

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <http://orca.cf.ac.uk/110049/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Ahmad, Shahzad, Bannister, Christian, van der Lee, Sven J., Vojinovic, Dina, Adams, Hieab H.H., Ramirez, Alfredo, Escott-Price, Valentina, Sims, Rebecca, Baker, Emily, Williams, Julie, Holmans, Peter, Vernooij, Meike W., Ikram, M. Arfan, Amin, Najaf and van Duijn, Cornelia M. 2018. Disentangling the biological pathways involved in early features of Alzheimer's disease in the Rotterdam Study. *Alzheimer's and Dementia* 14 (7) , pp. 848-857. 10.1016/j.jalz.2018.01.005 file

Publishers page: <http://dx.doi.org/10.1016/j.jalz.2018.01.005>
<<http://dx.doi.org/10.1016/j.jalz.2018.01.005>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



1 Disentangling the biological pathways involved in early features of
2 Alzheimer's disease in the Rotterdam Study
3

4 **Authors list:**

5 Shahzad Ahmad^a, Christian Bannister^b, Sven J. van der Lee^a, Dina Vojinovic^a, Hieab H.H. Adams^{a,c},
6 Alfredo Ramirez^{d,e,f}, Valentina Escott-Price^b, Rebecca Sims^b, Emily Baker^b, Julie Williams^b, Peter
7 Holmans^b, Meike W. Vernooij^{a,c}, M. Arfan Ikram^{a,c}, Najaf Amin^a, Cornelia M. Van Duijn*^a

8 **Affiliations:**

9 ^a Department of Epidemiology, Erasmus Medical Centre, Rotterdam, Netherlands

10 ^b MRC Centre for Neuropsychiatric Genetics & Genomics, Institute of Psychological Medicine and
11 Clinical Neurosciences, Cardiff University, Cardiff, United Kingdom

12 ^c Department of Radiology and Nuclear Medicine, Erasmus Medical Centre, Rotterdam, Netherlands

13 ^d Department of Psychiatry and Psychotherapy, University of Bonn, Bonn, Germany

14 ^e Department of Psychiatry and Psychotherapy, University Hospital Cologne, Cologne, Germany

15 ^f Institute of Human Genetics, University of Bonn, Bonn, Germany
16
17

18 *Corresponding Author

19 Cornelia M. Van Duijn

20 Department of Epidemiology

21 Erasmus MC, Rotterdam

22 Direct dial: +31107043394

23 Email: c.vanduijn@erasmusmc.nl

24 Postal Address: P.O. box 2040

25 3000CA Rotterdam, the Netherlands
26
27

28 **Abstract**

29 **INTRODUCTION,**

30 Exploring the role of Alzheimer's disease (AD) implicated pathways in pre-dementia phase may
31 provide new insight for preventive and clinical trials targeting disease specific pathways.

32 **METHODS,**

33 We constructed weighted Genetic Risk Scores, first based on 20 genome-wide significant AD
34 risk variants, second clustering these variants within pathways. Risk scores were investigated
35 for their association with AD, mild cognitive impairment (MCI) and brain magnetic resonance
36 imaging phenotypes including white matter lesions, hippocampal volume and brain volume.

37 **RESULTS,**

38 The risk score capturing *endocytosis* pathway was significantly associated with MCI ($P =$
39 1.44×10^{-4}). *Immune response* ($P = 0.016$) and *clathrin/AP2 adaptor complex* pathway ($P =$
40 3.55×10^{-3}) excluding apolipoprotein E (*APOE*) also showed modest association with white
41 matter lesions but did not sustain Bonferroni correction ($P = 9.09 \times 10^{-4}$).

42 **DISCUSSION,**

43 Our study suggests that the clinical spectrum of early AD pathology is explained by different
44 biological pathways in particular, the *endocytosis*, *immune response* and *clathrin/AP2 adaptor*
45 *complex* pathways that are independent of *APOE*.

46 **Keywords:** Genetic Risk Score, Alzheimer's disease, White matter lesions, Mild cognitive
47 impairment, Endocytosis, Immune response

48 **Introduction**

49 Alzheimer's disease (AD) is heterogeneous and genetically complex disease with a high
50 heritability (56-79 %) [1]. It has been known since the end of the previous century that a
51 polymorphism in the apolipoprotein E (*APOE*) gene is the strongest common genetic risk factor
52 [2-4]. This finding fueled speculations on the role of *lipid metabolism* and *cholesterol transport*
53 pathway in AD in addition to the *amyloid cascade* and *tau phosphorylation* mechanism [5, 6].
54 Furthermore, large-scale genome-wide association studies (GWAS) have discovered over 20
55 novel common genetic variants that influence the risk of late-onset AD [7-13]. These common
56 genetic variants have been mapped to eight biological pathways including *immune response*,
57 *endocytosis*, *cholesterol transport*, *hematopoietic cell lineage*, *protein ubiquitination*,
58 *hemostasis*, *clathrin/AP2 adaptor complex* and *protein folding*, each having a distinct biological
59 function [14-16]. These eight pathways are not independent in that genes may be part of more
60 than one biological pathway. For instance, *APOE* is part of four of the eight pathways namely
61 *cholesterol transport*, *hematopoietic cell lineage*, *clathrin/AP2 adaptor complex* and *protein*
62 *folding* pathways; clusterin (*CLU*) encoding for apolipoprotein J is involved in six pathways;
63 phosphatidylinositol binding clathrin assembly protein(*PICALM*) and complement factor 1 (*CR1*)
64 are involved in 2 pathways [14-16].

65 These diverse biological pathways may be responsible for the clinically heterogeneous
66 manifestation of AD [17-19], which include endophenotypes such as changes in structural and
67 functional magnetic resonance imaging (MRI) phenotypes, most notably hippocampal volume,
68 total brain volume and white matter lesions [20, 21]. Furthermore, these biological pathways
69 may also modulate the prodromal stages of AD such as mild cognitive impairment (MCI) [22-

70 24]. Owing to heterogeneity during pre-dementia phase, one important unanswered question is
71 whether the different biological pathways that are implicated in AD relate to the pleiotropy of
72 clinical endophenotypes. We hypothesized that some biological pathways are involved in
73 distinct clinical endophenotypes while others may be involved in multiple or even all.
74 Disentangling the connection of biological pathways to various aspects of AD related early
75 pathology may be a crucial step towards improving our understanding of the pathogenesis of
76 AD and a first step towards a more informative and powerful read-out for preventive and
77 therapeutic trials targeting specific pathways.

