Association Between Episodic Memory and Genetic Risk Factors for Alzheimer’s Disease in South Asians from the Longitudinal Aging Study in India–Diagnostic Assessment of Dementia (LASI-DAD)

Jennifer A. Smith, PhD,* Wei Zhao, PhD,* Miao Yu, MS,* Kalee E. Rumfelt, BS,* Priya Moorjani, PhD,**§ Andrea Ganna, PhD,¶ Aparajit B. Dey, MD,∥ Jinkook Lee, PhD,** and Sharon L.R. Kardia, PhD*

BACKGROUND/OBJECTIVES: Genetic factors play an important role in Alzheimer’s disease (AD) and cognitive aging. However, it is unclear whether risk loci identified in European ancestry (EA) populations have similar effects in other groups, such as South Asians.

DESIGN: We investigated the allelic distribution and cognitive associations of 56 known AD risk single-nucleotide polymorphisms (SNPs) identified from three EA genome-wide association studies (EA-GWASs) in a South Asian population. Single SNP and genetic risk score (GRS) associations with measures of episodic memory were assessed.

SETTING: The Diagnostic Assessment of Dementia for the Longitudinal Aging Study in India (LASI-DAD).

RESULTS: Although only a few SNPs were significantly associated with memory scores ($P < .05$), effect estimates from the EA-GWAS and the LASI-DAD showed moderate correlation (0.35–0.88) in the expected direction. GRSs were also associated with memory scores, although percentage variation explained was small (0.1%–0.6%).

CONCLUSIONS: Discrepancies in allele frequencies and cognitive association results suggest that genetic factors found predominantly through EA-GWASs may play a limited role in South Asians. However, the extent of differences in the genetic architecture of AD and cognition in EA and South Asians remains uncertain. There is also a critical need to perform a more comprehensive assessment of the mutational spectrum of South Asia to identify novel genetic variants associated with AD and cognition in this population.

INTRODUCTION

Alzheimer’s disease (AD) is a progressive form of dementia with pronounced impairment in memory. Recently, genetic factors have been linked to AD, and the estimated heritability of AD is as high as 80%.1-3 The
strongest genetic risk factor is the ε4 allele of the apolipoprotein E (APOE) gene, with an odds ratio (OR) of 14.9 in individuals who carry two ε4 alleles versus noncarriers. However, the ε2 allele of APOE is protective from AD.

Genome-wide association studies (GWASs) have identified other genetic risk factors for AD. The first large-scale GWAS of AD, conducted by Lambert and colleagues using 74,046 AD cases and controls, identified 19 AD risk loci in addition to APOE. Recently, Kunkle et al expanded the discovery cohort (N = 94,437) and identified 25 AD risk loci in total, 7 of which were new. A third GWAS, conducted by Jansen et al, added UK Biobank data and included both clinically diagnosed AD and AD-by-proxy (through parental diagnosis) cases (N = 455,258). They identified 29 risk loci, including 12 that were novel. Taken together, these AD GWASs identified 66 unique AD-associated single-nucleotide polymorphisms (SNPs) from 38 loci in addition to APOE (Supplementary Table S1).

Importantly, all three AD GWASs were conducted among European ancestry (EA) participants. Currently, there are no large-scale AD GWASs in Indian or South Asian populations. This is problematic because populations may have different linkage disequilibrium (LD) patterns due to their unique evolutionary history, potentially resulting in genetic heterogeneity across ancestries. Most South Asian groups descended from a mixture of two genetically divergent populations: ancestral North Indians related to Central Asians, Middle Easterners, Caucasians, and Europeans; and ancestral South Indians not related to any groups outside of the subcontinent. As a result, LD structure and allele frequencies differ substantially not only between South Asian and European populations, but also across neighboring groups in India. Further, genetic predictors of dementia may differ between EA and South Asian populations due to differences in environmental factors (including diet, toxicants, or socioeconomic status), which may interact with or overshadow genetic risk factors. Therefore, is not clear whether genetic loci previously identified in EA-GWASs will perform similarly in South Asians.

To assess the transferability of genetic variants discovered through EA-based AD GWASs to South Asians, we evaluated the effect of these variants with cognitive function in participants of the Longitudinal Aging Study in India–Diagnostic Assessment of Dementia (LASI-DAD). We tested 66 unique AD-associated SNPs from three GWASs to present an inclusive set of SNPs, representing both highly replicated, larger-effect AD risk variants and newly discovered, smaller-effect AD risk variants from studies with larger sample sizes and/or broader case definitions. We also evaluated the effect of APOE ε4 and ε2 alleles. Both single SNP association and genetic risk score (GRS) associations were evaluated.

