Journal of Alzheimer's Disease 84 (2021) 1323–1335
DOI 10.3233/JAD-210700
IOS Press

Association Between Polygenic Risk Score and the Progression from Mild Cognitive Impairment to Alzheimer's Disease

Hongliang Liu[a,b], Michael Lutz[c] and Sheng Luo[d,*] for the Alzheimer's Disease Neuroimaging Initiative[1]

[a]*Duke Cancer Institute, Duke University Medical Center, Durham, NC, USA*
[b]*Department of Population Health Sciences, Duke University School of Medicine, Durham, NC, USA*
[c]*Division of Translational Brain Sciences, Department of Neurology, Duke University Medical Center, Durham, NC, USA*
[d]*Department of Biostatistics and Bioinformatics, Duke University, Durham, NC, USA*

Handling Associate Editor: Jin-Tai Yu

Accepted 13 September 2021
Pre-press 13 October 2021

Abstract.
Background: Mild cognitive impairment (MCI) is a heterogeneous condition and MCI patients are at increased risk of progression to dementia due to Alzheimer's disease (AD).
Objective: In this study, we aim to evaluate the associations between polygenic risk scores (PRSs) and 1) time to AD progression from MCI, 2) changes in longitudinal cognitive impairment, and 3) biomarkers from cerebrospinal fluid and imaging.
Methods: We constructed PRS by using 40 independent non-*APOE* SNPs from well-replicated AD GWASs and tested its association with the progression time from MCI to AD by using 767 MCI patients from the ADNI study and 1373 patients from the NACC study. PRSs calculated with other methods were also computed.
Results: We found that the PRS constructed with SNPs that reached genome-wide significance predicted the progression from MCI to AD (beta = 0.182, SE = 0.061, $p = 0.003$) after adjusting for the demographic and clinical variables. This association was replicated in the NACC dataset (beta = 0.094, SE = 0.037, $p = 0.009$). Further analyses revealed that PRS was associated with the increased ADAS-Cog11/ADAS-Cog13/ADASQ4 scores, tau/ptau levels, and cortical amyloid burdens (PiB-PET and AV45-PET), but decreased hippocampus and entorhinal cortex volumes ($p < 0.05$). Mediation analysis showed that the effect of PRS on the increased risk of AD may be mediated by $A\beta_{42}$ (beta = 0.056, SE = 0.026, $p = 0.036$).
Conclusion: Our findings suggest that PRS can be useful for the prediction of time to AD and other clinical changes after the diagnosis of MCI.

Keywords: Alzheimer's disease, longitudinal analysis, mild cognitive impairment, neuroimaging, polygenic risk score, survival analysis

[1]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

*Correspondence to: Sheng Luo, Department of Biostatistics and Bioinformatics, Duke University, Durham, NC 77030, USA. Tel.: +1 919 668 8038; Fax: +1 919 668 8038; E-mail: sheng.luo@duke.edu.

INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disorder with the main neuropathological features of neuronal loss and accumulation of intracellular tau-containing neurofibrillary tangles and extracellular amyloid-β (Aβ) deposits in the brain [1]. It is estimated there are around 5.8 million AD patients in the US and 700,000 people age 65 years and older will die from AD in 2020 [2]. Late-onset AD (LOAD) has onset after the age of 65 years and accounts for 97% of AD in the US. LOAD is sporadic and up to 60-80% of cases are inheritable. As a complex disease, the genetic etiology of AD is still not completely understood. Carrying the E4 allele of apolipoprotein E (*APOE4*) is a well-known genetic risk factor for AD [3, 4]. Genome-wide association studies (GWASs) have identified about 40 AD-related loci. These loci map by proximity to genes involve in lipid metabolism, inflammation, immune function, endocytosis, and other biological pathways [5, 6]. However, with the exception of the coding SNPs in the *APOE* gene, SNPs in other regions have minor effects on AD risk and age of onset.

Polygenic risk scores (PRS) is an approach that combines the effects of multiple SNPs based on their effect sizes and can be used for the prediction of risks of multiple diseases including AD [7, 8]. By using the summary data from the GWAS study of the International Genomics of Alzheimer's Project (IGAP), PRS has been successfully applied to predict life-time AD risk, brain structure changes (e.g., hippocampal cortical thickness, hippocampal volume), cognitive ability changes, and biomarkers from cerebrospinal fluid and plasma [9–15]. Moreover, a Polygenic Hazard Score (PHS) was recently developed with *APOE* E2 and E4 alleles, and 31 SNPs selecting from the IGAP summary data after applying the step-wise Cox regression and used to quantify individual differences with age-specific genetic risk for AD [16]. Both the PRS and PHS models have comparable performance on the prediction of AD risk [11].

Mild cognitive impairment (MCI) is an intermediate state between the expected cognitive decline of normal aging and AD dementia. About one-third of MCI patients will develop AD over time [17, 18]. The early detection of MCI patients with high risk of AD may provide an opportunity to apply preventive or therapeutic interventions, which may be helpful for preventing or delaying the further conversion to AD [19]. Therefore, it is important to develop models that distinguish clinically heterogeneous MCI with different risks of AD. In this study, we constructed polygenic risk scores (PRS) by using the SNPs from four well-replicated, consortium GWASs and tested its association with the progression from mild cognitive impairment (MCI) to AD as well as cognitive ability and imaging features by using the clinical and biological data from the ADNI study.

METHODS

Participants

In the discovery stage, we used the dataset from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study (http://adni.loni.ucla.edu). There were 2,253 participants from the ADNI 1, ADNI GO, and ADNI 2 phases with clinical and biological data available (September 2005–January 2020), which included clinical diagnosis (Normal Cognition, MCI, and AD), longitudinal cognitive phenotypes, and neuropathological phenotypes including biomarkers from cerebrospinal fluid (CSF), and the imaging results of PET and MRI. Of them, 916 participants diagnosed with MCI at baseline or during follow-up and 767 samples with GWAS data available remained for further analysis. There were 294 subjects diagnosed with AD during the follow-up period.

An independent sample of MCI patients was selected from the National Alzheimer's Coordinating Center (NACC) to replicate the findings from ADNI. There were a total of 41,459 participants in the requested NACC dataset, of which 12,068 presented with MCI at baseline or during the following visits. Of these patients, there were 1,373 patients with at least two visits and who had GWAS data available. During the follow up visits, 864 out of the 1,373 MCI patients progressed to AD and 509 patients did not progress to AD but continued to have a diagnosis of MCI at the last visit. The flowcharts of the participant selection and further details of the two datasets have been presented in Supplementary Figure 1 and in the Supplementary Material.

The ADNI and NACC studies were approved by local institutional review boards, and all participants or participant's guardians provided written informed consent. Additional information about ADNI and NACC studies are available at http://www.adni-info.org and https://www.alz.washington.edu/WEB/study_pop.html, respectively.