78 The current study aims to capture the different biological pathways involved in AD using
79 genetic risk scores to evaluate their role in AD and pre-dementia endophenotypes including
80 MCI, white matter lesion, total brain and hippocampal volume.

81 **Methodology**

82 **Study population**

83 This study included samples from the Rotterdam study (RS). RS is a prospective population
84 based study [25] designed to investigate the etiology of age related disorders. At the baseline
85 examination in 1990-93, study recruited 7983 subjects ≥ 55 years of age from the Ommoord
86 district of Rotterdam (RS-I). At the baseline entry and after every 3 to 4 years, all the study
87 participants were extensively interviewed and physically examined at the dedicated research
88 center. During 2000 to 2001, the baseline cohort (RS-I) was expanded by adding 3011 subjects
89 ≥ 55 years of age, who were not yet part of RS-I (RS-II). Second expansion of RS was performed
90 by recruiting 3932 persons having ≥ 45 years of age during 2006-2008 (RS-III). The study has

91 been approved by the Medical Ethical Committee of Erasmus Medical Center and by the
92 Ministry of Health, Welfare and Sport of the Netherlands. Written Informed consents were also
93 obtained from each study participant to participate and to collect information from their
94 treating physicians. Details of AD, dementia and MCI diagnosis is provided elsewhere [26, 27].
95 In the current analyses, we included **in total** 1270 late-onset AD cases **and** 7623 controls (age at
96 follow-up \geq 65 years and dementia free) **from RS-I (1118 cases, 4736 controls), RS-II (134 cases,**
97 **1928 controls) and RS-III (18 cases, 959 controls) cohorts, from follow-up conducted during**
98 **2009-2013. 10370 dementia free (Normal) participants were also included in study from all**
99 **three RS baseline cohorts and followed for an average of 11 years to analyze their progression**
100 **into incident AD.** Further, we included 360 MCI cases and 3245 **cognitively normal** controls from
101 **RS-I (235 cases, 1943 controls) and RS-II (125 cases, 1302 controls)** who were **first time** assessed
102 during 2002-2005 **in RS** (Table 1).

103 ***Genotyping***

104 Blood was drawn for genotyping from participants of RS cohort during their first visit and DNA
105 genotyping was performed at the internal genotyping facility of Erasmus Medical Center,
106 Rotterdam. All samples were genotyped with the 550K, 550K duo, or 610K Illumina arrays.
107 Genotyping quality control criteria include, call rate $< 95\%$, Hardy-Weinberg equilibrium $P <$
108 1.0×10^{-6} and Minor Allele Frequency (MAF) $< 1\%$. Moreover, study samples with excess
109 autosomal heterozygosity, call rate $< 97.5\%$, ethnic outliers and duplicate or family
110 relationships were excluded during quality control analysis. Genetic variants were imputed
111 from the Haplotype Reference Consortium (HRC) reference panel (version 1.0) [28], using the
112 Michigan imputation server [29]. The server uses SHAPEIT2 (v2.r790) [30] to phase the

113 genotype data and performs imputation with Minimac 3 software [31]. For this study we used
114 only genetic variants that had imputation quality (R-squared) > 0.5.

115 *MRI scanning*

116 *Image acquisition*

117 MRI scanning is assessed on a 1.5-T MRI unit with a dedicated eight-channel head coil (Signa HD
118 platform, GE Healthcare, Milwaukee, USA) since the induction of a dedicated MRI machine in
119 the Rotterdam Study in 2005. The MRI protocol was based on several high-resolution axial
120 sequences, including a T1-weighted (slice thickness 0.8 mm), T2-weighted (1.6 mm), and fluid
121 attenuated inversion recovery sequence (2.5 mm). A detailed description of the MRI protocol is
122 described previously [32].

123 *Image processing*

124 we excluded 251 persons with stroke and/or dementia from the total 5899 subjects who
125 came for MRI, since this may affect image processing. All T1 images were segmented into
126 supratentorial gray matter, white matter and cerebrospinal fluid using a k-nearest neighbor
127 (kNN) algorithm [33]. White matter lesions were segmented based on T1 tissue maps and
128 an automatically detected threshold for the intensity of fluid-attenuated inversion recovery
129 (FLAIR) scans [34]. After visual inspection of all segmentations, an additional 313 subjects
130 were excluded due to poor quality, leaving 5335 for the analysis. The hippocampus was
131 segmented using a fully automated method, described previously [35]. Semi-quantitative
132 MRI post-processing software were used to measure intracranial volume and brain volume
133 which included Elastix and custom-built software [36]. To calculate intracranial volume,
134 non-brain tissues (skull, eyes, dura) were removed by non-linearly registering all brain

135 scans to a manually created template in which nonbrain tissues were masked. In all scans,
136 visual checks were performed and if needed any segmentation errors manually corrected
137 [36-38]. After excluding subjects whose genotyping information was not available, we
138 ended up with 4527 cognitively normal subjects collectively from RS cohorts including RS-I
139 (968), RS-II (1074) and RS-III (2485) cohorts for our current analyses.

140 **Statistical analysis**

141 ***Genetic Risk Score computation***

142 To construct the genetic risk score, we selected late-onset AD associated single nucleotide
143 polymorphisms (SNPs) reaching genome-wide significance level ($P < 5.0 \times 10^{-8}$; Supplementary
144 Table 1), including one rare *TREM2* variant [7, 39]. In common variants, we considered only
145 variants identified by the International Genomics of Alzheimer's Project (IGAP) meta-analyses.
146 Additionally, we considered *APOE*4* (rs429358) variant for genetic risk score construction.
147 From a total of 21 SNPs, *HLA-DRB1-HLA-DRB5* (rs9271192) variant was excluded from GRS
148 calculation because of its low imputation quality (R-squared = 0.31) in RS. This led to a final
149 selection of 20 independent genome-wide significant AD associated variants. Weighted genetic
150 risk score was constructed using the effect sizes (log of OR) of the genome-wide significant
151 variants from IGAP meta-analysis [7] as weights and their respective allele dosages from
152 imputed genotype data of our study cohorts. Genetic risk score was constructed as the sum of
153 the products of SNP dosages and their corresponding weights in R software ([https://www.R-](https://www.R-project.org/)
154 [project.org/](https://www.R-project.org/)). We constructed genetic risk score in two ways; 1) Combining all 20 selected
155 variants and 2) Clustering the variants into their respective pathways.

