

## Polygenic resilience score may be sensitive to preclinical Alzheimer's disease changes

Jaclyn M. Eissman<sup>1,2</sup>, Greyson Wells<sup>1</sup>, Omair A. Khan<sup>3</sup>, Dandan Liu<sup>3</sup>, Vladislav A. Petyuk<sup>4</sup>, Katherine A. Gifford<sup>1</sup>, Logan Dumitrescu<sup>1,2</sup>, Angela L. Jefferson<sup>1</sup>, and Timothy J. Hohman<sup>1,2†</sup>

<sup>1</sup>*Vanderbilt Memory and Alzheimer's Center, Vanderbilt University Medical Center, Nashville, TN 37212, USA*

<sup>2</sup>*Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN 37212, USA*

<sup>3</sup>*Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN 37212, USA*

<sup>4</sup>*Biological Sciences Division and Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99354, USA*

<sup>†</sup>*Email: timothy.j.hohman@vumc.org*

Late-onset Alzheimer's disease (LOAD) is a polygenic disorder with a long prodromal phase, making early diagnosis challenging. Twin studies estimate LOAD as 60-80% heritable, and while common genetic variants can account for 30% of this heritability, nearly 70% remains "missing". Polygenic risk scores (PRS) leverage combined effects of many loci to predict LOAD risk, but often lack sensitivity to preclinical disease changes, limiting clinical utility. Our group has built and published on a resilience phenotype to model better-than-expected cognition given amyloid pathology burden and hypothesized it may assist in preclinical polygenic risk prediction. Thus, we built a LOAD PRS and a resilience PRS and evaluated both in predicting cognition in a dementia-free cohort (N=254). The LOAD PRS had a significant main effect on baseline memory ( $\beta=-0.18$ ,  $P=1.68E-03$ ). Both the LOAD PRS ( $\beta=-0.03$ ,  $P=1.19E-03$ ) and the resilience PRS ( $\beta=0.02$ ,  $P=0.03$ ) had significant main effects on annual memory decline. The resilience PRS interacted with CSF A $\beta$  on baseline memory ( $\beta=-6.04E-04$ ,  $P=0.02$ ), whereby it predicted baseline memory among A $\beta^+$  individuals ( $\beta=0.44$ ,  $P=0.01$ ) but not among A $\beta^-$  individuals ( $\beta=0.06$ ,  $P=0.46$ ). Excluding *APOE* from PRS resulted in mainly LOAD PRS associations attenuating, but notably the resilience PRS interaction with CSF A $\beta$  and selective prediction among A $\beta^+$  individuals was consistent. Although the resilience PRS is currently somewhat limited in scope from the phenotype's cross-sectional nature, our results suggest that the resilience PRS may be a promising tool in assisting in preclinical disease risk prediction among dementia-free and A $\beta^+$  individuals, though replication and fine-tuning are needed.

*Keywords:* Alzheimer's disease, polygenic risk, resilience, preclinical, cognition

### 1. Introduction

Late-onset Alzheimer's disease (LOAD) is a highly polygenic disorder, characterized by a neuropathological cascade resulting in neurodegeneration and cognitive impairment.<sup>1</sup> Notably, LOAD is characterized by a long prodromal phase in which pathology begins to accumulate prior to the onset of clinical disease. The prodromal stage thus represents decades of pathological changes before cognitive deficits are detected (e.g., dementia), making early clinical dementia diagnosis quite challenging,<sup>1</sup> yet imperative. Additionally, LOAD is a highly heritable trait, with twin studies estimating LOAD heritability to be 60-80%,<sup>2,3</sup> though the source for much of the genetic variation

driving LOAD heritability has yet to be elucidated.<sup>2,3</sup> Genome-wide association studies (GWAS) have been integral in beginning to uncover narrow-sense heritability, defined as the additive genetic component of heritability. As of 2022, LOAD GWAS have identified and replicated 33 risk and protective loci.<sup>2,4</sup> However, the effect sizes of known LOAD GWAS loci are small to moderate,<sup>5</sup> accounting for ~8% of total LOAD heritability, with ~6% out of this ~8% coming from the *APOE*  $\epsilon$ 2 and  $\epsilon$ 4 risk and protective alleles.<sup>5</sup> Furthermore, studies have estimated the portion of LOAD narrow-sense heritability driven by common variants in the population, including and in addition to *APOE*. For example, Ridge and colleagues calculated that ~30% of LOAD phenotypic variance can be explained by a summation of effects of common GWAS variants,<sup>5</sup> suggesting a substantial heritable component remains unexplained or missing.