Genotyping and imputation

We downloaded the GWAS data from the ADNI database on February 20, 2020, Genotyping was performed with the platforms of Illumina Human610-Quad, Illumina Human OmniExpress, and Illumina Omni 2.5M on 1,674 MCI participants from ADNI 1, GO, and ADNI 2 phases, respectively. We then used SHAPEIT for phasing and performed imputation with minimac4 on the Michigan imputation server (https://imputationserver.sph.umich.edu) with the HRC reference panel (Version r1.1 2016) consisting of 64,940 haplotypes of predominantly European ancestry. For imputation, a set of high-quality SNPs were used: MAF >0.01; call rate $>95\%$, Hardy-Weinberg equilibrium test $p > 10^{-6}$; allele frequency difference ≤ 0.20 between the sample data and the reference panel. The genotyping data was processed by using PLINK 1.90/2.0.

For NACC, we requested the GWAS data for the 10,256 subjects in NACC AD Centers 1-7, of which genotyping was performed with the platforms of Human660W-Quad_v1_A, HumanOmni Express-12v1_A/H, and humanomniexpressexome-8v1-2_a, respectively (https://www.alz.washington.edu/ADGC/GENOtype.html) [20]. We then conducted quality control and imputation with the same procedures of ADNI.

PRS/PHS construction

There were 40 loci associated with AD risk reported by four well-powered AD GWASs/GWAXs (genome-wide association study by proxy) and 78 leading SNPs were found in those loci, which included two SNPs (rs41289512 and rs12691088) in the genomic region of *APOE* locus (19q13.32) [5, 6, 9, 21, 22]. All but one SNP rs9271058 were included in the imputation data. We performed clump analysis with PLINK 1.90 for these 77 SNPs based on the ADNI GWAS dataset and the reported p values from the previous GWASs. 41 independent SNPs, including one *APOE* SNP rs41289512, were selected based on an r-squared threshold greater than 0.1. We did not include rs41289512 in the PRS calculation since this SNP is located in the *APOE* region and had substantially greater effect size than other SNPs. The other *APOE* SNP rs12691088 was not included in the PRS calculation as it has moderate linkage disequilibrium ($r^2 = 0.53$) with rs41289512. To account for the effect of the *APOE* gene, we included the *APOE* E2 and E4 allele counts as covariates in the final regression model instead of including them in PRS calculation. The remaining 40 SNPs were used to construct PRS, weighted by the effect size estimates from the GWAS with the largest sample size for association with LOAD. Genomic data on the PRS SNPs is included in Supplementary Table 1. The PRS was calculated using PRSICE-2 and standardized to z-scores (centering by mean and scaling by standard deviation) [23]. For comparison, we also constructed PRSs by using the results of 4,465 SNPs with $p \leq 1 \times 10^{-5}$ in a recent AD GWAS Meta study (including 71,880 cases and 383,378 controls) [21]. SNPs within the 250kb region surrounding *APOE* (chr19:45159053-45662650; GRCh37/hg19 assembly) were excluded and five different thresholds (i.e., $p \leq 1 \times 10^{-5}$, 5×10^{-6}, 1×10^{-6}, 5×10^{-7}, and 1×10^{-7}) were applied for SNPs selection for the PRS construction. The summary results of this GWAS were downloaded from the website of the CTG laboratory (https://ctg.cncr.nl/documents/p1651/AD_sumstats_Jansenetal_2019sept.txt.gz). Desikan et al. [16] reported that PHS combining the effects of *APOE* E2 and E4 alleles, as well as 31 GWAS SNPs was associated with the age of AD onset. However, for comparison with PRS, we constructed PHS by only using those 31 reported SNPs, without including the *APOE* E2 and E4 alleles. After linkage-disequilibrium (LD) based clumping, 25 independent SNPs remained ($r^2 < 0.1$) in both the ADNI and NACC datasets and their reported effect sizes [log(HR)] in the same publication were used as the weights for PHS construction. Information of the 25 SNPs can be found in Supplementary Table 2 [16]. For method comparisons, we also constructed a Bayesian PRS by using the method of LDpred2 [24, 25]. Briefly, this method estimates posterior mean causal effect sizes for each SNP with a Gibbs sampler by assuming a prior for the genetic architecture and LD information from a reference panel. PRS was then constructed using the posterior effect sizes.

Statistical analysis

The association between demographic and clinical characteristics, PRS, and the progression from MCI to AD was tested using Cox proportional hazards regression with adjustment for age at baseline, sex, years of education, race, top three principal components (PCs) from the genetic data, and the allele copies of *APOE* E2 and E4. The endpoint event was the occurrence of AD. Survival time (in years)

Table 1
Characteristic distributions of the demographic and clinical variables in the ADNI dataset

Variables	Total	Event (AD, %)	Beta[1]	SE[1]	p^1
Overall	767	294 (38.3)			
Follow-up time (y)					
Median (Min-max)	3.00 (0.39–13.00)	2.03 (0.46–11.50)			
Age (y)			0.028	0.008	**0.001**
Mean ± SD	73.4 ± 7.40	74.3 ± 7.00			
Median (Min-max)	73.6 (55.0–91.4)	74.4 (55.0–88.4)			
Sex			–0.013	0.119	0.912
Female	299	114 (38.1)			
Male	468	180 (38.5)			
Education (y)			–0.001	0.021	0.951
mean ± SD	15.9 ± 2.82	15.9 ± 2.69			
Race[2]			–0.249	0.238	0.295
Non-Hispanic White	700	275 (39.3)			
Other races	67	19 (28.4)			
Baseline tau[3]			1.358	0.167	**5.07×10^{-16}**
Mean ± SD [min-max]	5.56 ± 0.43	5.76 ± 0.41			
Median (Min-max)	5.55 (4.58–6.72)	5.76 (4.62–6.71)			
Baseline ptau[3]			1.261	0.145	**$< 2 \times 10^{-16}$**
Mean ± SD [min-max]	3.19 ± 0.49	3.43 ± 0.46			
Median (Min-max)	3.18 (2.11–4.52)	3.44 (2.17–4.52)			
Baseline Aβ[3]			–1.523	0.155	$< \times 10^{-16}$
Mean ± SD [min-max]	6.79 ± 0.47	6.55 ± 0.42			
Median (Min-max)	6.77 (5.35–7.44)	6.51 (5.59–7.44)			
APOE-E4			0.554	0.080	**3.35×10^{-12}**
0	385	108 (28.1)			
1	302	143 (47.4)			
2	80	43 (53.8)			
APOE-E2			–0.492	0.250	**0.049**
0	708	277 (39.1)			
1	59	17 (28.8)			

[1]Univariate Cox regression. [2]There are 222 participants with missing data on tau/ptau/Aβ levels. [3]Log-transformed.

was defined as the duration from baseline for patients with MCI or the date of first diagnosis of MCI to AD occurrence or censoring. The HR from the Cox model analysis indicated the risk of progressing to AD caused by the per 1 SD change in PRS.

We performed log-transformation for the CSF biomarkers (Aβ, tau, and ptau). The volumes of five brain regions [i.e., Ventricles, Hippocampus, Entorhinal, Fusiform gyrus, and Middle temporal gyrus (MidTemp)] were represented as the percentage to the intracerebral volume (ICV). The correlations between PRS and these longitudinal phenotypes were investigated with a linear mixed model by including a random intercept and a random slope of follow-up time with adjustment for age at baseline, sex, years of education, race, top three principal components (PCs), and the allele copies of *APOE* E2 and E4. The *survival* and *nlme* packages in R were used for these analyses. Mediation analysis was performed to quantify the direct and indirect relationships between PRS, $A\beta_{42}$, PTau, and AD progression by using the method and related R code from one previous publication [26]. Direct and indirect effects were estimated with generalized linear and Cox regression analysis with adjustment for age at baseline, sex, years of education, race, top three PCs from the genetic data, and the number of allele copies of *APOE* E2 and E4. Bootstrapping was used for significance testing and standard error estimations of these effects. All analyses were conducted with R (version 3.5.1) if not mentioned otherwise.