156 **1-Combined Genetic Risk Score (GRS1)**

157 GRS1 was constructed in two ways, i.e., 1) using all the 20 selected SNPs and 2) excluding the
158 *APOE*4* variant to identify the joint independent effect of all other genome-wide significant
159 SNPs.

160 **2-Pathway-specific Genetic Risk Score (GRS2)**

161 For GRS2, the genome-wide significant AD SNPs were divided into pathways (*immune response,*
162 *endocytosis, cholesterol transport, hematopoietic cell lineage, protein ubiquitination,*
163 *hemostasis, clathrin/AP2 adaptor complex and protein folding pathway*) identified by Jones *et*
164 *al.* 2014 [16] (Supplementary Table 2). Classifying genome-wide significant AD SNPs into
165 pathways, we also utilized information from Guerreiro *et al.*, 2013 [14], in which the authors
166 reviewed the possible division of known AD associated genes into biological pathways [14].
167 Further, Gene Network database (<http://129.125.135.180:8080/GeneNetwork/>) was used to
168 confirm the allocated pathways. Of the 20 SNPs 14 could be clustered into 7 non mutually
169 exclusive pathways (Supplementary Table 2). Similar to GRS1, we also constructed GRS2 with
170 and without the *APOE*4* variant. *APOE*4* variant was grouped under four pathways including
171 *cholesterol transport [14], hematopoietic cell lineage, clathrin/AP2 adaptor complex and protein*
172 *folding [16]*. GRS2 was constructed for only those pathways, which could be assigned at least
173 two SNPs, therefore *protein ubiquitination* pathway, which contained only one SNP, was
174 excluded from all analyses, while *hematopoietic cell lineage and protein folding* pathways were
175 also not considered in the analyses excluding *APOE*4* variant.

176 **Association analyses of GRS1, GRS2**

177 To test the association of AD and MCI with **the risk scores** we used logistic regression analysis in
178 R software (www.R-project.org), using disease status as the outcome, **risk scores** as predictor
179 and age and sex as covariates. In order to assess the possible inflation of association results
180 between AD and **risk scores**, we repeated the association analysis excluding 625 AD cases which
181 were part of IGAP meta-analysis [7] from total 1270 AD cases of the RS cohort. **Further, we**
182 **performed prospective analysis using Cox-proportional hazards model (N=1057 incident AD**
183 **cases) in R software using ‘survival’ package [40] and reported results as hazard ratio (HR) per 1**
184 **standard deviation (SD) increase in risk score and 95% confidence interval.** The association of
185 single variants with AD and MCI in a logistic regression model adjusted for age and sex. Results
186 of association analyses were reported as unstandardized regression coefficient and *P* values.
187 To test the association of MRI phenotypes including total brain volume, white matter lesions
188 and hippocampal volume **with the risk scores** we used linear regression adjusted for age, sex
189 and intracranial volume at MRI scan. Single variant association analysis was also performed for
190 MRI phenotypes. **Bonferroni correction (0.05/(11 risk scores x 5 phenotypes); $P = 9.09 \times 10^{-04}$**
191 **was used to correct for multiple testing.**

192 **Results**

193 **Association of the GRS1 with AD, MCI and MRI endophenotypes**

194 The risk score containing all SNPs i.e., GRS1 both including *APOE*4* (effect = 0.73, $P = 6.53 \times 10^{-74}$)
195 and excluding *APOE*4* (effect = 0.69, $P = 1.12 \times 10^{-11}$) was significantly associated with an
196 increased risk of AD (Table 2). This association remained significant (*APOE* excluding; effect =

197 0.66, $P = 8.47 \times 10^{-7}$) after removing the patients that were included in the IGAP meta-analysis
198 [7] (Supplementary Table 3). GRS1 was also significantly associated with progression of normal
199 subjects into incident AD including (HR = 1.69, $P = 6.64 \times 10^{-83}$) and excluding APOE*4 (HR = 1.27,
200 $P = 4.88 \times 10^{-15}$; Supplementary Table 4). GRS1 was associated with MCI when APOE*4 was
201 included (effect = 0.19, $P = 0.012$) but the association was stronger with MCI when APOE*4 was
202 excluded from the analysis (effect = 0.59, $P = 9.51 \times 10^{-4}$; Table 3), but these associations did not
203 pass multiple testing correction. No association of GRS1 was observed with any of the MRI
204 phenotypes: white matter lesions, hippocampal volume, and total brain volume (Table 4).

205 Association of the GRS2 with AD

206 Among GRS2 of which APOE*4 is a part, *cholesterol transport*, *hematopoietic cell lineage*,
207 *clathrin/AP2 adaptor complex* and *protein folding* were significantly associated with AD (effect
208 ≥ 0.71 , $P < 3.22 \times 10^{-64}$) only when APOE*4 was included in the risk scores (Table 2). Among the
209 non-APOE pathways, AD was significantly associated with GRS2 capturing *immune response*
210 (effect = 0.69, $P = 3.20 \times 10^{-5}$) and *endocytosis pathway* (effect = 0.75, $P = 1.28 \times 10^{-5}$) and
211 association sustained (*Immune response*; effect = 0.68, $P = 2.22 \times 10^{-3}$ and *endocytosis*; effect =
212 0.79, $P = 5.37 \times 10^{-4}$) after removing the patients that were included in the IGAP meta-analysis
213 [7] (Supplementary Table 3). GRS2 capturing *immune response* (HR = 1.14, $P = 1.19 \times 10^{-5}$),
214 *endocytosis* (HR = 1.19, $P = 5.16 \times 10^{-8}$) and APOE*4 excluded *clathrin/AP2 adaptor complex* (HR
215 = 1.09, $P = 5.98 \times 10^{-3}$) pathway showed association with conversion risk from normal into
216 incident AD. Both *Immune response* and *endocytosis* pathways were significant after multiple
217 testing. GRS2 including APOE*4 were also significantly associated with normal to AD conversion
218 (HR ≥ 1.60 , $P \leq 1.44 \times 10^{-69}$; Supplementary Table 4). Comparatively, except for APOE*4, CR1 and

219 *BIN1*, no single variant showed significant evidence of association (Supplementary Table 5).