METHODS

Study Population

The LASI is a nationally representative sample of greater than 70,000 adults from India aged 45 years or older, with an inclusive set of SNPs, representing both highly replicated, larger-effect AD risk variants and newly discovered, smaller-effect AD risk variants from studies with larger sample sizes and/or broader case definitions. We also evaluated the effect of APOE ε4 and ε2 alleles. Both single SNP association and genetic risk score (GRS) associations were evaluated.

METHODS

Study Population

The LASI is a nationally representative sample of greater than 70,000 adults from India aged 45 years or older.
scores. We excluded variants in the APOE region from the three GRSs and treated APOE as an independent signal. Each GRS was calculated as $GRS_j = \sum \beta_i x_{ij}$, with $\beta_i$ being the effect size associated with the risk allele for SNP $i$, and $x_{ij}$ being the dosage of the risk allele for SNP $i$ in individual $j$. The effect size of each SNP was calculated as the $\ln(OR)$ reported in the corresponding GWAS article. We assessed whether each GRS was associated with total learning or delayed recall using the regression models above. We then combined the three GRSs into a single multivariable model to assess the total variance in cognitive function explained before and after adding APOE $\varepsilon2$ and $\varepsilon4$.

### Sensitivity Analysis of the APOE Region

Because APOE is known to be the most important risk locus for AD, we also tested the association between all SNPs in APOE (plus 2 kb upstream to capture the promoter region) and memory scores. For SNPs associated with any memory score, we tested for interaction with $\varepsilon4$ and $\varepsilon2$.

### RESULTS

#### Descriptive Statistics

The LASI-DAD sample was 44% male, and study participants were 69 years old (standard deviation (SD) = 7.2 years) on average (Table 1). Most participants attained lower than secondary level education (70.0%), and over half were from a rural area. Participants had a mean of 12.0 and 3.4 words (SD = 5.2 and 2.4 words) for total learning and delayed recall scores, respectively. The correlation between memory scores was $r = 0.73$ ($P < .0001$).

### Genotype and Imputation Quality in LASI-DAD

Among those genotyped, the median call rate was excellent (99.95%), and the estimated error rate was low ($1.5 \times 10^{-6}$). The mean EmpRsq (correlation between the true genotypes and imputed dosages calculated by masking the given SNP) was 0.86 for common variants (minor allele frequency [MAF] > 0.05), indicating relatively high quality, but was lower ($0.66$) for rare variants (MAF $\leq 0.05$, comprising 84.7% of the measured variants).

### Allelic Distribution of the AD Risk Loci

Among the 68 unique SNPs, 27 were directly genotyped on the GSA, and another 29 were successfully imputed with high quality ($r^2 > 0.8$). One SNP was not available because it was not included in the 1000G reference panel, and another 12 SNPs were excluded due to poor imputation quality in LASI-DAD. As a result, 56 unique SNPs were investigated (Supplementary Tables S1 and S2).

Figure 1 compares the risk allele frequencies of all 56 unique SNPs in LASI-DAD versus the corresponding AD GWAS. Although the risk allele frequencies are correlated as expected ($r = 0.91$; $P < .0001$), 47 SNPs had a significantly different allele frequency between LASI-DAD and the AD GWAS samples ($P \leq .05$; Supplementary Table S3). We
Table 2. SNPs with at Least One Nominally Significant Association in LASI-DAD (P ≤ .05), and APOE ε2 and ε4 Alleles