RESULTS

Characteristics of the study populations

There were 916 participants diagnosed with MCI at baseline or during follow-up. After merging the PRS data and clinical data, 767 samples remained for further analysis and 294 subjects were diagnosed with AD during the follow-up period. It should be noted that there were missing values in the baseline tau/ptau and Aβ levels for 222 subjects. The distributions of demographic and clinical variables were

Table 2
Survival analysis of polygenic risk score (PRS)/polygenic hazard score (PHS) and the progression of MCI to AD in ADNI dataset

	PRS			PHS		
Variables	Beta[1]	SE[1]	p^1	Beta[2]	SE[2]	p^2
Age (y)	0.045	0.009	4.59×10^{-7}	0.038	0.008	**4.46×10^{-6}**
Male in sex	–0.140	0.126	0.267	–0.146	0.118	0.214
Education (y)	0.019	0.022	0.382	0.015	0.020	0.459
Other races	–0.103	0.240	0.668	–0.186	0.228	0.414
PC3	3.276	2.730	0.230	5.604	2.715	**0.039**
PC9	–8.854	3.537	0.012	–7.408	2.855	**0.009**
PC10	–5.879	3.184	0.065	–4.698	2.567	0.067
APOE E4 (allele number)	0.613	0.088	3.82×10^{-12}	0.573	0.082	**3.10×10^{-12}**
APOE E2 (allele number)	–0.295	0.253	0.243	–0.170	0.229	0.458
PRS/PHS	0.182	0.061	0.003	0.068	0.056	0.221

[1]Multivariate Cox regression model including PRS, age at baseline, sex, years of education, race, significant principal components, the allele numbers of *APOE* E4 and *APOE* E2 (#total = 767 and #event = 294). [2]Multivariate Cox regression model including PHS, age at baseline, sex, years of education, race, significant principal components, the allele numbers of *APOE* E4 and *APOE* E2 (#total = 767 and #event = 294).

presented in Table 1. In the univariate analysis, we found that age, $A\beta_{42}$, tau, and ptau levels at baseline, number of *APOE* E2 and E4 alleles were all significantly associated with the progression to AD among MCI patients (Table 1, $p < 0.05$).

There were 1,373 MCI patients with both the genotyping data and survival data in NACC, of which 864 subjects were diagnosed with AD during the follow-up period. The distributions of age, sex, education years, race, *APOE* E2/3/4 alleles can be found in Supplementary Table 3. We found significant associations between race, number of *APOE* E2, E3, and E4 alleles and AD survival, which partially replicated the results in the ADNI study.

Survival analysis of PRS/PHS and the progression of MCI to AD

We investigated the association between the PRS/PHS and the progression of MCI to AD by applying Cox proportional hazard models with adjustment for age at baseline, sex, years of education, race, ethnicity, top three significant PCs from the genetic data, and the number of allele copies of *APOE* E4 and *APOE* E2. As presented in Table 2, we found that the PRS was associated with the time interval of progression of MCI to AD (beta = 0.182, SE = 0.061, $p = 0.003$), which indicated that MCI patients with higher PRS scores have increased risk developing AD after MCI. Survival probability was also estimated by the Kaplan-Meier approach. Patients from the ADNI study were divided into three groups based on the PRS quartiles: group 1 (PRS < the 25th percentile), group 2 (< the 75th percentile), and group3 ($\geq$ the 75th percentile). As shown in Fig. 1a, patients in group 2 and 3 had a significantly shorter survival time than those in group 1 [median progression time = NA (as less than 50% patients developing AD at the end of follow-up), 6.04, and 4.03 years for MCI patients in group 1, 2, and 3, respectively; Log-rank $p = 0.003$]. However, no evidence was found for the correlation between PHS and the progression of MCI to AD ($p = 0.221$ in Table 2). We also tested the correlation between the Bayesian PRS constructed with the LDpred2 method and MCI progression and found consistent results with the PRS using the Prsice2 method (beta = 0.281, SE = 0.125, $p = 0.024$) (Supplementary Table 4). We constructed PRSs using SNPs with different thresholds (i.e., $p \leq 1 \times 10^{-5}$, 5.0×10^{-6}, 1.0×10^{-6}, 5.0×10^{-7}, and 1.0×10^{-7}) and tested their associations with MCI progression ($p = 0.058$, 0.044, 0.084, 0.129, and 0.149, respectively). Only PRS constructed with SNPs with $p \leq 5.0 \times 10^{-6}$ showed significant association with MCI to AD progression (beta = 0.115, SE = 0.058, $p = 0.044$) (Supplementary Table 5). However, this association was not replicated in the NACC dataset (Supplementary Table 6; $p = 0.224$).

A previous study reported that Cox proportional hazard and Logistic regression models provided very similar estimates of regression coefficients in studies with short follow up time (5 years or less) [27]. We tested the association of PRS and AD status by using logistic regression in the ADNI dataset and found similar result between PRS and AD risk (beta = 0.195, SE = 0.080, $p = 0.015$) (Supplementary Table 7).

We performed stratified analysis by the numbers of *APOE* E2/E3/E4 alleles and found that the PRS using the Prsice2 method was significantly associated with the risk of AD progression in the strata of

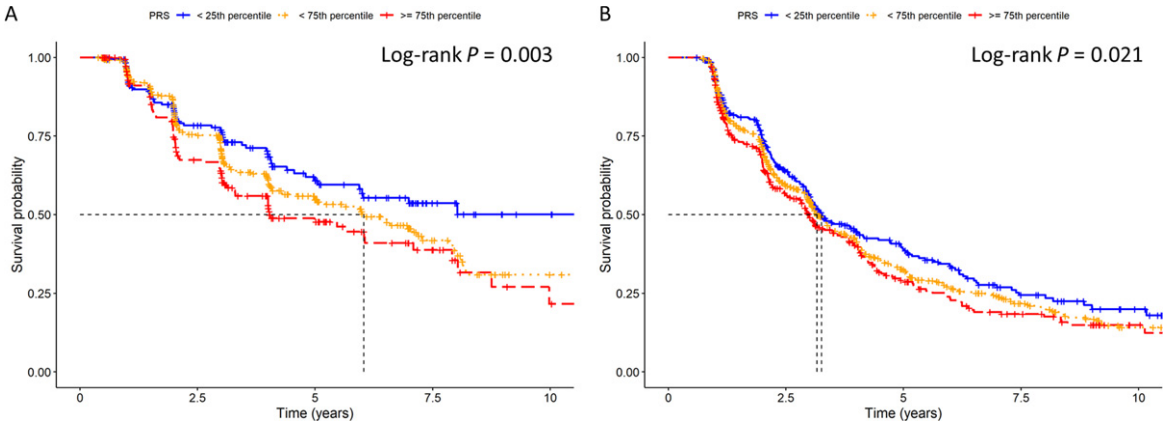

Fig. 1. Kaplan Meier plots of AD progression based on the trichotomized polygenic risk score (PRS). Patients were divided into three groups: group 1 (PRS < the 25th percentile), group 2 (< the 75th percentile), and group 3 (≥ the 75th percentile). A) Patients in the ADNI dataset (Log-rank $p = 0.003$): $n = 767$ (with # of deaths = 294), median progression time = 6.04, 4.03 years for MCI patients in group 2, and 3, respectively; B) Patients in the NACC dataset (Log-rank $p = 0.021$): $n = 1373$ (with # of deaths = 864), median progression time = 3.26, 3.16, 2.99 years for MCI patients group 1, 2, and 3, respectively.