220 *BIN1* as a part of *endocytosis* pathways also partially explains the association of GRS2 capturing
221 *endocytosis* with AD.

222 Association of the GRS2 with MCI

223 In GRS2, only *endocytosis* pathway showed significant evidence for association (effect = 1.16, P
224 = 1.44×10^{-4} ; Table 3) with MCI and it retained significance after multiple testing. Although the
225 significance of the association is similar to that of the overall risk score (GRS1), the effect
226 estimate is considerably higher (1.16 versus 0.59 overall). In the single variant analysis, the
227 strongest association of MCI was observed with rs6733839 in the *BIN1* gene (effect = 0.262, P =
228 1.12×10^{-3} ; Supplementary Table 5) but this association lost significance after Bonferroni
229 correction. Whereas *BIN1* is part of the *endocytosis* pathway, which partially explains the
230 association between MCI and GRS2 capturing *endocytosis*.

231 Association of the GRS2 with MRI phenotypes

232 White matter lesions were associated with GRS2 capturing *immune response* (effect = 0.15, P =
233 0.016), and *clathrin/AP2 adaptor complex* excluding *APOE*4* (effect = 0.26, P = 3.55×10^{-3}). If we
234 consider multiple testing, both these associations loses significance after accounting for all
235 tested phenotypes and risk scores. Of note is that the association of the GRS2 capturing the
236 *clathrin/AP2 adaptor complex* loses its association when *APOE*4* is included in the GRS2 (effect
237 = 0.011, P = 0.507). We did not observe association of GRS2 with hippocampal volume and total
238 brain volume. In the single variant analysis association of white matter lesions is seen with
239 variants in *PICALM*, *CLU* genes ($P \leq 0.05$). Hippocampal volume shows association with variants

240 in *BIN1* and *CELF1* genes ($P < 0.05$; Supplementary Table 6). None of the single variant
241 association sustained Bonferroni correction.

242 Discussion

243 Combined GRS1 including and excluding *APOE*4* is significantly associated with AD but not with
244 MCI and MRI phenotypes in our study. Our study shows that the GRS2 capturing *immune*
245 *response* and *endocytosis* pathways are not only significantly associated with AD, normal to AD
246 conversion but also its endophenotypes, for instance, the GRS2 capturing the *endocytosis*
247 pathway also associates significantly with MCI, a group at high risk of developing AD [41, 42],
248 while the GRS2 capturing *Immune response* and *clathrin/AP2 adaptor complex* showed modest
249 association with the presence of white matter lesions at MRI in cognitively normal subjects in
250 the RS cohort (Supplementary Figure 1).

251 In our study, the association of GRS1 with AD is consistent with other similar studies on AD [43-
252 45]. GRS1 association with MCI did not pass Bonferroni correction while other studies observed
253 significant association of combined risk score with MCI [46, 47]. We did not find association of
254 GRS1 with any of the studied MRI endophenotypes. These findings are consistent with those of
255 Mormino *et al.* 2016 [48] and Lupton *et al.* 2016 [49], both studies did not find association of
256 hippocampal volume with combined GRS1 based on genome-wide significant AD variants but
257 Mormino *et al.* 2016 [48] observed this association only with risk score based on non-genome
258 wide significant AD variants. The largest study so far that included RS, however, reported
259 significant evidence of association of risk score based on all genome-wide significant AD variants
260 with hippocampal volume and total brain volume [27].

261 This is the first study that addressed the role of specific pathways in AD and its early clinical
262 manifestations i.e., MCI and MRI phenotypes. Our study shows that **GRS2** based on *immune*
263 *response* pathway was significantly associated with AD, **normal to AD conversion**. **We also**
264 **observed evidence of association of *immune response* with white matter lesions at MRI but this**
265 **did not pass Bonferroni correction, therefore should be considered carefully while interpreting**
266 **results**. These findings are converging with studies showing enrichment of *immune system*
267 *pathway* with non-genome wide significant AD variants [16, 50]. The genes clustered in *immune*
268 *response* pathway (*CLU, CRI, INPP5D, MS4A6A, TREM2, MEF2C, EPHA1*) are mainly expressed in
269 **microglial cells and play a part in the innate *immune response* in central nervous system** [51-
270 55]. Microglial cells are also thought to play a role in amyloid plaque clearance [56, 57]. It has
271 been hypothesized that activation of immune system and the subsequent inflammatory
272 response are involved in neuronal damage including axonal loss and white matter pathology
273 due to demyelination [58]. This may explain the association of the AD genes involved in the
274 *immune response* with white matter lesions that we observed in the present study [59]. White
275 matter lesions are associated with increased risk of cognitive decline, developing dementia [21]
276 and AD [60, 61]. White matter lesions are also more frequently observed in AD patients than
277 controls [62, 63].