In recent years, LOAD polygenic risk scores (PRS) have leveraged the effects of multiple genetic loci to predict LOAD risk, but these PRS have not had expected clinical utility. One reason is that LOAD PRS are often built from case/control GWAS, which may represent later-stage disease processes, resulting in a loss of sensitivity when applied to preclinical disease.<sup>6</sup> Thus, LOAD PRS may be most beneficial in identifying symptomatic MCI or LOAD cases.<sup>6</sup> At the same time, some studies have found that LOAD PRS can be built in a sensitive manner to predict MCI or LOAD risk in younger, dementia-free individuals.<sup>7,8</sup> Yet, it also remains unclear if LOAD PRS hold more predictive power than simply predicting genetic risk from *APOE* genotype alone. Many studies have found that LOAD PRS hold predictive power for LOAD risk above and beyond *APOE* genotype,<sup>9,10</sup> while other studies have found that *APOE* genotype is still the best predictor.<sup>6,11</sup>

While neuropathology is a hallmark of LOAD and other related disorders, it is notable that a subset of individuals can maintain normal cognition in the face of neuropathology. In fact, ~30% of elderly adults who meet NIA-AA Reagan neuropathological criteria for AD at autopsy remain cognitively unimpaired throughout life.<sup>12,13</sup> These elderly individuals are characterized as “resilient” in frameworks of cognitive reserve and resilience.<sup>14,15</sup> Our group has defined a continuous measure of resilience, representing better-than-expected cognition given amyloid pathology burden, and leveraged this measure for genomic analysis.<sup>16</sup> The purpose of our original resilience GWAS was to identify common genetic variants that relate to cognition in the face of amyloid. By design, the residual metric of resilience is not correlated with amyloid, but is strongly predictive of future memory performance among people who are  $A\beta$ +.<sup>16,17</sup> Notably, we found resilience to be 20-25% heritable,<sup>17</sup> and found it has a genetic architecture distinct from that of clinical AD.<sup>17</sup>

However, to our knowledge, very few studies have examined polygenic resilience scores for complex traits, but these few have laid a framework for polygenic resilience scores as a tool to study complex, heritable traits. In 2021, Hess and colleagues created a method of calculating a “polygenic resilience score” for schizophrenia. In brief, this method takes marginal SNP effects from a trait, builds a weighted summary score from these SNP associations, and then selects the controls and the cases with the highest scores.<sup>18</sup> Hou and colleagues applied this method to look at LOAD in the context of resilience and observed that a higher polygenic resilience score was associated with lower LOAD risk penetrance among high-risk LOAD individuals.<sup>19</sup> A caveat of Hou and colleague’s study is that their findings attenuated when only examining their score among high-risk *APOE*  $\epsilon$ 4 carriers, and Hou et al. reiterates that PRS contributions above and beyond *APOE* is mixed in the literature.<sup>19</sup>

Additionally, a limitation of this polygenic resilience score method is that it uses trait-based GWAS and binning to determine “resilient” individuals, limiting the scope of the analysis.

We felt we could extend the polygenic resilience score framework by 1) leveraging our continuous, quantitative resilience phenotype 2) clarifying if a resilience PRS could predict risk above and beyond *APOE* 3) examine the relationship of a resilience PRS with amyloid pathology, which has been scarcely analyzed in LOAD PRS studies. Thus, we generated a LOAD<sup>4</sup> PRS and a cognitive resilience<sup>17</sup> PRS. In a dementia-free cohort, we assessed the association of each PRS with baseline memory and with annual memory decline and tested to see if amyloid modified the association of each PRS with memory performance. We hypothesized that while the LOAD PRS would be useful in predicting annual memory decline due to neuropathological build up, the resilience PRS would be more predictive of baseline memory in the presence of amyloid pathology, by differentiating the heterogeneity in memory performance among Aβ<sup>+</sup> individuals.

## 2. Methods

### 2.1. Participants

Participants were recruited as part of a case-control, longitudinal, observational design study, the Vanderbilt Memory and Aging Project (VMAP) which takes place at the Vanderbilt University Medical Center in Nashville, Tennessee.<sup>20</sup> VMAP began in 2012 and recruited individuals who were 60+ years of age, English speakers, had auditory/visual capacity for testing, and had a study partner. Each participant was given a Clinical Dementia Rating (CDR) interview and NIA-AA criteria was leveraged to classify individuals into cognitively unimpaired or mild cognitive impairment (MCI).<sup>20</sup> All protocols for the VMAP cohort were IRB-approved and informed consent for each participant was obtained prior to enrollment. Please see **Table 1** for an overview of the VMAP cohort.

Table 1. VMAP Cohort Demographics.

Cohort Characteristics	
Number of participants	334
Number of participants with genetic data	76.05% (254)
Total number of visits	3.83 +/- 0.76
Longitudinal follow-up (years)	2.27 +/- 1.97
Demographics and Health Characteristics	
Age at baseline (years)	72.74 +/- 6.89
Sex (% female)	27.54% (92)
Education (years)	16.13 +/- 2.56
<i>APOE</i> ε4 (% positive)	26.35% (88)
Amyloid status (% positive)	49.70% (166)
Diagnosis at baseline (% MCI)	29.34% (98)

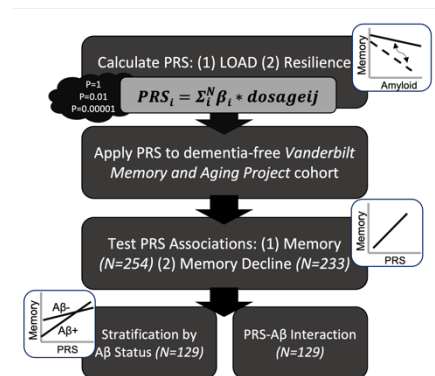


Figure 1. Flow-chart summary of analytical workflow.