Table 3
Associations of polygenic risk score (PRS)/polygenic hazard score (PHS) and longitudinal cognitive impairments for in ADNI dataset

Phenotype	Participants	Observations	PRS			PHS		
			Beta[1]	SE[1]	p^1	Beta[1]	SE[1]	p^1
MMSE	766	3992	–0.091	0.066	0.168	–0.044	0.064	0.490
CDR-SB	763	3956	0.008	0.034	0.826	–0.028	0.034	0.404
MOCA	482	2372	–0.258	0.125	**0.039**	–0.002	0.120	0.989
FAQ	763	3952	0.085	0.142	0.551	–0.125	0.146	0.393
ADAS-Cog11	766	3987	0.499	0.161	**0.002**	0.113	0.159	0.477
ADAS-Cog13	765	3964	0.800	0.240	$\mathbf{9.0 \times 10^{-4}}$	0.209	0.236	0.377
ADASQ4	767	3993	0.373	0.086	$\mathbf{1.4 \times 10^{-5}}$	0.145	0.081	0.077

[1]Adjusted for age at baseline, sex, years of education, race, significant principal components, and the allele numbers of *APOE* E4 and *APOE* E2.

participants with one copy of *APOE E4*, null copy of *APOE E2*, or one copy of *APOE E3* ($p < 0.05$) (Supplementary Table 8). This association was not significant in participants possessing two copies of *APOE* E4. We then tested the interaction effect of PRS and *APOE* alleles by including interaction terms in model but did not find statistical significance for any of the interactions ($p = 0.564$, 0.800, and 0.931 for the interaction terms between PRS and *APOE* E2, E3, and E4, respectively). No significance was found for the *APOE* stratified results of PHS association with progression from MCI to AD (Supplementary Table 8).

Correlation between PRS/PHS and longitudinal cognitive impairments

The correlation between the PRS/PHS and longitudinal cognitive scores (MMSE, MOCA, CDR-SB, FAQ, and ADAS-Cog 11/13/Q4) were presented in Table 3 adjusted for age, sex, years of education, race, ethnicity, top 3 significant PCs, and the numbers of *APOE* E4 and *APOE* E2 alleles. We found that the PRS was significantly correlated with the worse cognitive performance measured by ADAS-Cog11, ADAS-Cog13, and ADASQ4 (beta = 0.499, 0.800, and 0.373; $p = 0.002$, 9.0×10^{-4}, and 1.4×10^{-5}, respectively). PRS constructed with GWAS SNPs with $p \leq 5 \times 10^{-6}$ also showed significant association with the three cognitive measures (beta = 0.345, 0.650, and 0.346; $p = 0.034$, 0.007, and 0.0001 for ADAS-Cog11, ADAS-Cog13, and ADASQ4, respectively; Supplementary Table 5).

However, no significant association was found for the other cognitive phenotypes. We also did not find any significant correlation between PHS and the longitudinal cognitive scores.

Table 4
Correlation between polygenic risk score (PRS)/polygenic hazard score (PHS) and longitudinal biomarkers (from CSF, PET, and MRI imaging) in subjects with MCI at baseline or diagnosed during following visits in ADNI dataset

Biomarker	#Participants	#Observations	PRS			PHS		
			Beta[1]	SE[1]	p^1	Beta[1]	SE[1]	p^1
Aβ[2]	295	712	–0.037	0.023	0.116	–0.018	0.024	0.447
tau[2]	295	711	0.068	0.022	**0.002**	0.012	0.023	0.598
ptau[2]	294	709	0.082	0.025	**0.001**	0.014	0.026	0.599
PIB	34	81	0.175	0.068	**0.017**	–0.061	0.069	0.387
FDG	385	1,278	–0.012	0.007	0.072	–0.015	0.006	**0.015**
AV45	312	817	0.040	0.011	$\mathbf{3.0 \times 10^{-4}}$	0.009	0.010	0.374
Ventricles[3]	700	2,878	0.009	0.040	0.812	-0.033	0.038	0.391
Hippocampus[3]	652	2,553	–0.011	0.003	$\mathbf{1.0 \times 10^{-4}}$	-0.004	0.003	0.142
Entorhinal[3]	626	2,431	–0.006	0.002	$\mathbf{7.0 \times 10^{-4}}$	–0.003	0.002	0.113
Fusiform[3]	626	2,431	–0.003	0.006	0.608	–0.003	0.006	0.656
MidTemp[3]	626	2,431	–0.013	0.007	0.054	–0.012	0.007	0.062

[1]Adjusted for age at baseline, sex, years of education, race, significant principal components, and the allele copies of *APOE* E4 and *APOE* E2. [2]Log-transformed. [3]These dependent variables were expressed as the percentages to intracranial volume.

Table 5
Association between polygenic risk score (PRS) and the longitudinal clinical outcomes in NACC dataset

Outcome	#participants	#events	Beta[1]	SE[1]	p^1
Survival (MCI to AD)	1,373 #participants	864 #observations	0.094	0.037	**0.009**
MMSE	1,227	4,294	–0.138	0.067	**0.039**
CDR-SB	1,360	5,227	0.033	0.035	0.346
CDR_Global	1,360	5,227	0.004	0.006	0.513
Ventricles[2]	52	118	–0.005	0.004	0.992
Hippocampus[2]	52	118	–0.034	0.039	0.393
Entorhinal[2]	52	118	–0.089	0.097	0.303
Fusiform[2]	52	118	–0.076	0.055	0.174
MidTemp[2]	52	118	–0.038	0.027	0.171

[1]Adjusted for age at baseline, sex, years of education, race, and the allele copies of *APOE* E4 and E2. [2]These dependent variables were expressed as the percentages to intracranial volume.

