haplotype Reference Consortium panel. Nat Genet. 10 2016;48(11):1443-8. 11 24. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, et al. Next-12 generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284-7. 13 25. Okada Y, Momozawa Y, Sakaue S, Kanai M, Ishigaki K, Akiyama M, et al. Deep 14 whole-genome sequencing reveals recent selection signatures linked to evolution and 15 disease risk of Japanese. Nat Commun. 2018;9(1):1631. 16 26. Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack C, Jagust W, et al. The 17 Alzheimer's disease neuroimaging initiative. Neuroimaging Clin N Am. 2005;15(4):869-18 77, xi-xii. 19 27. Saykin AJ, Shen L, Foroud TM, Potkin SG, Swaminathan S, Kim S, et al. 20 Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: 21 Genetics core aims, progress, and plans. Alzheimers Dement. 2010;6(3):265-73. 22 28. Chen CY, Pollack S, Hunter DJ, Hirschhorn JN, Kraft P, Price AL. Improved 23 ancestry inference using weights from external reference panels. Bioinformatics. 24 2013;29(11):1399-406. 25 29. Prive F, Vilhjalmsson BJ, Aschard H, Blum MGB. Making the Most of Clumping 26 and Thresholding for Polygenic Scores. Am J Hum Genet. 2019;105(6):1213-21. 27 30. Leonenko G, Baker E, Stevenson-Hoare J, Sierksma A, Fiers M, Williams J, et 28 al. Identifying individuals with high risk of Alzheimer's disease using polygenic risk 29 scores. Nat Commun. 2021;12(1):4506. 30 31. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. 31 Bioinformatics. 2015;31(9):1466-8. 32 32. Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. 33 Genome-wide meta-analysis identifies new loci and functional pathways influencing 34 Alzheimer's disease risk. Nat Genet. 2019;51(3):404-13. 35 33. Kasuga K, Kikuchi M, Tsukie T, Suzuki K, Ihara R, Iwata A, et al. Different

36 AT(N) profiles and clinical progression classified by two different N markers using total

All rights reserved. No reuse allowed without permission.

tau and neurofilament light chain in cerebrospinal fluid. BMJ Neurol Open.
2022;4(2):e000321.

3 34. Fujishima M, Kawaguchi A, Maikusa N, Kuwano R, Iwatsubo T, Matsuda H, et
al. Sample Size Estimation for Alzheimer's Disease Trials from Japanese ADNI Serial
5 Magnetic Resonance Imaging. J Alzheimers Dis. 2017;56(1):75-88.

6 35. Ikari Y, Nishio T, Makishi Y, Miya Y, Ito K, Koeppe RA, et al. Head motion
7 evaluation and correction for PET scans with 18F-FDG in the Japanese Alzheimer's
8 disease neuroimaging initiative (J-ADNI) multi-center study. Ann Nucl Med.
9 2012;26(7):535-44.

36. Silverman DH, Small GW, Chang CY, Lu CS, Kung De Aburto MA, Chen W, et
al. Positron emission tomography in evaluation of dementia: Regional brain metabolism
and long-term outcome. JAMA. 2001;286(17):2120-7.

13 37. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al.
14 Metascape provides a biologist-oriented resource for the analysis of systems-level
15 datasets. Nat Commun. 2019;10(1):1523.

16 38. Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA,
17 et al. Genetic structure of human populations. Science. 2002;298(5602):2381-5.

18 39. Najar J, van der Lee SJ, Joas E, Wetterberg H, Hardy J, Guerreiro R, et al.
19 Polygenic risk scores for Alzheimer's disease are related to dementia risk in APOE
20 varepsilon4 negatives. Alzheimers Dement (Amst). 2021;13(1):e12142.

40. Stocker H, Mollers T, Perna L, Brenner H. The genetic risk of Alzheimer's
disease beyond APOE epsilon4: systematic review of Alzheimer's genetic risk scores.
Transl Psychiatry. 2018;8(1):166.

24 41. Zhang Q, Sidorenko J, Couvy-Duchesne B, Marioni RE, Wright MJ, Goate AM,
25 et al. Risk prediction of late-onset Alzheimer's disease implies an oligogenic architecture.
26 Nat Commun. 2020;11(1):4799.

42. Kerminen S, Martin AR, Koskela J, Ruotsalainen SE, Havulinna AS, Surakka I,
et al. Geographic Variation and Bias in the Polygenic Scores of Complex Diseases and
Traits in Finland. Am J Hum Genet. 2019;104(6):1169-81.

30 43. Altmann A, Scelsi MA, Shoai M, de Silva E, Aksman LM, Cash DM, et al. A
31 comprehensive analysis of methods for assessing polygenic burden on Alzheimer's
32 disease pathology and risk beyond APOE. Brain Commun. 2020;2(1):fcz047.