### 2.2. Cerebrospinal fluid amyloid

A subset of participants (N=155) consented to and successfully completed lumbar puncture. Cerebral spinal fluid (CSF) was collected, spun down, and supernatant was analyzed through enzyme-linked immunosorbent assays (ELISA). One assay conducted was the INNOTEST® β-

AMYLOID<sub>(1-42)</sub>, which includes autoantibodies for neo-epitopes of amino acids 1 and 42 of the A $\beta$ 1-42 amino acid peptides, ensuring specificity for A $\beta$ 1-42 peptides. Binarized amyloid status was determined for each participant based on CSF A $\beta$ 1-42 measurements. A published cut-point of CSF A $\beta$ 1-42 530ng/L was implemented, thus defining A $\beta$ <sup>+</sup> individuals with CSF A $\beta$ 1-42 values under 530ng/L.<sup>21</sup> A more detailed protocol is described in a prior paper by our group.<sup>20</sup>

### 2.3. Neuropsychological composites

Participants completed a series of neuropsychological tests that covered domains including memory, and a memory composite score was defined in a prior paper by our group.<sup>22</sup> Memory composites were calculated from item-level data, to reduce multiple testing burden. The composite score leveraged test item-level data from the California Verbal Learning Test, Second Edition, and the Biber Figure Learning Test. Composite scores were calculated with a bifactor latent variable model, and final memory composite scores were on a z-score scale.<sup>22</sup>

### 2.4. Genetic data quality control and imputation

Individuals consenting to genotyping (N=333) were genotyped from whole blood on the Illumina MEGA<sup>EX</sup> genotyping array. Raw genetic data were processed as follows. First, variant-level filtering removed variants with >5% missingness, <1% minor allele frequency (MAF), and non-autosomal variants. Next, sample-level filtering removed individuals with >1% missingness, those who were related, those with mismatched self-reported and genetically determined sex, and heterozygosity outliers. Then genetic data were filtered to keep self-reported non-Hispanic white individuals, and genetic ancestry outliers (e.g., principal component analysis – PCA) were removed. Variants were also filtered for Hardy-Weinberg equilibrium (HWE) exact test  $P < 1 \times 10^{-6}$ . Finally, genetic data were lifted over to hg38 and compared and aligned to the Trans-Omics for Precision Medicine (TOPMed) reference panel,<sup>23-25</sup> dropping variants that failed lift-over or mismatched with the reference panel.

Cleaned genetic data were next phased (Eagle phasing) and imputed on the TOPMed imputation server.<sup>23-25</sup> Raw, imputed data were filtered to remove variants with an imputed  $R^2 < 0.8$  or duplicated/multi-allelic variants. Additionally, original genotypes were merged back in with the imputed data. Another HWE exact test was performed filtering for  $P < 1 \times 10^{-6}$ , and variants with MAF <1% were removed. Once again, genetic ancestry outliers determined by a PCA were subsequently filtered. The final, cleaned, imputed VMAP genetic data included 255 non-Hispanic white participants and 8,689,730 variants. Additionally, *APOE* genotypes were determined by the TaqMan genotyping assay for rs7412 and rs429358 performed on DNA extracted from whole blood.<sup>20</sup>

### 2.5. Statistical analyses

See **Figure 1** for an overview of our analytical plan.

#### 2.5.1. Polygenic risk score generation

Two PRS were calculated leveraging Kunkle et al. LOAD case/control genome-wide meta-analysis<sup>4</sup> and our group's recent genome-wide meta-analysis on resilience<sup>17</sup>. No participants in VMAP were

included in either of the original GWAS. First, when applicable, GWAS were lifted to hg38. Next, GWAS variants were compared to the VMAP genetic data. Any ambiguous, palindromic variants were filtered out. Then overlapping variants between the GWAS and the VMAP genetic data were retained and then were compared for variants on opposite strands between the GWAS and the genetic data, and strand differences were resolved. Then, linkage disequilibrium (LD) clumping was performed with PLINK<sup>26</sup> in the VMAP genetic data ( $r^2=0.5$ , window=250kb), to choose the variant with the most significant phenotypic association within each genetically-linked genomic region. Each PRS was built with three different P-value thresholds:  $P=1$ ,  $P=0.01$ , and  $P=0.00001$ , wherein variants were included in the PRS only if their phenotypic association was less than the given threshold. The LD-clumped genetic data were then leveraged to calculate each PRS with PLINK's profile function<sup>27</sup> which calculates scores as follows: Weights were retrieved from the variant associations with LOAD or with resilience from the respective GWAS. For each variant the given weight was multiplied by 0, 1, or 2, based on how many risk alleles an individual had. The summation of this process results in a summary score for an individual. Since *APOE* polymorphism is a robust risk factor for LOAD, PRS were calculated with and without the *APOE* region, defined by a 1Mb region up- and downstream of the *APOE* gene.

### 2.5.2. Baseline and longitudinal linear models

We performed a series of linear models and linear mixed effects models in R (v. 4.2) for each PRS. Fixed effects in our models included baseline age, self-reported sex, and the given PRS. Linear mixed effects models included a PRS-by-interval term, where interval was determined by the difference between a participant's age at each cognitive visit and their baseline age. Additionally, linear mixed effects models allowed slope and intercept to vary for each participant. In addition, we performed identical sets of models with the addition of a PRS-by-amyloid term in linear models and a PRS-by-amyloid-by-interval term for linear mixed effects models, with amyloid measured by the CSF A $\beta$ 1-42 assay outlined above. The outcome of our models were baseline memory or annual memory decline for linear models and linear mixed effect models, respectively. Each set of models above was performed again stratifying by amyloid status. Sensitivity analyses were performed for all models leveraging PRS generated without the *APOE* region.