Correlation between PRS/PHS and longitudinal biomarkers from CSF and imaging

We investigated the correlations between the PRS/PHS and longitudinal biomarkers from CSF as well as the PET- and MRI-based image data. The distributions of the number of observations in these analyses were presented in Supplementary Table 9. As shown in Table 4, there were significant correlations between the PRS and increased levels of tau and ptau in CSF (beta = 0.068 and 0.082; $p = 0.002$ and 0.001, respectively). We also observed significant correlations between PRS and cortical amyloid burden, which was quantified using ^{11}C-PiB PET and AV45 (^{18}F florbetapir) PET (beta = 0.175 and 0.040; $p = 0.017$ and 3.0×10^{-4}, respectively). MCI patients with different PRS scores had different volumes in hippocampus, entorhinal cortex, and middle temporal gyrus ((beta = –0.011, –0.006, and –0.013; $p = 1.0 \times 10^{-4}$, 7.0×10^{-4}, and 0.054, respectively). We also observed significant correlations between the PHS and FDG and MidTemp (beta = –0.015; $p = 0.015$). The PRS constructed with SNPs with $p \leq 5 \times 10^{-6}$ in GWAS also presented significant associations with the increases of Aβ (beta = –0.059; $p = 0.010$), tau (beta = 0.066; $p = 0.003$), ptau (beta = 0.078; $p = 0.002$) in CSF, amyloid burden measured with AV45 PET (beta = 0.025; $p = 0.020$), and the volume of hippocampus (beta = –0.007; $p = 0.016$) in the ADNI dataset (Supplementary Table 5).

Replication in NACC study

We performed an independent replication using the NACC dataset. As shown in Table 5, there is a significant association between the PRS constructed using the Prsice2 method and increased risk of AD progression (beta = 0.094, SE = 0.037, $p = 0.009$), which is consistent with the finding in the ADNI dataset.

We observed similar results using the Bayesian PRS constructed with the LDpred2 method (beta = 0.200, SE = 0.075, $p = 0.007$) (Supplementary Table 4).

The survival probability was also estimated with the Kaplan-Meier approach. As shown in Fig. 1b, patients in groups 2 and 3 had significantly shorter survival times than those in group 1 (median progression time = 3.26, 3.16, 2.99 years for MCI patients in group 1, 2, and 3, respectively; Log-rank $p = 0.021$).

In the longitudinal analysis, we found a significant correlation between the PRS and the change in of MMSE for MCI patients. However, we could not replicate the associations between the PRS and MRI markers identified in the ADNI study. We performed stratified analysis by age, sex, race, and numbers of *APOE* E2/E3/E4 alleles and found that the PRS was significantly associated with the risk of AD progression in the strata of older ($\geq$55), females, Whites, *APOE* E4 heterozygous individuals and *APOE* E2 non-carriers ($p < 0.05$) (Supplementary Table 10).

Model performance comparison

We evaluated the predictive accuracy of different models by using the Harrell's C-statistic (Supplementary Table 11). As shown, the *APOE* model including numbers of *APOE* E2 and E4 alleles had a significantly increased C-statistic compared with the demographic model that included only demographic variables in the ADNI data (0.635 versus 0.557 for the *APOE* model and demographic model, respectively; $p = 7.37 \times 10^{-7}$). After adding PRS calculated by the PRSICE2 method to the *APOE* model, there is a slight increase in the Harrell's C-statistic (0.644 versus 0.635 for the PRS1 model and *APOE* model, respectively; $p = 0.124$). Similar results were observed for PRS2 model calculated using the Bayesian PRS (Harrell's C-statistic = 0.642; $p = 0.119$ compared with the *APOE* model) and the model including PRS3 constructed with GWAS SNPs with $p \leq 5.0\text{x}10^{-6}$ (Harrell's C-statistic = 0.638; $p = 0.384$ compared with the *APOE* model).

Models that included the PHS in contrast with the *APOE* model did not show any significant improvement in predictive accuracy (0.627 versus 0.627 for the PHS model and *APOE* model, respectively; $p = 0.814$). In the NACC dataset, we found that the *APOE* model had better performance on prediction than the demographic model (Harrell's C-statistic: 0.575 versus 0.524 for the *APOE* model and demographic model, respectively; $p = 3.52 \times 10^{-4}$); however, inclusion of PRS or PHS in the *APOE* model did not significantly increase the prediction accuracy (Harrell's C-statistic = 0.585 and 0.615 for PRS and PHS models, respectively; $p = 0.418$ and 0.579 for prediction comparisons of PRS model versus *APOE* model, and PHS model versus *APOE* model, respectively). We also estimated the prediction accuracy of model including the Bayesian PRS and found its Harrell's C-statistic = 0.574 ($p = 0.394$ compared with the *APOE* model). Non-significant results were found for the PRS constructed with GWAS SNPs with $p \leq 5 \times 10^{-6}$ (Harrell's C-statistic = 0.574; $p = 0.538$ comparing with the *APOE* model).

Mediation analysis

Reduction of CSF $A\beta_{42}$ levels and elevation of phosphorylated tau protein are two important pathologic changes in the brains of MCI and AD patients. In this study, we observed that the PRS were associated with the baseline levels of CSF $A\beta_{42}$ and phosphorylated tau in MCI patients (Supplementary Figure 2). We performed mediation analysis in the ADNI dataset (547 MCI patients with 198 AD event) to test if the effect of the PRS on AD progression are mediated by the two biomarkers. As shown in Supplementary Figure 2, we found that the PRS contributed the increased risk of AD progression through the mechanism mediated by $A\beta_{42}$ (beta = 0.056, SE = 0.026, $p = 0.036$). No statistical significance was observed for the direct effect of PRS (beta = 0.067, SE = 0.072, $p = 0.355$) as well as the indirect effect mediated by phosphorylated tau (beta = 0.036, SE = 0.020, $p = 0.075$). Our results suggested that about 35.2% [0.056/(0.067 + 0.056 + 0.036) = 0.352] effect of PRS on AD progression was mediated by $A\beta_{42}$.

DISCUSSION

In this study, we investigated the association between PRSs/PHS and the time of progression to AD in patients with MCI. We found that the PRS constructed with SNPs reported in previously AD GWASs was significantly associated with the progression from MCI to AD in both the ADNI cohort and the NACC study after adjusting for demographic and clinical variables. Further analysis showed that this PRS was associated with increases of ADAS-Cog11/ADAS-Cog13/ADASQ4 scores, tau/ptau levels in CSF, amyloid burden in brain regions, and the decreased volumes of hippocampus and entorhinal cortex. Although several studies have reported the association between PRS and AD risk as well

as with AD pathologic-related biomarkers [11, 12, 27–33], there are at least three differences between this study and those previously published. First, in the study design, we used the NACC dataset as an independent replication of our findings on the PRS in the ADNI dataset, while the results of most previous studies were only based on the ADNI dataset. Second, we performed a comprehensive investigation of the correlation between the constructed PRS and longitudinal CSF biomarkers (i.e., $A\beta_{42}$, tau/ptau levels) and imaging biomarkers (i.e., amyloid burden and the volumes of various structures of the brain). Other studies were limited to investigation of the association of AD PRSs and baseline-levels of $A\beta_{42}$/tau/ptau, and/or longitudinal changes of volumes. Third, in this study, we performed mediation analysis to explore possible mechanisms underlying the association between the AD PRS and AD progression, which has not been reported in other studies. We found that the effect of the PRS on the time of progression to AD from MCI was partly mediated via $A\beta_{42}$. Finally, we compared the results of conventional PRS (based on different p value thresholds and LD pruning) with a PRS calculated by a recently published Bayesian method (LDPred2) [24] and with PRSs constructed using SNPs with different p-value selection thresholds. Overall, the study and specifically the mediation analysis places the PRS association results in context with the National Institutes for Health/Alzheimer's Association (NIH/AA) Research Framework for observational and interventional research in AD [34–37]. The framework provides guidelines to interpret the definition of AD in terms of specific biomarkers related to Aβ deposition, pathologic tau accumulation and neurodegeneration [AT(N)]. The framework considers AD as a continuum where the biomarkers help to stage the severity of the disease and symptoms in context of underlying biology and pathophysiology. The mediation analysis provides estimates of the contribution of the PRS, derived from numerous biological pathways to specific AT(N) biomarker-related processes to the progression of AD. Overall, our findings suggested that the PRS derived from the AD GWASs is useful for the prediction of AD risk and other clinical changes in MCI patients.