278 The present study reveals further that the genes capturing the *endocytosis pathway* not only
279 strongly associate to AD but also to MCI. **We also showed that *endocytosis pathway* is also**
280 **associated with progression from normal (dementia free) to AD in average 11 years of follow-**
281 **up**. This pathway is independent of *APOE* and includes the *BIN1, PICALM, CD2AP, SORL1* genes.
282 We show that the association of GRS1 with MCI status is mainly attributed to the genes

283 involved in the *endocytosis* pathway. Omitting the AD genes not related to the *endocytosis*
284 pathway makes the association of the pathway with MCI even stronger. The association
285 suggests that the *endocytosis* pathway plays a critical role in an early prodromal phase of AD
286 and converges with studies suggesting activation of *endocytic pathway* is the earliest reported
287 intracellular manifestation of AD [41, 42, 64]. Based on the effect estimates of association of
288 *endocytosis pathway* with AD and MCI (0.79 vs 1.59), we also speculated that *endocytosis* is
289 more strongly associated to MCI than to AD. MCI is considered a prodromal stage in AD
290 patients that suggest *endocytosis* pathway is associated with early pathology of AD. The
291 *endocytosis pathway* is involved in neuronal uptake of macromolecules and secretory vesicles
292 during synaptic transmission. As efficient uptake of extracellular cholesterol is critical for
293 neuronal functions such as repair, synapse formation and exon elongation [65], normal
294 neuronal work needs smooth functioning of *endocytosis* pathways [66]. Post-mortem studies
295 have also demonstrated reduced brain cholesterol levels in the brain areas responsible for
296 memory and learning, among late-onset AD cases and age matched controls [67]. These facts
297 suggest that defects in *endocytosis* which derive the cholesterol uptake could lead to impaired
298 neurotransmitter release and synaptic function [68]. Dysfunction in *endocytosis* can also
299 contribute to accumulation of abnormal A β peptide [69]. Based on this finding, we can suggest
300 that *endocytosis* pathway is a common molecular mechanism between MCI and AD that starts
301 manifesting at early stages of disease. Risk contributed by variants clustered in this pathway at
302 various stages of AD progression can possibly provide clue about disease trajectory.

303 Similar to the *immune response* pathway, the *clathrin/AP2 adaptor complex* pathway modestly
304 associated with white matter lesions. Although, the association failed to pass the multiple

305 testing, but combined AD risk score did not capture the association with white matter lesions in
306 our and even in large studies [27]. Capturing this association in small sample indicates the
307 importance of this pathway in explaining the white matter lesions pathology. Two variants
308 tagging *PICALM* and *CLU* genes cluster in the *clathrin/AP2 adaptor complex* pathway. Each
309 variant independently shows nominally significant association with white matter lesions in our
310 analyses but combining their effect are additive and improve the strength of association. There
311 is a strong evidence that the two protein encoded by the genes interact at molecular level [70,
312 71]. *PICALM* is involved in VAMP2 trafficking that is a crucial process to maintain functional
313 integrity of synapses which are crucial to cognitive function [72, 73]. *PICALM* is also found to be
314 expressed in the white matter and, immune-labeling of human brain tissue shows that *PICALM*
315 is mainly found in blood vessel walls [74]. *CLU* clustered in *clathrin/AP2 adaptor* is involved in
316 efflux of free insoluble amyloid beta (A β) peptide through blood brain barrier [75]. Increased
317 plasma levels of *CLU* were found to be associated with increased burden of A β peptide in
318 healthy elderly population and brain atrophy in AD [76, 77] and decreased integrity of white
319 matter in young adults [78]. Demyelination of white matter is reported to occur even before
320 the accumulation of A β plaques and neurofibrillary tangles [79]. The findings of the present
321 study suggest that the increased genetic burden of risk variants in *clathrin/AP2 adaptor*
322 complex (*clathrin mediated endocytosis*) and *immune response* pathway may play a role in early
323 pathogenesis of AD through white matter pathology.

324 Among pathways including *APOE (Cholesterol transport, hematopoietic cell lineage,*
325 *clathrin/AP2 adaptor complex, protein folding)*, significant association with AD and normal to
326 AD conversion suggests that *APOE*4* appear to be the driving genetic factor for these

327 associations. Only *the clathrin/AP2 adaptor complex* shows evidence of association ($P = 0.036$)
328 to AD and normal to AD conversion ($P = 5.98 \times 10^{-3}$) when *APOE*4* variant is excluded from the
329 analysis but did not pass multiple testing correction.

330 Our study provides a readout of pathway based **risk score** association with AD and its pre-
331 dementia endophenotypes. Main clinical significance of our findings is that they will allow to
332 determine whether a certain biological pathway is involved in an individual patient. This will
333 permit targeted interventions based on predicted pathological pathways. Similar as the case of
334 cardiovascular diseases [80], a heterogeneous disease treatment can be followed based on
335 pathway biomarkers (e.g., glucose level, total cholesterol and high density lipid levels, liver
336 enzymes in case of cardiovascular disease) [81] but rather on genetic basis. This require
337 reference pathways and treatment portfolio. In the meantime, the pathway based genetic risk
338 score will allow stratification of the high risk patients in clinical trials based on causal pathways
339 involved in patients. This may improve both the power and efficiency of future clinical and
340 preventive trials.

341 Our study is a step forward to use known genetic and pathway information for disentangling
342 the mechanisms of AD but it has one major limitation that pathway information is based on
343 known AD variants identified so far. This will further improve in future with improved genetic
344 risk information that can better capture the underlying pathways. Another possible limitation of
345 our study is that 625 cases of RS-I was a part of meta-analysis performed by IGAP [7] which can
346 contribute to possible inflation in our results of association of **risk score** with AD. However,
347 excluding these patients, the results of this study largely remained unchanged.

348 In conclusion we found different pathways are implicated in different endophenotypes of AD.
349 *Endocytosis* pathway is involved in MCI and AD, while *immune response* associates with AD.
350 Further, *Immune response* and *clathrin/AP2 adaptor complex* pathways are involved in white
351 matter lesions but their association should be carefully considered due to multiple testing
352 limitation. Interestingly, all the observed associations with early AD pathology are shown by
353 *APOE* excluding pathways. Future findings from genomic research will improve the quality of
354 the pathway-specific genetic scores.