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>GWAS study</th>
<th>Risk allele</th>
<th>Risk AF</th>
<th>Chr</th>
<th>Position</th>
<th>Total Learning Score</th>
<th></th>
<th></th>
<th>Delayed Recall Score</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>rs2830500</td>
<td>ADAMTS1</td>
<td>Kunkle et al⁹</td>
<td>C</td>
<td>0.77</td>
<td>21</td>
<td>28,156,856</td>
<td>−.66</td>
<td>.019</td>
<td>−.5</td>
<td>.059</td>
<td>−.44</td>
</tr>
<tr>
<td>rs10948363</td>
<td>CD2AP</td>
<td>Lambert et al⁸</td>
<td>G</td>
<td>0.20</td>
<td>6</td>
<td>47,487,762</td>
<td>.62</td>
<td>.025</td>
<td>.52</td>
<td>.041</td>
<td>.4</td>
</tr>
<tr>
<td>rs9473117</td>
<td>CD2AP</td>
<td>Kunkle et al⁹</td>
<td>C</td>
<td>0.18</td>
<td>6</td>
<td>47,431,284</td>
<td>.57</td>
<td>.064</td>
<td>.48</td>
<td>.091</td>
<td>.4</td>
</tr>
<tr>
<td>rs4147929</td>
<td>ABCA7</td>
<td>Lambert et al⁸</td>
<td>A</td>
<td>0.20</td>
<td>19</td>
<td>1,063,443</td>
<td>.55</td>
<td>.062</td>
<td>.67</td>
<td>.015</td>
<td>.24</td>
</tr>
<tr>
<td>rs3752246</td>
<td>ABCA7</td>
<td>Kunkle et al⁹</td>
<td>G</td>
<td>0.20</td>
<td>19</td>
<td>1,056,492</td>
<td>.51</td>
<td>.098</td>
<td>.59</td>
<td>.037</td>
<td>.2</td>
</tr>
<tr>
<td>rs1859788</td>
<td>ZCWPW1</td>
<td>Jansen et al¹⁰</td>
<td>G</td>
<td>0.72</td>
<td>7</td>
<td>99,971,834</td>
<td>−.59</td>
<td>.018</td>
<td>−.41</td>
<td>.072</td>
<td>−.1</td>
</tr>
<tr>
<td>rs6733839</td>
<td>BIN1</td>
<td>Lambert et al⁸ and Kunkle et al⁹</td>
<td>T</td>
<td>0.41</td>
<td>2</td>
<td>127,892,810</td>
<td>−.43</td>
<td>.059</td>
<td>−.44</td>
<td>.034</td>
<td>−.18</td>
</tr>
<tr>
<td>rs4663105</td>
<td>BIN1</td>
<td>Jansen et al¹⁰</td>
<td>C</td>
<td>0.49</td>
<td>2</td>
<td>127,891,427</td>
<td>−.39</td>
<td>.078</td>
<td>−.43</td>
<td>.036</td>
<td>−.18</td>
</tr>
<tr>
<td>rs7185636</td>
<td>IQCK</td>
<td>Kunkle et al⁹</td>
<td>T</td>
<td>0.69</td>
<td>16</td>
<td>19,808,163</td>
<td>−.59</td>
<td>.017</td>
<td>−.53</td>
<td>.021</td>
<td>−.10</td>
</tr>
<tr>
<td>rs11218343</td>
<td>SORL1</td>
<td>Lambert et al⁸, Kunkle et al⁹ and Jansen et al¹⁰</td>
<td>T</td>
<td>0.92</td>
<td>11</td>
<td>121,435,587</td>
<td>−.83</td>
<td>.046</td>
<td>−.89</td>
<td>.02</td>
<td>−.56</td>
</tr>
<tr>
<td>rs429358</td>
<td>APOE</td>
<td>ε4</td>
<td>C</td>
<td>0.10</td>
<td>19</td>
<td>45,411,941</td>
<td>−.30</td>
<td>.438</td>
<td>−.20</td>
<td>.572</td>
<td>−.15</td>
</tr>
<tr>
<td>rs7412</td>
<td>APOE</td>
<td>ε2</td>
<td>C</td>
<td>0.95</td>
<td>19</td>
<td>45,412,079</td>
<td>−.49</td>
<td>.36</td>
<td>−.19</td>
<td>.702</td>
<td>−.17</td>
</tr>
</tbody>
</table>

Note: P ≤ .05 is indicated by bold text. Effects sizes (β values) are calculated with respect to the Alzheimer’s disease risk allele.
Abbreviations: AF, allele frequency; Chr, chromosome; GWAS, genome-wide association study; LASI-DAD, Longitudinal Aging Study in India–Diagnostic Assessment of Dementia; SNP, single-nucleotide polymorphism.
recall score (Table 2). Among the 10 SNPs, 3 (rs2830500 nominally associated with total learning and/or delayed are presented in Table 2. Of the 56 SNPs assessed, 10 were

**Association Between Single SNPs and Cognitive Measures**

Associations between each AD risk SNP and each cognitive measure were assessed using two models. Model 1 adjusted for age, sex, and the top 10 genetic PCs, and model 2 additionally adjusted for education. Q-Q plots indicate that in both models, P values from the 56 SNPs were generally smaller than expected by chance alone (Supplementary Figure S1). This suggests some evidence of association between the AD risk SNPs and cognitive measures.