33 44. Bellenguez C, Kucukali F, Jansen IE, Kleineidam L, Moreno-Grau S, Amin N,
34 et al. New insights into the genetic etiology of Alzheimer's disease and related dementias.
35 Nat Genet. 2022;54(4):412-36.

36 45. Tesi N, van der Lee SJ, Hulsman M, Jansen IE, Stringa N, van Schoor NM, et al.

All rights reserved. No reuse allowed without permission.

1 Immune response and endocytosis pathways are associated with the resilience against

2 Alzheimer's disease. Transl Psychiatry. 2020;10(1):332.

46. Leonenko G, Shoai M, Bellou E, Sims R, Williams J, Hardy J, et al. Genetic risk
for alzheimer disease is distinct from genetic risk for amyloid deposition. Ann Neurol.
2019;86(3):427-35.

6 47. Kikuchi M, Kobayashi K, Itoh S, Kasuga K, Miyashita A, Ikeuchi T, et al.
7 Identification of mild cognitive impairment subtypes predicting conversion to
8 Alzheimer's disease using multimodal data. Comput Struct Biotechnol J. 2022;20:52969 308.

10 48. Nunomura A, Chiba S, Eto M, Saito M, Makino I, Miyagishi T. Apolipoprotein
11 E polymorphism and susceptibility to early- and late-onset sporadic Alzheimer's disease

12 in Hokkaido, the northern part of Japan. Neurosci Lett. 1996;206(1):17-20.

13

14

All rights reserved. No reuse allowed without permission.



1



### 3 those in the CU and MCI groups.

- 4 The PRS.noAPOEs in each group were represented by violin plots (CU, n=145;
- 5 MCI, n=220; ADD, n=139). Each violin plot includes the kernel probability
- 6 density of the data at different values and the box plots with the median value
- 7 and the interquartile range. Tukey's HSD test was used to perform multiple
- 8 comparisons of PRSs among each group. We normalized the PRS distribution
- 9 to have a mean of 0 and an SD of 1. CN = cognitively normal; MCI = mild
- 10 cognitive impairment; ADD = Alzheimer's disease dementia.



3 Figure 2. The PRS.noAPOE correlated with CSF Tau/Aβ42 ratios in the 4 MCI.

- 5 CSF tTau/Aβ42 (A) and pTau/Aβ42 (B) ratios by decile of PRS are shown in
- 6 each diagnostic group. The participants were divided into ten groups based on
- 7 the PRS.noAPOE, ranging from the lowest group (1<sup>st</sup> decile) to the highest
- group (10<sup>th</sup> decile). CN = cognitively normal; MCI = mild cognitive impairment; 8
- ADD = Alzheimer's disease dementia. 9



3 Figure 3. The high-PRS group was more likely to convert to AD than the

#### low-PRS group in the APOE ε4 non-carrier individuals with MCI. 4

- Kaplan-Meier survival curves for conversion rates of MCI to AD in the low-PRS 5
- group (1<sup>st</sup> tertile) and the high-PRS group (3<sup>rd</sup> tertile). The shaded area 6
- represents the 95% confidence interval. 7
- 8



- 1
- 2 Figure S1. The excluded region around the *APOE* gene.
- 3 We removed the *APOE* region, consisting of ±500 kb, from around the top-hit
- 4 SNP rs1160985 (chr19:45403412) in our data. Each data point indicates GWAS
- 5 p values from Jansen *et al.* [32] used as SNP weights in the PRS calculation.
- 6

All rights reserved. No reuse allowed without permission.



- 1
- 2 Figure S2. Associations between the PRS and covariates.
- 3 Age at baseline examination and years of education were examined by
- 4 Spearman correlation. Sex and doses of APOE ε4 and ε2 alleles were analysed

- All rights reserved. No reuse allowed without permission.
- 1 by t tests or ANOVAs. CN = cognitively normal; MCI = mild cognitive
- 2 impairment; ADD = Alzheimer's disease dementia.
- 3



4



- 6 APOE ε4 non-carriers with high PRS values.
- 7 Kaplan–Meier survival curves for the rates of conversion from MCI to AD in
- 8 APOE £4 carriers and APOE £4 non-carriers with high PRS values. The shaded
- 9 area represents the 95% confidence interval.

All rights reserved. No reuse allowed without permission.



# Low PRS vs High PRS

#### Nonconverter vs Converter



1

### 2 Figure S4. Age differences between the low- and high-PRS groups and

- 3 between the nonconverters and converters.
- 4 Baseline ages were compared between groups using the Wilcoxon rank-sum
- 5 test. Each violin plot includes the kernel probability density of the data at
- 6 different values and the box plots with the median value and the interquartile
- 7 range.