355

356 **Acknowledgements**

357 This study was funded by the PERADES Program (Defining Genetic, Polygenic and
358 Environmental Risk for Alzheimer’s disease, using multiple powerful cohorts, focused
359 Epigenetics and Stem cell metabolomics), Project number 733051021. The Rotterdam Study is
360 funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands
361 Organization for the Health Research and Development (ZonMw), the Research Institute for
362 Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for
363 Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of
364 Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam
365 Study and the participating general practitioners and pharmacists. The generation and
366 management of GWAS genotype data for the Rotterdam Study (RS-I, RS-II, RS-III) was executed
367 by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal
368 Medicine, Erasmus MC, Rotterdam, The Netherlands. The GWAS datasets are supported by the
369 Netherlands Organization of Scientific Research NWO Investments (Project number
370 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal
371 Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2),
372 the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research
373 (NWO) Netherlands Consortium for Healthy Aging (NCHA), project number 050-060-810. We
374 thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters, MSc, and
375 Carolina Medina-Gomez, MSc, for their help in creating the GWAS database, and Karol Estrada,
376 PhD, Yurii Aulchenko, PhD, and Carolina Medina-Gomez, MSc, for the creation and analysis of
377 imputed data.

378 **References**

- 379 [1] Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, et al. Role of genes
380 and environments for explaining Alzheimer disease. *Arch Gen Psychiatry*. 2006;63:168-74.
- 381 [2] Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al.
382 Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset
383 families. *Science*. 1993;261:921-3.
- 384 [3] Bu G. Apolipoprotein E and its receptors in Alzheimer's disease: pathways, pathogenesis and
385 therapy. *Nat Rev Neurosci*. 2009;10:333-44.
- 386 [4] Huang Y, Mucke L. Alzheimer mechanisms and therapeutic strategies. *Cell*. 2012;148:1204-
387 22.
- 388 [5] Kopke E, Tung YC, Shaikh S, Alonso AC, Iqbal K, Grundke-Iqbal I. Microtubule-associated
389 protein tau. Abnormal phosphorylation of a non-paired helical filament pool in Alzheimer
390 disease. *J Biol Chem*. 1993;268:24374-84.
- 391 [6] Hardy JA, Higgins GA. Alzheimer's disease: the amyloid cascade hypothesis. *Science*.
392 1992;256:184-5.
- 393 [7] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-
394 analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat*
395 *Genet*. 2013;45:1452-8.
- 396 [8] Genin E, Hannequin D, Wallon D, Sleegers K, Hiltunen M, Combarros O, et al. APOE and
397 Alzheimer disease: a major gene with semi-dominant inheritance. *Mol Psychiatry*. 2011;16:903-
398 7.
- 399 [9] Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, et al. Genome-wide
400 association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat*
401 *Genet*. 2009;41:1094-9.
- 402 [10] Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, et al.
403 Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA*.
404 2010;303:1832-40.
- 405 [11] Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM, et al.
406 Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated
407 with Alzheimer's disease. *Nat Genet*. 2011;43:429-35.
- 408 [12] Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buross J, et al. Common variants
409 at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's
410 disease. *Nat Genet*. 2011;43:436-41.
- 411 [13] Desikan RS, Schork AJ, Wang Y, Thompson WK, Dehghan A, Ridker PM, et al. Polygenic
412 Overlap Between C-Reactive Protein, Plasma Lipids, and Alzheimer Disease. *Circulation*.
413 2015;131:2061-9.
- 414 [14] Guerreiro R, Bras J, Hardy J. SnapShot: genetics of Alzheimer's disease. *Cell*.
415 2013;155:968- e1.
- 416 [15] Medway C, Morgan K. Review: The genetics of Alzheimer's disease; putting flesh on the
417 bones. *Neuropathol Appl Neurobiol*. 2014;40:97-105.
- 418 [16] International Genomics of Alzheimer's Disease C. Convergent genetic and expression data
419 implicate immunity in Alzheimer's disease. *Alzheimers Dement*. 2015;11:658-71.
- 420 [17] Van der Flier WM. Clinical heterogeneity in familial Alzheimer's disease. *Lancet Neurol*.
421 2016;15:1296-8.

422 [18] Becker JT, Huff FJ, Nebes RD, Holland A, Boller F. Neuropsychological function in
423 Alzheimer's disease. Pattern of impairment and rates of progression. *Arch Neurol.* 1988;45:263-
424 8.

425 [19] Mann UM, Mohr E, Gearing M, Chase TN. Heterogeneity in Alzheimer's disease:
426 progression rate segregated by distinct neuropsychological and cerebral metabolic profiles. *J*
427 *Neurol Neurosurg Psychiatry.* 1992;55:956-9.

428 [20] Dubois B, Hampel H, Feldman HH, Scheltens P, Aisen P, Andrieu S, et al. Preclinical
429 Alzheimer's disease: Definition, natural history, and diagnostic criteria. *Alzheimers Dement.*
430 2016;12:292-323.

431 [21] Debette S, Beiser A, DeCarli C, Au R, Himali JJ, Kelly-Hayes M, et al. Association of MRI
432 markers of vascular brain injury with incident stroke, mild cognitive impairment, dementia, and
433 mortality: the Framingham Offspring Study. *Stroke.* 2010;41:600-6.

434 [22] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive
435 impairment: clinical characterization and outcome. *Arch Neurol.* 1999;56:303-8.

436 [23] Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, et al. Mild cognitive
437 impairment. *Lancet.* 2006;367:1262-70.

438 [24] Petersen RC. Mild cognitive impairment as a diagnostic entity. *Journal of internal medicine.*
439 2004;256:183-94.

440 [25] Hofman A, Brusselle GG, Darwish Murad S, van Duijn CM, Franco OH, Goedegebure A, et
441 al. The Rotterdam Study: 2016 objectives and design update. *Eur J Epidemiol.* 2015;30:661-708.

442 [26] Ott A, Breteler MM, van Harskamp F, Stijnen T, Hofman A. Incidence and risk of
443 dementia. The Rotterdam Study. *Am J Epidemiol.* 1998;147:574-80.

444 [27] Chauhan G, Adams HH, Bis JC, Weinstein G, Yu L, Toghiofer AM, et al. Association of
445 Alzheimer's disease GWAS loci with MRI markers of brain aging. *Neurobiol Aging.*
446 2015;36:1765 e7-16.

447 [28] McCarthy S, Das S, Kretschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference
448 panel of 64,976 haplotypes for genotype imputation. *Nat Genet.* 2016;48:1279-83.

449 [29] Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation
450 genotype imputation service and methods. *Nat Genet.* 2016;48:1284-7.