All nominally significant association results (P \( \leq .05 \)) are presented in Table 2. Of the 56 SNPs assessed, 10 were nominally associated with total learning and/or delayed recall score (Table 2). Among the 10 SNPs, 3 (rs2830500 (ADAMTS1), rs10948363 (CD2AP6), and rs11218343 (SORL1)) were associated with both memory scores in model 1. Two SNPs (rs1859788 (ZCWPW1) and rs7185636 (IQCK2)) were associated with total learning score only, and one SNP (rs9473117 (CD2AP)) was associated with delayed recall score only in model 1. Significance was attenuated for rs2830500 (ADAMTS1) and rs1859788 (ZCWPW1) for total learning score after controlling for education (model 2). In contrast, four SNPs (rs4147929 and rs2752246 in ABCA7 and rs6733839 and rs4663105 in BDN1) show a stronger association with total learning and/or delayed recall in model 2 than model 1. However, no associations remained significant after Bonferroni correction (P \( \leq 8.9 \times 10^{-4} \)). Among the 10 SNPs, most demonstrated an association in the expected direction, where AD risk alleles were associated with lower memory score. Neither the APOE ε2 nor ε4 allele was associated (P \( \leq .05 \)) with memory scores, although the effect directions were as expected.

Although only a small number of the 56 SNPs were significant, most had the expected effect directions (62.5% for total learning and 57.1% for delayed recall in model 1; Supplementary Table S4). AD risk alleles were generally associated with decreased cognition in LASI-DAD (r between effect estimates ranges from −0.35 to −0.88; Supplementary Table S5, Supplementary Figure S2). SNPs with larger GWAS effect sizes tended to have the expected direction of effect in our study; however, those with smaller effect sizes did not always have the expected effect direction. Thus, the relatively strong correlation may not necessarily reflect consistent effects of AD risk variants on cognition.

Among 56 tested SNPs, 6 interacted with age at P \( \leq .05 \) in one or multiple models (Supplementary Table S6), but none was significant after Bonferroni correction, suggesting that the SNP effects do not vary substantially by age.

**Association Between GRSs and Cognitive Measures**
The distribution of the three GRSs (constructed from three AD GWASs) are presented in Supplementary Figure S3. As

### Table 3. Correlations of GRSs Calculated from Alzheimer’s Disease GWASs

<table>
<thead>
<tr>
<th>Correlation</th>
<th>GRS (Lambert et al)</th>
<th>GRS (Kunkle et al)</th>
<th>GRS (Jansen et al)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRS (Lambert et al)</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRS (Kunkle et al)</td>
<td>0.798</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>GRS (Jansen et al)</td>
<td>0.567</td>
<td>0.592</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: All GRSs are standardized to a N(0,1) distribution. Abbreviations: GRS, genetic risk score; GWAS, genome-wide association study.

### Table 4. GRS Associations with Cognitive Function

| Article | Total learning score | | Delayed recall score | | |
|---------|----------------------| |----------------------| | |
| | Model 1 | Model 2 | Model 1 | Model 2 | |
| | \( \beta \) | \( P \) value | \( R^2 \), % | \( \beta \) | \( P \) value | \( R^2 \), % | \( \beta \) | \( P \) value | \( R^2 \), % | \( \beta \) | \( P \) value | \( R^2 \), % |
| GRS (Lambert et al) | \(-.313\) | .050 | 0.4 | \(-.252\) | .089 | 0.2 | \(-.094\) | .219 | 0.1 | \(-.066\) | .361 | 0.1 |
| GRS (Kunkle et al) | \(-.420\) | .008 | 0.6 | \(-.317\) | .032 | 0.4 | \(-.179\) | .018 | 0.5 | \(-.136\) | .059 | 0.3 |
| GRS (Jansen et al) | \(-.215\) | .178 | 0.2 | \(-.145\) | .330 | 0.1 | \(-.068\) | .376 | 0.1 | \(-.039\) | .588 | 0.02 |
| GRS (Lambert et al) + GRS (Kunkle et al) | 0.7 | | 0.4 | | | | | | | | |
| GRS (Lambert et al) + GRS (Jansen et al) | 0.7 | | 0.4 | | | | | | | | |
| GRS (Lambert et al) + GRS (Kunkle et al) + GRS (Jansen et al) | 0.8 | | 0.4 | | | | | | | | |