It has been established that the progression from MCI to AD can be predicted by the biomarkers from MRI imaging and CSF, *APOE* genotypes, and different cognitive test batteries [11, 38, 39]. Recent studies also suggested that both PRS and PHS can be used to predict an age-specific risk for developing AD [11, 31, 40]. In this study, we investigated the association between PRSs/PHS and the time from MCI to AD and found that PRSs constructed with SNPs reported in previous AD GWASs can be used for the prediction of progression from MCI to AD in both the discovery and replication studies. Our results are similar with previous studies, which constructed PRS using the non-*APOE* SNPs from the IGAP GWAS and reported a predictive ability for LOAD progression from MCI of 61–67%, especially in MCI patients with amyloid deposition [28, 33]. By using the effect size estimates from a recent AD GWAS meta-analysis [6], Altmann et. al. constructed PRSs with two p-value cut-offs: $p = 0.5$ and 1.0×10^{-5} (genome-wide suggestive) and found the two PRSs had significantly different distributions between cognitively normal individuals and MCI patients, as well as between cognitively normal individuals and AD patients [27]. However, both PRS did not show significant differences between MCI and AD patients. Desikan et al. reported that the PHS constructed from 31 SNPs can significantly predict time to progression from MCI to AD in the ADNI dataset [32]. However, we did not replicate this association in our study. Although both studies used the ADNI dataset, this discrepancy might be due to the difference in SNPs used for PHS construction. Our study applied LD clumping before PHS construction, which resulted in 25 independent SNPs with pairwise r-squared < 0.1, and also excluded the two SNPs in *APOE* region, while the PHS in the Desikan study was constructed with 31 SNPs including the two *APOE* SNPs, which had stronger effects compared with SNPs from other loci [16]. Such results suggest that *APOE* SNPs in the PHS reported in this paper [16] may play the major role in the prediction of AD risk in MCI patients.

In this study, we found that the AD PRS was associated with CSF biomarkers (i.e., CSF Aβ and tau/ptau) and amyloid burden in brain regions. Our results were consistent with two previous studies which reported significant associations between the PRS based on significant SNPs identified in the International Genomics of Alzheimer's Disease GWAS and baseline levels of CSF Aβ and tau/ptau [27, 29]. The mediation analysis suggested that the association between PRS and AD progression might via a mechanism mediated by Aβ. Since both Aβ plaque and tau deposition precede memory impairment [41], these results suggest that PRS may reflect pathophysiology in early AD stages. However, further functional studies are warranted. It should be noted that the PRSs constructed in another study with two

different *P*-value thresholds showed significant association with the baseline levels of tau/ptau, but not Aβ. Such inconsistencies might be a consequence that PRSs constructed with variants with strong effects are more robust to population structure differences [27, 42].

Both the entorhinal cortex and the hippocampus are essential parts of the medial temporal lobe system that supports declarative memory [43]. It has been shown that early AD associated Braakian atrophy includes the hippocampus, entorhinal and cingulate cortices [44, 45]. In this study, we found that PRS had significant associations with the volume changes of hippocampus and entorhinal cortex. This was consistent with a previous study, which found that the AD PRS comprised of non-*APOE* SNPs was associated with reduced hippocampal volume in healthy older adults and those with MCI [13].

The present study had several limitations. First, the sample size may not be large enough to detect certain associations, especially for analyses using samples with CSF biomarker data of Aβ and tau. Second, we did not replicate the association between the PRS and imaging markers observed in the ADNI study in the NACC study, which might be due to the small sample size in the NACC dataset for individuals with longitudinal biomarker data. Third, different diagnostic criteria of AD were used across different UDS versions and Alzheimer's disease centers (ADCs) contributing data to the NACC study, which might cause population heterogeneity for AD diagnoses. Validation of our findings in a larger study is warranted. Finally, AD develops over a period of decades often without evidence of cognitive decline. The interval between MCI and development of dementia represents only one part of the cycle of neurodegenerative events. Other factors, including vascular pathology that is independent of the genetic factors modeled in the PRS may influence the rate of progression of AD development for individuals.

In conclusion, a PRS derived from 40 common AD risk variants with small effect size was found to be associated with the progression time from MCI to AD independent of *APOE* E4 and E2 risk alleles. The correlation between the PRS and early AD markers (tau/ptau, and atrophy of hippocampus and entorhinal cortex) shows evidence for a genetic modulation of neurodegeneration, and the potential for a combination of PRS and other brain biomarkers to aid in the prediction of MCI patients at risk of AD. Mediation analysis and interpretation of the results in context of the NIH/AA research framework enable use of the PRS in a more precise approach for the design of clinical trials with specific agents that may affect Aβ, pathologic tau, or other biological pathways.

ACKNOWLEDGMENTS

The research of Sheng Luo was supported by National Institute of Aging (grant numbers: R01A G064803, P30AG072958, and P30AG028716).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.;Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.;Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (http://www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by the NIA-funded ADCs: P30 AG019610 (PI Eric Reiman, MD), P30 AG013846 (PI Neil Kowall, MD), P30 AG062428-01 (PI James Leverenz, MD) P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P50 AG047266 (PI

Todd Golde, MD, PhD), P30 AG010133 (PI Andrew Saykin, PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P30 AG062421-01 (PI Bradley Hyman, MD, PhD), P30 AG062422-01 (PI Ronald Petersen, MD, PhD), P50 AG005138 (PI Mary Sano, PhD), P30 AG008051 (PI Thomas Wisniewski, MD), P30 AG013854 (PI Robert Vassar, PhD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG010161 (PI David Bennett, MD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P30 AG062429-01(PI James Brewer, MD, PhD), P50 AG023501 (PI Bruce Miller, MD), P30 AG035982 (PI Russell Swerdlow, MD), P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG049638 (PI Suzanne Craft, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P30 AG062715-01 (PI Sanjay Asthana, MD, FRCP), P50 AG005681 (PI John Morris, MD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD).

Authors' disclosures available online (https://www.j-alz.com/manuscript-disclosures/21-0700r1).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: https://dx.doi.org/10.3233/JAD-210700.

REFERENCES

[1] Karch CM, Goate AM (2015) Alzheimer's disease risk genes and mechanisms of disease pathogenesis. *Biol Psychiatry* **77**, 43-51.

[2] (2020) 2020 Alzheimer's disease facts and figures. *Alzheimers Dement* **16**, 391-460.

[3] Liu CC, Liu CC, Kanekiyo T, Xu H, Bu G (2013) Apolipoprotein E and Alzheimer disease: Risk, mechanisms and therapy. *Nat Rev Neurol* **9**, 106-118.

[4] Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, Myers RH, Pericak-Vance MA, Risch N, van Duijn CM (1997) Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA* **278**, 1349-1356.