## 3. Results

We performed a series of linear models and linear mixed effects models investigating each PRS association with baseline memory or annual memory decline, respectively. All main effect associations are presented in **Figure 2** and/or **Table 2**. The LOAD PRS had a significant main effect on baseline memory (**Figure 2A**; **Table 2**), but when *APOE* was excluded from the PRS, this result attenuated to nonsignificant

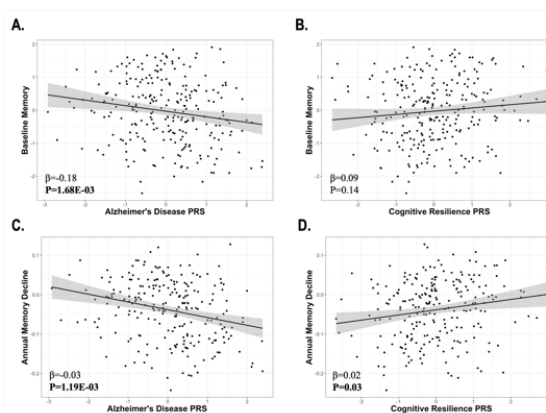


Figure 2. Main effect PRS associations ( $P=0.01$  threshold; with *APOE*) with baseline memory (A, B) and annual memory decline (C, D).

(Table 2). Both the LOAD PRS (Figure 2C; Table 2) and the resilience PRS (Figure 2D; Table 2) had significant main effects on annual memory decline irrespective of *APOE* inclusion in PRS.

Next, we performed a second series of models with a PRS-by-CSF-A $\beta$  interaction term to determine if amyloid modified the association of each PRS with memory performance. Additionally, we performed amyloid status-stratified models to determine if A $\beta$ - individuals or A $\beta$ + individuals (or neither) were driving any observed significant interactions. All CSF-A $\beta$  interaction and amyloid-status stratified results are presented in Figure 3 and/or Table 2.

The LOAD PRS did not interact with CSF A $\beta$  on either baseline memory (Figure 3A; Table 2) or annual memory decline (Figure 3C; Table 2), and this was consistent when *APOE* was excluded from PRS (Table 2).

However, the LOAD PRS significantly predicted annual memory decline more strongly among A $\beta$ + individuals (Figure 3C; Table 2), albeit this result is difficult to interpret with the PRS-by-CSF-A $\beta$  interaction term being nonsignificant. The resilience PRS significantly interacted with CSF A $\beta$  on baseline memory (Figure 3B; Table 2), whereby it significantly predicted baseline memory among A $\beta$ + individuals (Figure 3B; Table 2) but not among A $\beta$ - individuals (Figure 3B; Table 2). These results remained consistent when *APOE* was excluded.

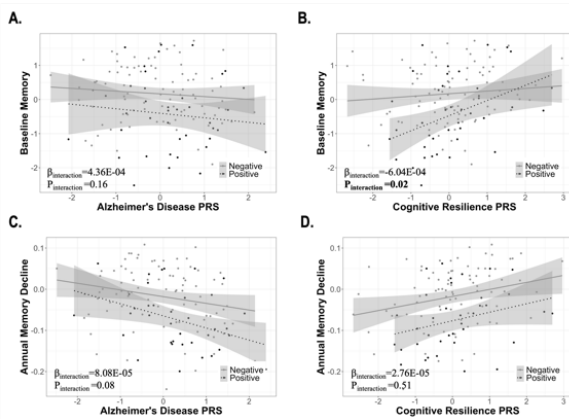


Figure 3. PRS associations ( $P=0.01$  threshold; with *APOE*) with baseline memory (A, B) and annual memory decline (C, D) stratified by A $\beta$  status.

2). All results were consistent across all three thresholds unless denoted in the following paragraph. The LOAD PRS without *APOE* fell just under significance in the main effect association on annual memory decline at the  $P=1$  and  $P=0.00001$  thresholds. The resilience PRS did not have a main effect on annual memory decline at  $P=1$  or  $P=0.00001$  (with or without *APOE*). Additionally, the resilience PRS-by-CSF-A $\beta$  interaction trended significant at  $P=1$ , but still significantly predicted baseline memory among A $\beta$ + individuals. Lastly, both the LOAD PRS and the resilience PRS varied by threshold – and by *APOE* inclusion for the LOAD PRS – in predicting annual memory decline among A $\beta$ - individuals and/or among A $\beta$ + individuals.

Table 2. PRS Associations with Baseline Memory and Annual Memory Decline.