Note: P \( \leq .05 \) is indicated by bold text. GRS effects sizes (\( \beta \) values) are calculated with respect to an increasing number of Alzheimer’s disease risk alleles. Model 1: cognitive measure \( \sim \) GRS + sex + age + principal component (PC) 1 – PC10. Model 2: cognitive measure \( \sim \) GRS + sex + age + education + PC1 – PC10. \( R^2 \) represents the variance in the cognitive measure explained by the individual GRS, three GRSs combined, or GRSs combined with APOE variants. Abbreviation: GRS, genetic risk score.
expected, the GRSs were positively correlated with each other (Table 3). However, none of the correlations was strong, suggesting that they capture similar but distinct genetic risk profiles in LASI-DAD.

In regression model 1, the Lambert et al.\textsuperscript{8} GRS and the Kunkle et al.\textsuperscript{9} GRS were negatively associated with total learning score (Table 4), and the Kunkle et al.\textsuperscript{9} GRS was negatively associated with delayed recall score. In model 2, the Kunkle et al.\textsuperscript{9} GRS was associated with total learning score, but all other GRSs were attenuated. The proportion of variance in total learning score explained by each GRS ranged from 0.2% to 0.6% in model 1 and from 0.1% to 0.4% in model 2. When combined together, the three GRSs explained 0.7% and 0.4% of variance in model 1 and model 2, respectively. When further combined with APOE e4 and e2, they explained 0.8% in model 1 and 0.4% in model 2. Compared with total learning score, the proportion of variance in delayed recall score explained by each GRS showed a similar pattern but was smaller.

### Sensitivity Analysis of the APOE Region

There were 10 SNPs within APOE (plus the 2-kb promoter) region in LASI-DAD. Among the 10 SNPs, all 3 SNPs in the promoter region (rs7256173, rs7259620, and rs405509) and 2 SNPs within the gene (rs440446 and rs769450) were nominally associated with total learning score in models 1 and/or 2 at \(P \leq 0.05\) (Supplementary Table S7). Two SNPs (rs7259620 and rs440446) remained significant after Bonferroni correction \((P \leq 0.005)\).

We further tested the interaction of these SNPs with both rs429358 (e4) and rs7412 (e2). We observed a nominally significant interaction \((P \leq 0.05)\) between rs7256173 and rs429358 (e4) for both total learning and delayed recall scores in models 1 and 2, and a significant interaction between rs440446 and rs7412 (e2) on delayed recall score in model 2 (Supplementary Table S8). The interaction between rs7256173 and rs429358 (e4) on total learning score remained significant after Bonferroni correction \((P \leq 0.01)\). Suggestive evidence of interactions \((P < 0.1)\) were observed for rs405509 by rs429358 (e4) on total learning score in model 2, rs440446 by rs7412 (e2) on both memory scores in models 1 and 2, and rs405509 by rs7412 (e2) for both memory scores in model 2.

### Discussion

We investigated the association between AD-associated SNPs identified from EA-GWASs and cognitive function in a South Asian cohort. Many of the AD risk variants had different allele frequencies in LASI-DAD compared with EA-GWAS samples. This could be due to differences in study design. Neither LASI-DAD nor the EA-GWASs are fully population-representative. EA-GWAS samples are mostly case-control studies, and LASI-DAD is a sub-study of the population-representative LASI study, with oversampling those at high risk for cognitive impairment. On the other hand, there could be true differences in allele frequency between the two populations. This would not be surprising due to strong founder events in Indian/South Asian populations.\textsuperscript{18} Differences in LD patterns between European and Indian/South Asian ancestries may also cause different SNPs at the same genes/loci to be more strongly associated with AD in each population. Further, because Indians/South Asians derive more than half of their ancestry from a founding population that no longer exists in an unmixed form and is significantly divergent from other extant populations, it is unlikely that all of the variants identified in the EA populations are directly transferrable to Indian/South Asian populations.\textsuperscript{11,14}