451 [30] Delaneau O, Marchini J, Genomes Project C, Genomes Project C. Integrating sequence and
452 array data to create an improved 1000 Genomes Project haplotype reference panel. *Nat*
453 *Commun.* 2014;5:3934.

454 [31] Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate
455 genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet.*
456 2012;44:955-9.

457 [32] Ikram MA, van der Lugt A, Niessen WJ, Koudstaal PJ, Krestin GP, Hofman A, et al. The
458 Rotterdam Scan Study: design update 2016 and main findings. *European journal of*
459 *epidemiology.* 2015;30:1299-315.

460 [33] Vrooman HA, Cocosco CA, van der Lijn F, Stokking R, Ikram MA, Vernooij MW, et al.
461 Multi-spectral brain tissue segmentation using automatically trained k-Nearest-Neighbor
462 classification. *Neuroimage.* 2007;37:71-81.

463 [34] De Boer R, Vrooman HA, Van Der Lijn F, Vernooij MW, Ikram MA, Van Der Lugt A, et
464 al. White matter lesion extension to automatic brain tissue segmentation on MRI. *Neuroimage.*
465 2009;45:1151-61.

466 [35] Fischl B. FreeSurfer. *Neuroimage.* 2012;62:774-81.

467 [36] Klein S, Staring M, Pluim JPW. Comparison of gradient approximation techniques for
468 optimisation of mutual information in nonrigid registration. 2005. p. 192-203.

469 [37] Vrooman HA, Cocosco CA, van der Lijn F, Stokking R, Ikram MA, Vernooij MW, et al.
470 Multi-spectral brain tissue segmentation using automatically trained k-Nearest-Neighbor
471 classification. *Neuroimage*. 2007;37:71-81.

472 [38] Ikram MA, Fornage M, Smith AV, Seshadri S, Schmidt R, Debette S, et al. Common
473 variants at 6q22 and 17q21 are associated with intracranial volume. *Nat Genet*. 2012;44:539-44.

474 [39] Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaevea E, Majounie E, et al. TREM2
475 variants in Alzheimer's disease. *N Engl J Med*. 2013;368:117-27.

476 [40] Therneau T. A Package for Survival Analysis in S. R package version 2.37-7. 2014. URL
477 <http://cran.r-project.org/package=survival>. 2015.

478 [41] Funderburk SF, Marcellino BK, Yue Z. Cell "self-eating" (autophagy) mechanism in
479 Alzheimer's disease. *Mt Sinai J Med*. 2010;77:59-68.

480 [42] Ginsberg SD, Mufson EJ, Alldred MJ, Counts SE, Wu J, Nixon RA, et al. Upregulation of
481 select rab GTPases in cholinergic basal forebrain neurons in mild cognitive impairment and
482 Alzheimer's disease. *J Chem Neuroanat*. 2011;42:102-10.

483 [43] Chouraki V, Reitz C, Maury F, Bis JC, Bellenguez C, Yu L, et al. Evaluation of a Genetic
484 Risk Score to Improve Risk Prediction for Alzheimer's Disease. *J Alzheimers Dis*. 2016;53:921-
485 32.

486 [44] Escott-Price V, Shoai M, Pither R, Williams J, Hardy J. Polygenic score prediction captures
487 nearly all common genetic risk for Alzheimer's disease. *Neurobiol Aging*. 2016.

488 [45] Slegers K, Bettens K, De Roeck A, Van Cauwenberghe C, Cuyvers E, Verheijen J, et al. A
489 22-single nucleotide polymorphism Alzheimer's disease risk score correlates with family history,
490 onset age, and cerebrospinal fluid Abeta42. *Alzheimers Dement*. 2015;11:1452-60.

491 [46] Adams HH, de Bruijn RF, Hofman A, Uitterlinden AG, van Duijn CM, Vernooij MW, et al.
492 Genetic risk of neurodegenerative diseases is associated with mild cognitive impairment and
493 conversion to dementia. *Alzheimers Dement*. 2015;11:1277-85.

494 [47] Lacour A, Espinosa A, Louwersheimer E, Heilmann S, Hernandez I, Wolfsgruber S, et al.
495 Genome-wide significant risk factors for Alzheimer's disease: role in progression to dementia
496 due to Alzheimer's disease among subjects with mild cognitive impairment. *Mol Psychiatry*.
497 2017;22:153-60.

498 [48] Mormino EC, Sperling RA, Holmes AJ, Buckner RL, De Jager PL, Smoller JW, et al.
499 Polygenic risk of Alzheimer disease is associated with early- and late-life processes. *Neurology*.
500 2016;87:481-8.

501 [49] Lupton MK, Strike L, Hansell NK, Wen W, Mather KA, Armstrong NJ, et al. The effect of
502 increased genetic risk for Alzheimer's disease on hippocampal and amygdala volume. *Neurobiol*
503 *Aging*. 2016;40:68-77.

504 [50] Lambert JC, Grenier-Boley B, Chouraki V, Heath S, Zelenika D, Fievet N, et al. Implication
505 of the immune system in Alzheimer's disease: evidence from genome-wide pathway analysis. *J*
506 *Alzheimers Dis*. 2010;20:1107-18.

507 [51] Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keefe S, et al. An RNA-
508 sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the
509 cerebral cortex. *J Neurosci*. 2014;34:11929-47.

510 [52] Kofler J, Bissel S, Wiley C, Stauffer M, Murdoch G. Differential microglial expression of
511 new Alzheimer's disease associated genes MS4A4A and MS4A6A. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*.8:P253.