PRS		Baseline Memory							
		Main Effect		A $\beta$ *PRS		A $\beta$ -		A $\beta$ +	
PRS	Threshold	$\beta$	P	$\beta$	P	$\beta$	P	$\beta$	P
LOAD	P=1	-0.13	<b>0.03</b>	2.73E-04	0.41	-0.11	0.21	-0.10	0.59
LOAD	P=0.01	-0.18	<b>1.68E-03</b>	4.36E-04	0.16	-0.08	0.37	-0.13	0.47
LOAD	P=0.00001	-0.25	<b>4.00E-05</b>	5.70E-04	0.13	0.02	0.87	-0.11	0.57
Resilience	P=1	0.09	0.12	-4.97E-04	0.08	0.10	0.21	0.57	<b>1.33E-03*</b>
Resilience	P=0.01	0.09	0.14	-6.04E-04	<b>0.02*</b>	0.06	0.46	0.44	<b>0.01*</b>
Resilience	P=0.00001	0.01	0.88	-7.83E-04	<b>0.02*</b>	-0.09	0.30	0.52	<b>0.02*</b>

<i>PRS</i>		<b>Annual Memory Decline</b>							
		<i>Main Effect</i>		<i>A<math>\beta</math>*PRS</i>		<i>A<math>\beta</math>-</i>		<i>A<math>\beta</math>+</i>	
<u>PRS</u>	<u>Threshold</u>	$\beta$	<u>P</u>	$\beta$	<u>P</u>	$\beta$	<u>P</u>	$\beta$	<u>P</u>
LOAD	P=1	-0.02	<b>0.02</b>	2.59E-05	0.63	-0.03	<b>0.04*</b>	-0.05	<b>0.01*</b>
LOAD	P=0.01	-0.03	<b>1.19E-03*</b>	8.08E-05	0.08	-0.03	<b>0.09#</b>	-0.05	<b>0.01*</b>
LOAD	P=0.00001	-0.03	<b>8.66E-04</b>	7.16E-05	0.21	-0.01	<b>0.54#</b>	-0.02	0.23
Resilience	P=1	4.74E-03	0.60	3.66E-05	0.42	9.88E-04	0.94	0.04	<b>4.69E-02*</b>
Resilience	P=0.01	0.02	<b>0.03*</b>	2.76E-05	0.51	0.03	<b>0.02*</b>	0.02	0.36
Resilience	P=0.00001	0.01	0.39	-6.30E-05	0.25	-2.41E-03	0.87	0.08	<b>6.60E-04*</b>

*Note: P-values with \* remain significant without APOE; # significant without APOE only*

## 4. Discussion

We built a LOAD PRS and a cognitive resilience PRS and evaluated each PRS in predicting memory outcomes among dementia-free elderly individuals. Both sets of PRS provided useful information and performed best in the spheres most closely related to the original phenotype in the GWAS. The LOAD PRS was predictive of annual memory decline in the whole sample and more strongly among A $\beta$ +. In contrast, the resilience PRS was a particularly strong predictor of baseline memory in the presence of amyloid pathology, reflecting that the original phenotype was built to represent better-than-expected memory performance among those with high levels of AD biomarkers. Together, our findings suggest that the complementary information of a resilience PRS could improve preclinical prediction. It also highlights the need to expand sample sizes allowing for incorporation of longitudinal cognitive data into genetic studies of resilience to improve polygenic risk score applications in the future.

### 4.1. LOAD PRS is a strong predictor of annual cognitive decline in later stages of disease

Our main effect findings (**Figure 2; Table 2**) highlight that the LOAD PRS had a significant main effect on both baseline memory and annual memory decline. While the LOAD PRS did not interact with CSF A $\beta$  on baseline memory or annual memory decline (**Figure 3; Table 2**), it more strongly predicted annual memory decline among A $\beta$ + individuals. LOAD PRS associations with cognitive decline have been replicated in other studies. For example, Kauppi and colleagues found that an AD PRS significantly predicted cognitive decline in a cohort of cognitive unimpaired individuals.<sup>28</sup> Ge and colleagues determined that a LOAD PRS predicted cognitive decline among A $\beta$ + cognitively unimpaired and MCI individuals.<sup>29</sup> Likewise, both Tan et al. and Desikan et al. observed that a polygenic hazard score was associated with cognitive decline.<sup>30-32</sup> More specifically, Tan et al. found that those that had a high polygenic hazard score, indicative of high polygenic risk for LOAD, and who were A $\beta$ +, showed steeper cognitive decline.<sup>30-32</sup> Taken together, it may be that the LOAD PRS reflects a number of heterogeneous routes to cognitive impairment that includes AD neuropathology, but also includes some non-AD processes. All the studies mentioned as well as ours, found consistent associations with cognitive decline and stronger associations among A $\beta$ + individuals than among A $\beta$ - individuals, though the difference in our non-demented cohort was negligible at best. It is notable that we did not observe a LOAD PRS-A $\beta$  interaction. Perhaps the LOAD PRS models later stages of disease where A $\beta$  accumulation has already occurred in many individuals and is but one contributor, while other pathways downstream and parallel to amyloidosis

are primarily contributing to cognitive decline. This idea was posited by Carrasquillo and colleagues<sup>6</sup> (and others) and appears to be supported by our findings.