Genetic predictors of dementia may also differ between populations when there are major differences in environmental risk factors. Dementia and cognitive decline may result from AD-related neurodegeneration, but also from other common disease processes, such as cardiovascular disease. The prevalence of hypertension has increased rapidly in India due to longer life expectancy and westernization in lifestyle.\textsuperscript{19} Previous studies indicate a larger burden of vascular dementia over AD in Asian populations, whereas AD-related dementia is more prevalent in EA.\textsuperscript{20} Also, LASI-DAD participants differ in other ways from EA-GWAS participants, including diet (e.g., enriched with curcumin spice), natural environment (e.g., greater exposure to various pollutants), and social environment (e.g., lower socioeconomic status and education levels). These environmental factors may be associated with dementia/cognition alone and/or through interaction with genetic risk factors.\textsuperscript{21-25} Thus, the underlying genetic risk factors may differ in these populations. It is therefore critical to validate the effects of AD risk variants in Indians/South Asians.

Although India is the second largest country in the world, it is rarely represented in genomic studies.\textsuperscript{12} The 1000 Genomes Project only contains a small proportion (13.6%) of Indians/South Asians, and they are not well represented in any other commonly used large reference panel, including the Trans-Omics for Precision Medicine and the Haplotype Reference Consortium.\textsuperscript{26} As a result, some of the SNPs had low imputation quality in LASI-DAD, and thus could not be evaluated in this study. Imputation quality was especially low for rare variants, as has been previously demonstrated.\textsuperscript{27} Therefore, large-scale, population-based Indian/South Asian sequencing projects are needed to attain a representative reference panel that would allow for higher-quality imputation for future Indian/South Asian-based studies and facilitate efforts to identify new variants linked to AD and other diseases.

This study suggests that EA-GWAS AD risk variants confer a small amount of genetic risk for decreased cognitive function in Indians/South Asians. Therefore, previously identified AD risk variants may not be the most prominent or the only risk variants in Indians/South Asians. At a nominal \(P\) value, risk variants in ADAMTS1, ZCWPW1, IQCK, B1NI, and SORL1 were associated with decreased cognitive function in LASI-DAD. Among them, B1NI, SORL1, and ZCWPW1 were the early identified risk loci in Lambert et al.\textsuperscript{8} B1NI is one of the most strongly associated loci for AD after APOE. The gene is expressed primarily in the central nervous system and is believed to activate a caspase-independent apoptotic process.\textsuperscript{28} ZCWPW1 is a histone modification reader and potentially involved in sphingosine 3-kinase signaling pathways in neurons.\textsuperscript{29} The association between both loci and AD has been reported in East Asians, although the implicated SNPs were not the same.\textsuperscript{29,30} Interestingly, the variant we examined at...
the ZCWPW1 locus (rs1859788) is a missense mutation and has been recently suggested to likely be a causal allele for AD. SORL1 encodes a type I transmembrane protein that helps facilitate lipid absorption through endocytosis and sorting of fats. The T allele of rs1121834 in SORL1 was associated with increased risk of late-onset AD in East Asians. The suggestive evidence of associations observed in LASI-DAD suggest that these loci are likely to have similar function in Indians/South Asians and EA.

ADAMTS1 and IQCK were novel loci that were identified by the recent AD GWASs (ADAMTS1 in both Kunkle et al and Jansen et al). ADAMTS1 encodes a zinc-binding enzyme excreted in several adult tissues that is required for normal ovulation and renal function, and has been linked to breast cancer. IQCK encodes a protein with an IQ motif that allows for binding of EF-hand proteins and has been linked to obsessive compulsive disorder. Additional studies are needed to understand the roles of these genes in AD and cognition.

Interestingly, two SNPs from CD2AP and two SNPs from ABCA7 were nominally associated with increases in total learning and delayed recall scores. CD2AP encodes a scaffolding protein that supports the integrity of intercellular junctions. ABCA7 encodes a multispan transmembrane protein that is highly expressed in the brain, facilitating transport of phospholipids and cholesterol across cell membranes. Associations between SNPs in these two genes and AD in East Asian cohorts have shown mixed results. Similar to our study, a study in approximately 850 Koreans found that several AD risk variants in Europeans were actually protective for AD at a nominal significance level, including the A allele of rs4147932, which also shows the unexpected direction of effect in our study as well. Thus, CD2AP and ABCA7 likely play a role in AD and/or cognition in Indians/South Asians, although the involved risk variants may differ from EA. We also noted that none of the results from these genes remained significant after multiple testing correction in our study, and thus should be interpreted cautiously.