[5] Andrews SJ, Fulton-Howard B, Goate A (2020) Interpretation of risk loci from genome-wide association studies of Alzheimer's disease. *Lancet Neurol* **19**, 326-335.

[6] Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, Boland A, Vronskaya M, van der Lee SJ, Amlie-Wolf A, 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.

[7] Ala-Korpela M, Holmes MV (2020) Polygenic risk scores and the prediction of common diseases. *Int J Epidemiol* **49**, 1-3.

[8] Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, Badarinarayan N, Gerad/Perades, consortia I, Morgan K, Passmore P, Holmes C, Powell J, Brayne C, Gill M, Mead S, Goate A, Cruchaga C, Lambert JC, van Duijn C, Maier W, Ramirez A, Holmans P, Jones L, Hardy J, Seshadri S, Schellenberg GD, Amouyel P, Williams J (2015) Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain* **138**, 3673-3684.

[9] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, et al. (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-1458.

[10] Wang T, Han Z, Yang Y, Tian R, Zhou W, Ren P, Wang P, Zong J, Hu Y, Jiang Q (2019) Polygenic risk score for Alzheimer's disease is associated with Ch4 volume in normal subjects. *Front Genet* **10**, 519.

[11] Leonenko G, Sims R, Shoai M, Frizzati A, Bossu P, Spalletta G, Fox NC, Williams J, GERAD consortium, Hardy J, Escott-Price V (2019) Polygenic risk and hazard scores for Alzheimer's disease prediction. *Ann Clin Transl Neurol* **6**, 456-465.

[12] Harrison TM, Mahmood Z, Lau EP, Karacozoff AM, Burggren AC, Small GW, Bookheimer SY (2016) An Alzheimer's disease genetic risk score predicts longitudinal thinning of hippocampal complex subregions in healthy older adults. *eNeuro* **3**, ENEURO.0098-16.2016.

[13] Lupton MK, Strike L, Hansell NK, Wen W, Mather KA, Armstrong NJ, Thalamuthu A, McMahon KL, de Zubicaray GI, Assareh AA, Simmons A, Proitsi P, Powell JF, Montgomery GW, Hibar DP, Westman E, Tsolaki M, Kloszewska I, Soininen H, Mecocci P, Velas B, Lovestone S, Alzheimer's Disease Neuroimaging I, Brodaty H, Ames D, Trollor JN, Martin NG, Thompson PM, Sachdev PS, Wright MJ (2016) The effect of increased genetic risk for Alzheimer's disease on hippocampal and amygdala volume. *Neurobiol Aging* **40**, 68-77.

[14] Darst BF, Koscik RL, Racine AM, Oh JM, Krause RA, Carlsson CM, Zetterberg H, Blennow K, Christian BT, Bendlin BB, Okonkwo OC, Hogan KJ, Hermann BP, Sager MA, Asthana S, Johnson SC, Engelman CD (2017) Pathway-specific polygenic risk scores as predictors of amyloid-beta deposition and cognitive function in a sample at increased risk for Alzheimer's disease. *J Alzheimers Dis* **55**, 473-484.

[15] Morgan AR, Touchard S, O'Hagan C, Sims R, Majounie E, Escott-Price V, Jones L, Williams J, Morgan BP (2017) The correlation between inflammatory biomarkers and polygenic risk score in Alzheimer's disease. *J Alzheimers Dis* **56**, 25-36.

[16] Desikan RS, Fan CC, Wang Y, Schork AJ, Cabral HJ, Cupples LA, Thompson WK, Besser L, Kukull WA, Holland D, Chen CH, Brewer JB, Karow DS, Kauppi K, Witoelar A, Karch CM, Bonham LW, Yokoyama JS, Rosen HJ, Miller BL, Dillon WP, Wilson DM, Hess CP, Pericak-Vance M, Haines JL, Farrer LA, Mayeux R, Hardy J, Goate AM, Hyman BT, Schellenberg GD, McEvoy LK, Andreassen OA, Dale AM (2017) Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score. *PLoS Med* **14**, e1002258.

[17] Petersen RC (2009) Early diagnosis of Alzheimer's disease: Is MCI too late? *Curr Alzheimer Res* **6**, 324-330.

[18] Mitchell AJ, Shiri-Feshki M (2009) Rate of progression of mild cognitive impairment to dementia–meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr Scand* **119**, 252-265.

[19] Torkamani A, Wineinger NE, Topol EJ (2018) The personal and clinical utility of polygenic risk scores. *Nat Rev Genet* **19**, 581-590.

[20] Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buros J, Gallins PJ, Buxbaum JD, Jarvik GP, Crane PK, et al. (2011) Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet* **43**, 436-441.

[21] Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, Sealock J, Karlsson IK, Hagg S, Athanasiu L, Voyle N, Proitsi P, Witoelar A, Stringer S, Aarsland D, Almdahl IS, Andersen F, Bergh S, Bettella F, Bjornsson S, Braekhus A, Brathen G, de Leeuw C, Desikan RS, Djurovic S, Dumitrescu L, Fladby T, Hohman TJ, Jonsson PV, Kiddle SJ, Rongve A, Saltvedt I, Sando SB, Selbaek G, Shoai M, Skene NG, Snaedal J, Stordal E, Ulstein ID, Wang Y, White LR, Hardy J, Hjerling-Leffler J, Sullivan PF, van der Flier WM, Dobson R, Davis LK, Stefansson H, Stefansson K, Pedersen NL, Ripke S, Andreassen OA, Posthuma D (2019) Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet* **51**, 404-413.

[22] Marioni RE, Harris SE, Zhang Q, McRae AF, Hagenaars SP, Hill WD, Davies G, Ritchie CW, Gale CR, Starr JM, Goate AM, Porteous DJ, Yang J, Evans KL, Deary IJ, Wray NR, Visscher PM (2018) GWAS on family history of Alzheimer's disease. *Transl Psychiatry* **8**, 99.

[23] Choi SW, O'Reilly PF (2019) PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience* **8**, giz082.

[24] Prive F, Arbel J, Vilhjalmsson BJ (2020) LDpred2: Better, faster, stronger. *Bioinformatics* **36**, 5424-5431.

[25] Vilhjálmsson BJ, Yang J, Finucane HK, Gusev A, Lindström S, Ripke S, Genovese G, Loh PR, Bhatia G, Do R, Hayeck T, Won HH; Schizophrenia Working Group of the Psychiatric Genomics Consortium, Discovery, Biology, and Risk of Inherited Variants in Breast Cancer (DRIVE) study, Kathiresan S, Pato M, Pato C, Tamimi R, Stahl E, Zaitlen N, Pasaniuc B, Belbin G, Kenny EE, Schierup MH, De Jager P, Patsopoulos NA, McCarroll S, Daly M, Purcell S, Chasman D, Neale B, Goddard M, Visscher PM, Kraft P, Patterson N, Price AL (2015) Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am J Hum Genet* **97**, 576-592.

[26] Huang YT, Yang HI (2017) Causal mediation analysis of survival outcome with multiple mediators. *Epidemiology* **28**, 370-378.