#### **4.2. Resilience PRS is a strong predictor of cognition in earlier stages of disease**

Only the resilience PRS significantly interacted with A $\beta$  on memory performance, whereby it predicted baseline memory among A $\beta$ + individuals but not among A $\beta$ - individuals. (**Figure 3; Table 2**). Notably, previous studies are mixed regarding if LOAD polygenic risk associates with A $\beta$  burden. Multiple studies have found associations between LOAD PRS and amyloid positivity, including Mormino and colleagues who also observed an association between their LOAD PRS and cognitive decline.<sup>7,27,33</sup> Other studies have found no association between a LOAD PRS and amyloid positivity, or an association that attenuated when *APOE* was excluded.<sup>11,29,34,35</sup> It is noteworthy that Ge and colleagues found no association between the LOAD PRS and baseline A $\beta$ , but did find an association of the LOAD PRS with cognitive decline among A $\beta$ +.<sup>29</sup> Ebenau and colleagues comment on the mixed literature surrounding LOAD PRS association with A $\beta$  positivity, pointing to heterogeneity in A $\beta$  progression across diagnostic status as a potential reason for disagreement.<sup>33</sup>

Our original resilience phenotype was designed to predict better-than expected cognition in the presence of amyloid pathology.<sup>16</sup> This matches what we are seeing with the resilience PRS, and the cross-sectional result we see with the PRS matches the cross-sectional nature of the phenotype.<sup>16</sup> A recent study showed that a LOAD PRS enriched for amyloid-positivity-associated loci was associated with cognitive decline, whereas simply a LOAD PRS was not associated.<sup>36</sup> This highlights that loci driving amyloidosis, which begins earlier in disease progression, may not be the same loci driving clinical dementia (downstream).<sup>36</sup> To address this limitation, the resilience PRS may be a complementary tool in this case, as based on our novel results, it can selectively predict baseline memory among A $\beta$ + individuals (**Figure 3; Table 2**). Since much of the elderly population is living with neuropathology,<sup>12</sup> determining those most at risk for future cognitive decline is imperative. Whereas the LOAD PRS may be working through amyloid pathology, performing similarly irrespective of amyloid pathology, the resilience PRS, in contrast, may be interacting with amyloid pathology, predicting genetic risk above and beyond amyloid pathology. It is noteworthy that all individuals in the VMAP cohort were dementia-free. Thus, our resilience PRS may be a tool that can best predict genetic risk for cognitive deficits among biomarker-positive individuals while they are still in the preclinical stage of disease. Our promising initial results indicate that we may have developed a novel PRS that 1) does not lose predictive power among those with A $\beta$  pathology 2) performs its best among this high-risk A $\beta$ + group, separating them out from those in the elderly population who may or may not have A $\beta$  in their brain, and 3) performs robustly irrespective of an individual's future clinical diagnosis. Replicating our findings, incorporating longitudinal data into resilience models, and increasing sample size will be necessary to fine-tune this PRS.

#### **4.3. PRS including more variants may have predictive power beyond the *APOE* locus**

Over the last decade of PRS as a tool for LOAD risk prediction, there has been much debate regarding if a LOAD PRS has more predictive power than *APOE* genotype alone. Studies have been mixed, with many demonstrating that LOAD PRS associate with LOAD risk and LOAD-



endophenotypes above and beyond *APOE*,<sup>9,27,37</sup> while some studies show that LOAD PRS without *APOE* attenuate to nonsignificant in predicting LOAD risk or endophenotype levels.<sup>11,29,34</sup> However, some of the studies that found PRS to contribute to risk prediction beyond that of *APOE* still underscore that *APOE* is contributing a large amount to polygenic risk.<sup>9,37</sup> One study positing that a LOAD PRS has predictive power beyond *APOE* also stated that 43.8% of the 61.0% total predictive power of the LOAD PRS on conversion from MCI to LOAD was coming from *APOE* alone.<sup>9</sup> Notably, our resilience PRS findings remained consistent when *APOE* was removed from PRS calculations, which makes sense as the resilience phenotype attempts to regress out effects of amyloidosis<sup>17</sup> which are often driven by *APOE*.<sup>34</sup> A resilience PRS like the one we built in this study may be promising in terms of its ability to predict LOAD-related cognitive outcomes above and beyond that of *APOE* but replicating our findings and larger sample sizes for future resilience GWAS are needed to fully elucidate this theory.

In addition, there is no gold standard for a singular P-value threshold to leverage for LOAD PRS calculations. Two recent studies examined LOAD PRS at a variety of different thresholds. Ge and colleagues observed fairly consistent results across thresholds spanning from  $P=0.01$  to  $P=1 \times 10^{-7}$ .<sup>38</sup> Another study observed that distinguishing between cognitively unimpaired and LOAD participants was best with a threshold of  $P=0.01$ , and in fact predictive power plateaued after  $P=0.01$ .<sup>27</sup> In this study, we tested three thresholds:  $P=1$ ,  $P=0.01$ , and  $P=0.00001$ . Our results were mostly consistent across the three thresholds, but the resilience PRS at  $P=0.01$  seemed to best predict annual memory decline. Overall, our results combined with some previous studies suggest that perhaps allowing for inclusion of more loci that fall below the stringent genome-wide threshold captures a wider variety of processes contributing to complex trait risk.<sup>18,19,27</sup>