The two APOE major isoforms were not associated with memory scores in LASI-DAD, although the effect estimates were in the expected direction. APOE is the strongest genetic risk locus for AD in multiple populations, including Asians. Two recent meta-analyses show that APOE ε4 is associated with AD in Indians, with similar effect sizes as in EA (OR ranging from 4.14 to 5.90). Interestingly, one study found that APOE ε3 instead of ε2 showed suggestive evidence of protection against AD in Indians. The lack of association between APOE major isoforms and cognition in this study may be due to our examination of memory scores rather than diagnosed AD, which are not necessarily good predictors of AD risk. Prior studies found that the association between APOE and memory performance/cognition was much weaker than with AD. In addition, the association may be age dependent, with stronger associations at older ages. Further, ε4 is more strongly associated with story-based than word-based memory tasks. Consistent with our study, multiple studies in Asians failed to find an association between APOE variants and cognition. Although poor episodic memory/cognition is a key symptom of AD, it may also be a consequence of other health conditions, such as Huntington’s disease, Parkinson’s disease, hypertension, and others. Thus, lack of association for APOE major isoforms, and possibly for other AD risk variants, may be due to the outcomes investigated in this study.

In addition to APOE major isoforms, many studies suggest that the variants in the promoter region and intron 1, in particular rs405509 and rs440446, were associated with cognitive function/AD independently or through interaction with the major APOE isoforms. Consistent with this literature, the same SNPs along with three other variants (two in the promoter region) were associated with total learning score in this study. Furthermore, we observed a significant interaction between a promoter SNP and ε4 on both total learning score and delayed recall score as well as suggestive evidence of interaction that includes both rs405509 and rs440446. This evidence suggests that APOE is probably an important risk locus for dementia in Indians/South Asians as well, but that multiple variants and mechanisms might be involved. Nonetheless, given that we investigated memory rather than AD, there is still tremendous uncertainty regarding the effect of these genetic risk factors on AD in South Asians. Future studies that focus on clinically diagnosed AD in this population will be the essential next step to compare the SNP effect sizes across populations and better understand the genetic architecture of AD in South Asians.

Although most AD risk variants were not strongly associated with cognitive function, GRSSs composed of these variants did show association with cognitive function in LASI-DAD. The GRS constructed from Kunkle et al performed best among all three GRSSs, possibly because this study is the best-powered GWAS with clinically diagnosed AD cases. Although Jansen et al have a much larger sample size, the case definition for this GWAS included both clinically diagnosed AD as well as AD by proxy (parental diagnosis). The additional cases identified through proxy may have introduced bias or noise to the analysis, reducing predictive ability of the GRS from this GWAS. Another contributing factor may be that approximately one-third of the risk variants from Jansen et al were from novel loci with potentially small and hard-to-replicate effects. Thus, the GRS from Jansen et al may have lower signal/noise ratio compared with other GRSSs. The combined GRSSs plus APOE ε4 and ε2 accounted for less than 1% of the variance in LASI-DAD memory scores. This explains drastically less than the 16% AD variance explained by AD risk variants in EA (APOE ε4 and ε2 explained 13%, whereas other genes explained 3%). Given the low variance explained in LASI-DAD, we caution against applying EA-based genetic knowledge to Indians/South Asians in a clinical setting.

In summary, this was the first study to comprehensively survey all AD risk variants identified in three large-scale AD GWASs and examine their association with cognitive function in an Indian/South Asian population. This study demonstrated some evidence of association between AD risk variants and lower cognition, although the effect sizes were small. A GRSS of the known AD risk variants explained less than a percentage of the total variance in cognitive function, suggesting that many AD risk loci in Indians/South Asians may remain to be identified. Therefore, future large-scale AD GWASs and whole genome sequence analysis in Indian/South Asian samples are warranted to better characterize and understand the genetic cause of AD in this population.
**ACKNOWLEDGEMENTS**

Financial Disclosure: This project is funded by the National Institute on Aging (R01 AG051125 and R1F AG055273). The National Institute on Aging had no role in preparing the data or the manuscript.

Conflict of Interest: The authors have no conflicts of interest to disclose.


**Sponsor’s Role:** The study sponsors had no role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript.

**REFERENCES**


SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.