[27] Altmann A, Scelsi MA, Shoai M, de Silva E, Aksman LM, Cash DM, Hardy J, Schott JM (2020) A comprehensive analysis of methods for assessing polygenic burden on Alzheimer's disease pathology and risk beyond APOE. *Brain Commun* **2**, fcz047.

[28] Chaudhury S, Brookes KJ, Patel T, Fallows A, Guetta-Baranes T, Turton JC, Guerreiro R, Bras J, Hardy J, Francis PT, Croucher R, Holmes C, Morgan K, Thomas AJ (2019) Alzheimer's disease polygenic risk score as a predictor of conversion from mild-cognitive impairment. *Transl Psychiatry* **9**, 154.

[29] Cruchaga C, Del-Aguila JL, Saef B, Black K, Fernandez MV, Budde J, Ibanez L, Deming Y, Kapoor M, Tosto G, Mayeux RP, Holtzman DM, Fagan AM, Morris JC, Bateman RJ, Goate AM, Dominantly Inherited Alzheimer N, Disease Neuroimaging I, study N-Lf, Harari O (2018) Polygenic risk score of sporadic late-onset Alzheimer's disease reveals a shared architecture with the familial and early-onset forms. *Alzheimers Dement* **14**, 205-214.

[30] Escott-Price V, Myers A, Huentelman M, Shoai M, Hardy J (2019) Polygenic risk score analysis of Alzheimer's disease in cases without APOE4 or APOE2 alleles. *J Prev Alzheimers Dis* **6**, 16-19.

[31] Escott-Price V, Myers AJ, Huentelman M, Hardy J (2017) Polygenic risk score analysis of pathologically confirmed Alzheimer disease. *Ann Neurol* **82**, 311-314.

[32] Kauppi K, Fan CC, McEvoy LK, Holland D, Tan CH, Chen CH, Andreassen OA, Desikan RS, Dale AM, Alzheimer's Disease Neuroimaging Initiative (2018) Combining polygenic hazard score with volumetric MRI and cognitive measures improves prediction of progression from mild cognitive impairment to Alzheimer's disease. *Front Neurosci* **12**, 260.

[33] Leonenko G, Shoai M, Bellou E, Sims R, Williams J, Hardy J, Escott-Price V, Alzheimer's Disease Neuroimaging Initiative (2019) Genetic risk for Alzheimer disease is distinct from genetic risk for amyloid deposition. *Ann Neurol* **86**, 427-435.

[34] Jack CR, Jr., Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Holtzman DM, Jagust W, Jessen F, Karlawish J, Liu E, Molinuevo JL, Montine T, Phelps C, Rankin KP, Rowe CC, Scheltens P, Siemers E, Snyder HM, Sperling R, Contributors (2018) NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* **14**, 535-562.

[35] Jack CR, Jr., Bennett DA, Blennow K, Carrillo MC, Feldman HH, Frisoni GB, Hampel H, Jagust WJ, Johnson KA, Knopman DS, Petersen RC, Scheltens P, Sperling RA, Dubois B (2016) A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology* **87**, 539-547.

[36] Knopman DS, Haeberlein SB, Carrillo MC, Hendrix JA, Kerchner G, Margolin R, Maruff P, Miller DS, Tong G, Tome MB, Murray ME, Nelson PT, Sano M, Mattsson N, Sultzer DL, Montine TJ, Jack CR, Jr., Kolb H, Petersen RC, Vemuri P, Canniere MZ, Schneider JA, Resnick SM, Romano G, van Harten AC, Wolk DA, Bain LJ, Siemers E (2018) The National Institute on Aging and the Alzheimer's Association Research Framework for Alzheimer's disease: Perspectives from the Research Roundtable. *Alzheimers Dement* **14**, 563-575.

[37] Knopman DS, Jack CR, Jr., Lundt ES, Weigand SD, Vemuri P, Lowe VJ, Kantarci K, Gunter JL, Senjem ML, Mielke MM, Machulda MM, Roberts RO, Boeve BF, Jones DT, Petersen RC (2016) Evolution of neurodegeneration-imaging biomarkers from clinically normal to dementia in the Alzheimer disease spectrum. *Neurobiol Aging* **46**, 32-42.

[38] Elias-Sonnenschein LS, Viechtbauer W, Ramakers IH, Verhey FR, Visser PJ (2011) Predictive value of APOE-epsilon4 allele for progression from MCI to AD-type dementia: A meta-analysis. *J Neurol Neurosurg Psychiatry* **82**, 1149-1156.

[39] Alexopoulos P, Werle L, Roesler J, Thierjung N, Gleixner LS, Yakushev I, Laskaris N, Wagenpfeil S, Gourzis P, Kurz A, Perneczky R, Alzheimer's Disease Neuroimaging Initiative (2016) Conflicting cerebrospinal fluid biomarkers

and progression to dementia due to Alzheimer's disease. *Alzheimers Res Ther* **8**, 51.

[40] Tan CH, Fan CC, Mormino EC, Sugrue LP, Broce IJ, Hess CP, Dillon WP, Bonham LW, Yokoyama JS, Karch CM, Brewer JB, Rabinovici GD, Miller BL, Schellenberg GD, Kauppi K, Feldman HA, Holland D, McEvoy LK, Hyman BT, Bennett DA, Andreassen OA, Dale AM, Desikan RS, Alzheimer's Disease Neuroimaging Initiative (2018) Polygenic hazard score: An enrichment marker for Alzheimer's associated amyloid and tau deposition. *Acta Neuropathol* **135**, 85-93.

[41] Camara H, De-Souza EA (2018) beta-amyloid accumulation slows earlier than expected in preclinical Alzheimer's disease patients. *J Neurosci* **38**, 9123-9125.

[42] Kerminen S, Martin AR, Koskela J, Ruotsalainen SE, Havulinna AS, Surakka I, Palotie A, Perola M, Salomaa V, Daly MJ, Ripatti S, Pirinen M (2019) Geographic variation and bias in the polygenic scores of complex diseases and traits in Finland. *Am J Hum Genet* **104**, 1169-1181.

[43] Squire LR, Zola-Morgan S (1991) The medial temporal lobe memory system. *Science* **253**, 1380-1386.

[44] Braak H, Braak E (1995) Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiol Aging* **16**, 271-278; discussion 278-284.

[45] Krumm S, Kivisaari SL, Probst A, Monsch AU, Reinhardt J, Ulmer S, Stippich C, Kressig RW, Taylor KI (2016) Cortical thinning of parahippocampal subregions in very early Alzheimer's disease. *Neurobiol Aging* **38**, 188-196.

[46] Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, Jack CR, Jr., Jagust WJ, Shaw LM, Toga AW, Trojanowski JQ, Weiner MW (2010) Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. *Neurology* **74**, 201-209.

[47] McDougall F, Edgar C, Mertes M, Delmar P, Fontoura P, Abi-Saab D, Lansdall CJ, Boada M, Doody R (2021) Psychometric properties of the Clinical Dementia Rating - Sum of Boxes and other cognitive and functional outcomes in a prodromal Alzheimer's disease population. *J Prev Alzheimers Dis* **8**, 151-160.

[48] Balsis S, Benge JF, Lowe DA, Geraci L, Doody RS (2015) How do scores on the ADAS-Cog, MMSE, and CDR-SOB correspond? *Clin Neuropsychol* **29**, 1002-1009.