#### 4.4. *Strengths and weaknesses*

Our study had multiple strengths. We leveraged a deeply-phenotyped cohort, the Vanderbilt Memory and Aging Project. This cohort has many important features including participants free of dementia, baseline biomarker status for participants, and longitudinal measurements of memory composite scores. However, our study did have some limitations. Our resilience PRS was not built with inclusion of measures of tau pathology or other known age-related neuropathologies. Sample size is a limiting factor for these measures of pathology, but as sample sizes increase in these cohorts, we plan to incorporate other pathology measures into our resilience models in addition to amyloid. Additionally, our sample size (in VMAP) was limited to those consenting to genotyping, neuropsychological testing, and lumbar puncture. Our study was limited to non-Hispanic white individuals, attenuating the generalizability of our findings to other populations. Currently, genetic data is becoming available for individuals across multiple ancestry groups, allowing groups including ours to expand diversity in GWAS studies, including cross-ancestry approaches. With more diverse GWAS, future studies will be able to build PRS in multiple ancestry groups, which will aid in our understanding of AD genetic risk in diverse populations. Lastly, some of the PRS associations reported in this study did not survive correction for multiple comparisons with the false discovery rate ( $FDR < 0.05$ ) procedure, likely due to power and sample size constraints of the original GWAS. The sample sizes of individuals with cognition, genotyping, and neuropathology data are

ever increasing, which we are leveraging to increase our sample sizes for our resilience GWAS, and this will contribute to increased power in an analysis like this one in the future.

#### 4.5. Conclusions

Although our study needs to be replicated, we find our initial novel findings to be promising that a cognitive resilience PRS may serve as a complementary clinical tool with a LOAD PRS in identifying those most at risk for future cognitive decline while individuals are still in the preclinical and prodromal stages of LOAD.

#### 5. Acknowledgements

Study data were obtained from the Vanderbilt Memory and Aging Project (VMAP). Data were collected by Vanderbilt Memory and Alzheimer's Center Investigators at Vanderbilt University Medical Center. This work was supported by NIA grants R01-AG034962, R01-AG056534, R01-AG062826, K24-AG046373, and Alzheimer's Association IIRG-08-88733. This work was additionally supported by: Swedish Research Council #2018-02532, Swedish Research Council #2018-02532, European Research Council, #681712, Swedish State Support for Clinical Research, #ALFGBG-720931, Swedish State Support for Clinical Research, #ALFGBG-720931, U24-AG074855, R01-AG061518, R01-AG061518, P20-AG068082, U01-AG068057, R21-AG059941, K01-AG049164, K12-HD043483, HHSN311201600276P, RF1-AG059869, R01-AG073439, and T32-GM080178.

#### References

1. Jack CR, Knopman DS, Jagust WJ, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology*. 2010;9(1):119.
2. Bellenguez C, Küçükali F, Jansen IE, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet*. 2022;54(4):412-436. doi:10.1038/s41588-022-01024-z
3. Gatz M, Reynolds CA, Fratiglioni L. Role of genes and environments for explaining alzheimer disease. *Archives of General Psychiatry*. 2006;63(2):168-174.
4. Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A $\beta$ , tau, immunity and lipid processing. *Nature genetics*. 2019;51(3):414-430. doi:10.1038/s41588-019-0358-2
5. Ridge PG, Mukherjee S, Crane PK, Kauwe JSK. Alzheimer's disease: analyzing the missing heritability. *PloS one*. 2013;8(11):e79771.
6. Carrasquillo MM, Crook JE, Pedraza O, et al. Late-onset Alzheimer's risk variants in memory decline, incident mild cognitive impairment, and Alzheimer's disease. *Neurobiol Aging*. 2015;36(1):60-67. doi:10.1016/j.neurobiolaging.2014.07.042
7. Li WW, Wang Z, Fan DY, et al. Association of Polygenic Risk Score with Age at Onset and Cerebrospinal Fluid Biomarkers of Alzheimer's Disease in a Chinese Cohort. *Neurosci Bull*. 2020;36(7):696-704. doi:10.1007/s12264-020-00469-8

8. Logue MW, Panizzon MS, Elman JA, et al. Use of an Alzheimer's disease polygenic risk score to identify mild cognitive impairment in adults in their 50s. *Mol Psychiatry*. 2019;24(3):421-430. doi:10.1038/s41380-018-0030-8
9. Chaudhury S, Brookes KJ, Patel T, et al. Alzheimer's disease polygenic risk score as a predictor of conversion from mild-cognitive impairment. *Transl Psychiatry*. 2019;9(1):154. doi:10.1038/s41398-019-0485-7
10. Ware EB, Faul JD, Mitchell CM, Bakulski KM. Considering the APOE locus in Alzheimer's disease polygenic scores in the Health and Retirement Study: a longitudinal panel study. *BMC Med Genomics*. 2020;13(1):164. doi:10.1186/s12920-020-00815-9
11. Darst BF, Kosciak RL, Racine AM, et al. 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*. 2017;55(2):473-484. doi:10.3233/JAD-160195
12. Sonnen JA, Santa Cruz K, Hemmy LS, et al. Ecology of the aging human brain. *Archives of neurology*. 2011;68(8):1049-1056.
13. Driscoll I, Troncoso J. Asymptomatic Alzheimers Disease: A Prodrome or a State of Resilience? *Current Alzheimer Research*. 2011;8(4):330-335.
14. Stern Y, Arenaza-Urquijo EM, Bartrés-Faz D, et al. Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's & Dementia*. 2020;16(9):1305-1311. doi:https://doi.org/10.1016/j.jalz.2018.07.219
15. Arenaza-Urquijo EM, Vemuri P. Improving the resistance and resilience framework for aging and dementia studies. *Alzheimer's research & therapy*. 2020;12(1):41. doi:10.1186/s13195-020-00609-2
16. Hohman TJ, McLaren DG, Mormino EC, Gifford KA, Libon DJ, Jefferson AL. Asymptomatic Alzheimer disease: Defining resilience. *Neurology*. 2016;87(23):2443-2450.
17. Dumitrescu L, Mahoney ER, Mukherjee S, et al. Genetic variants and functional pathways associated with resilience to Alzheimer's disease. *Brain*. 2020;143(8):2561-2575. doi:10.1093/brain/awaa209
18. Hess JL, Tylee DS, Mattheisen M, et al. A polygenic resilience score moderates the genetic risk for schizophrenia. *Mol Psychiatry*. 2021;26(3):800-815. doi:10.1038/s41380-019-0463-8
19. Hou J, Hess JL, Armstrong N, et al. Polygenic resilience scores capture protective genetic effects for Alzheimer's disease. *Transl Psychiatry*. 2022;12(1):296. doi:10.1038/s41398-022-02055-0
20. Jefferson AL, Gifford KA, Acosta LMY, et al. The Vanderbilt Memory & Aging Project: Study Design and Baseline Cohort Overview. *Journal of Alzheimer's Disease*. 2016;(Preprint):1-20.
21. Skillbäck T, Farahmand BY, Rosén C, et al. Cerebrospinal fluid tau and amyloid- $\beta$ 1-42 in patients with dementia. *Brain*. 2015;138(Pt 9):2716-2731. doi:10.1093/brain/awv181
22. Kresge HA, Khan OA, Wagener MA, et al. Subclinical Compromise in Cardiac Strain Relates to Lower Cognitive Performances in Older Adults. *Journal of the American Heart Association*. 2018;7(4). doi:10.1161/jaha.117.007562
23. Das S, Forer L, Schonherr S, et al. Next-generation genotype imputation service and methods. *Nature genetics*. 2016;48(10):1284-1287. doi:10.1038/ng.3656

24. Fuchsberger C, Abecasis GR, Hinds DA. minimac2: faster genotype imputation. *Bioinformatics*. 2014;31(5):782-784. doi:10.1093/bioinformatics/btu704
25. Taliun D, Harris DN, Kessler MD, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *bioRxiv*. Published online January 1, 2019.
26. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*. 2015;4(1). doi:10.1186/s13742-015-0047-8
27. Mormino EC, Sperling RA, Holmes AJ, et al. Polygenic risk of Alzheimer disease is associated with early- and late-life processes. *Neurology*. 2016;87(5):481-488. doi:10.1212/wnl.0000000000002922
28. Kauppi K, Rönnlund M, Nordin Adolfsson A, Pudas S, Adolfsson R. Effects of polygenic risk for Alzheimer's disease on rate of cognitive decline in normal aging. *Transl Psychiatry*. 2020;10(1):250. doi:10.1038/s41398-020-00934-y
29. Ge T, Sabuncu MR, Smoller JW, Sperling RA, Mormino EC. Dissociable influences of APOE epsilon4 and polygenic risk of AD dementia on amyloid and cognition. *Neurology*. 2018;90(18):e1605-e1612. doi:10.1212/wnl.0000000000005415
30. Tan CH, Hyman BT, Tan JJX, et al. Polygenic hazard scores in preclinical Alzheimer disease. *Annals of neurology*. 2017;82(3):484-488. doi:10.1002/ana.25029
31. Desikan RS, Fan CC, Wang Y, et al. Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score. *PLoS medicine*. 2017;14(3):e1002258. doi:10.1371/journal.pmed.1002258
32. Tan CH, Bonham LW, Fan CC, et al. Polygenic hazard score, amyloid deposition and Alzheimer's neurodegeneration. *Brain*. 2019;142(2):460-470. doi:10.1093/brain/awy327
33. Ebenau JL, van der Lee SJ, Hulsman M, et al. Risk of dementia in APOE ε4 carriers is mitigated by a polygenic risk score. *Alzheimers Dement (Amst)*. 2021;13(1):e12229. doi:10.1002/dad2.12229
34. Leonenko G, Shoai M, Bellou E, et al. Genetic risk for alzheimer disease is distinct from genetic risk for amyloid deposition. *Ann Neurol*. 2019;86(3):427-435. doi:10.1002/ana.25530
35. Tan CH, Fan CC, Mormino EC, et al. Polygenic hazard score: an enrichment marker for Alzheimer's associated amyloid and tau deposition. *Acta Neuropathologica*. 2018;135(1):85-93. doi:10.1007/s00401-017-1789-4
36. Xicota L, Gyorgy B, Grenier-Boley B, et al. Association of APOE-Independent Alzheimer Disease Polygenic Risk Score With Brain Amyloid Deposition in Asymptomatic Older Adults. *Neurology*. Published online May 23, 2022:10.1212/WNL.0000000000200544. doi:10.1212/WNL.0000000000200544
37. Karlsson IK, Escott-Price V, Gatz M, et al. Measuring heritable contributions to Alzheimer's disease: polygenic risk score analysis with twins. *Brain Commun*. 2022;4(1):fcab308. doi:10.1093/braincomms/fcab308
38. Ge T, Sabuncu MR, Smoller JW, Sperling RA, Mormino EC. Dissociable influences of APOE epsilon4 and polygenic risk of AD dementia on amyloid and cognition. *Neurology*. 2018;90(18):e1605-e1612. doi:10.1212/wnl.0000000000005415