512

513 [53] Colonna M. TREMs in the immune system and beyond. *Nat Rev Immunol.* 2003;3:445-53.
514 [54] Crehan H, Hardy J, Pocock J. Blockage of CR1 prevents activation of rodent microglia.
515 *Neurobiol Dis.* 2013;54:139-49.
516 [55] Zhang B, Gaiteri C, Bodea LG, Wang Z, McElwee J, Podtelezchnikov AA, et al. Integrated
517 systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell.*
518 2013;153:707-20.
519 [56] Lee CY, Landreth GE. The role of microglia in amyloid clearance from the AD brain. *J*
520 *Neural Transm (Vienna).* 2010;117:949-60.
521 [57] Heneka MT, Carson MJ, El Khoury J, Landreth GE, Brosseron F, Feinstein DL, et al.
522 Neuroinflammation in Alzheimer's disease. *Lancet Neurol.* 2015;14:388-405.
523 [58] Heppner FL, Ransohoff RM, Becher B. Immune attack: the role of inflammation in
524 Alzheimer disease. *Nat Rev Neurosci.* 2015;16:358-72.
525 [59] Gouw AA, Seewann A, van der Flier WM, Barkhof F, Rozemuller AM, Scheltens P, et al.
526 Heterogeneity of small vessel disease: a systematic review of MRI and histopathology
527 correlations. *J Neurol Neurosurg Psychiatry.* 2011;82:126-35.
528 [60] Debette S, Markus HS. The clinical importance of white matter hyperintensities on brain
529 magnetic resonance imaging: systematic review and meta-analysis. *BMJ.* 2010;341:c3666.
530 [61] Jeong H, Son S, Kim S-K, Park K-J, Choi N-C, Kwon O-Y, et al. Age, hypertension, and
531 genetic polymorphisms and their relative associations with white matter hyperintensities in
532 Korean patients with Alzheimer's disease. *Neurology Asia.* 2015;20:35-41.
533 [62] Smith EE, Egorova S, Blacker D, Killiany RJ, Muzikansky A, Dickerson BC, et al.
534 Magnetic resonance imaging white matter hyperintensities and brain volume in the prediction of
535 mild cognitive impairment and dementia. *Arch Neurol.* 2008;65:94-100.
536 [63] Wolf H, Ecke GM, Bettin S, Dietrich J, Gertz HJ. Do white matter changes contribute to the
537 subsequent development of dementia in patients with mild cognitive impairment? A longitudinal
538 study. *Int J Geriatr Psychiatry.* 2000;15:803-12.
539 [64] Cataldo AM, Peterhoff CM, Troncoso JC, Gomez-Isla T, Hyman BT, Nixon RA. Endocytic
540 pathway abnormalities precede amyloid beta deposition in sporadic Alzheimer's disease and
541 Down syndrome: differential effects of APOE genotype and presenilin mutations. *Am J Pathol.*
542 2000;157:277-86.
543 [65] Vance JE, Campenot RB, Vance DE. The synthesis and transport of lipids for axonal growth
544 and nerve regeneration. *Biochim Biophys Acta.* 2000;1486:84-96.
545 [66] Keating DJ, Chen C, Pritchard MA. Alzheimer's disease and endocytic dysfunction: clues
546 from the Down syndrome-related proteins, DSCR1 and ITSN1. *Ageing Res Rev.* 2006;5:388-
547 401.
548 [67] Svennerholm L, Gottfries CG. Membrane lipids, selectively diminished in Alzheimer
549 brains, suggest synapse loss as a primary event in early-onset form (type I) and demyelination in
550 late-onset form (type II). *J Neurochem.* 1994;62:1039-47.
551 [68] Woodruff G, Reyna SM, Dunlap M, Van Der Kant R, Callender JA, Young JE, et al.
552 Defective Transcytosis of APP and Lipoproteins in Human iPSC-Derived Neurons with Familial
553 Alzheimer's Disease Mutations. *Cell Reports.* 2016;17:759-73.
554 [69] Maxfield FR. Role of endosomes and lysosomes in human disease. *Cold Spring Harb*
555 *Perspect Biol.* 2014;6:a016931.
556 [70] Yang X, Li J, Liu B, Li Y, Jiang T. Impact of PICALM and CLU on hippocampal
557 degeneration. *Hum Brain Mapp.* 2016;37:2419-30.

558 [71] Tan L, Wang HF, Tan MS, Tan CC, Zhu XC, Miao D, et al. Effect of CLU genetic variants
559 on cerebrospinal fluid and neuroimaging markers in healthy, mild cognitive impairment and
560 Alzheimer's disease cohorts. *Sci Rep.* 2016;6:26027.

561 [72] Harel A, Wu F, Mattson MP, Morris CM, Yao PJ. Evidence for CALM in directing VAMP2
562 trafficking. *Traffic.* 2008;9:417-29.

563 [73] Yao PJ. Synaptic frailty and clathrin-mediated synaptic vesicle trafficking in Alzheimer's
564 disease. *Trends in neurosciences.* 2004;27:24-9.

565 [74] Baig S, Joseph SA, Tayler H, Abraham R, Owen MJ, Williams J, et al. Distribution and
566 expression of picalm in Alzheimer disease. *J Neuropathol Exp Neurol.* 2010;69:1071-7.

567 [75] Zlokovic BV. Neurovascular pathways to neurodegeneration in Alzheimer's disease and
568 other disorders. *Nat Rev Neurosci.* 2011;12:723-38.

569 [76] Thambisetty M, Simmons A, Velayudhan L, Hye A, Campbell J, Zhang Y, et al.
570 Association of plasma clusterin concentration with severity, pathology, and progression in
571 Alzheimer disease. *Arch Gen Psychiatry.* 2010;67:739-48.

572 [77] Weinstein G, Beiser AS, Preis SR, Courchesne P, Chouraki V, Levy D, et al. Plasma
573 clusterin levels and risk of dementia, Alzheimer's disease, and stroke. *Alzheimers Dement*
574 *(Amst).* 2016;3:103-9.

575 [78] Braskie MN, Jahanshad N, Stein JL, Barysheva M, McMahon KL, de Zubicaray GI, et al.
576 Common Alzheimer's disease risk variant within the CLU gene affects white matter
577 microstructure in young adults. *J Neurosci.* 2011;31:6764-70.

578 [79] Sachdev PS, Zhuang L, Braidy N, Wen W. Is Alzheimer's a disease of the white matter?
579 *Curr Opin Psychiatry.* 2013;26:244-51.

580 [80] Sun X, Jia Z. A brief review of biomarkers for preventing and treating cardiovascular
581 diseases. *J Cardiovasc Dis Res.* 2012;3:251-4.

582 [81] Upadhyay RK. Emerging risk biomarkers in cardiovascular diseases and disorders. *J Lipids.*
583 2015;2015:971453.

584

585

586

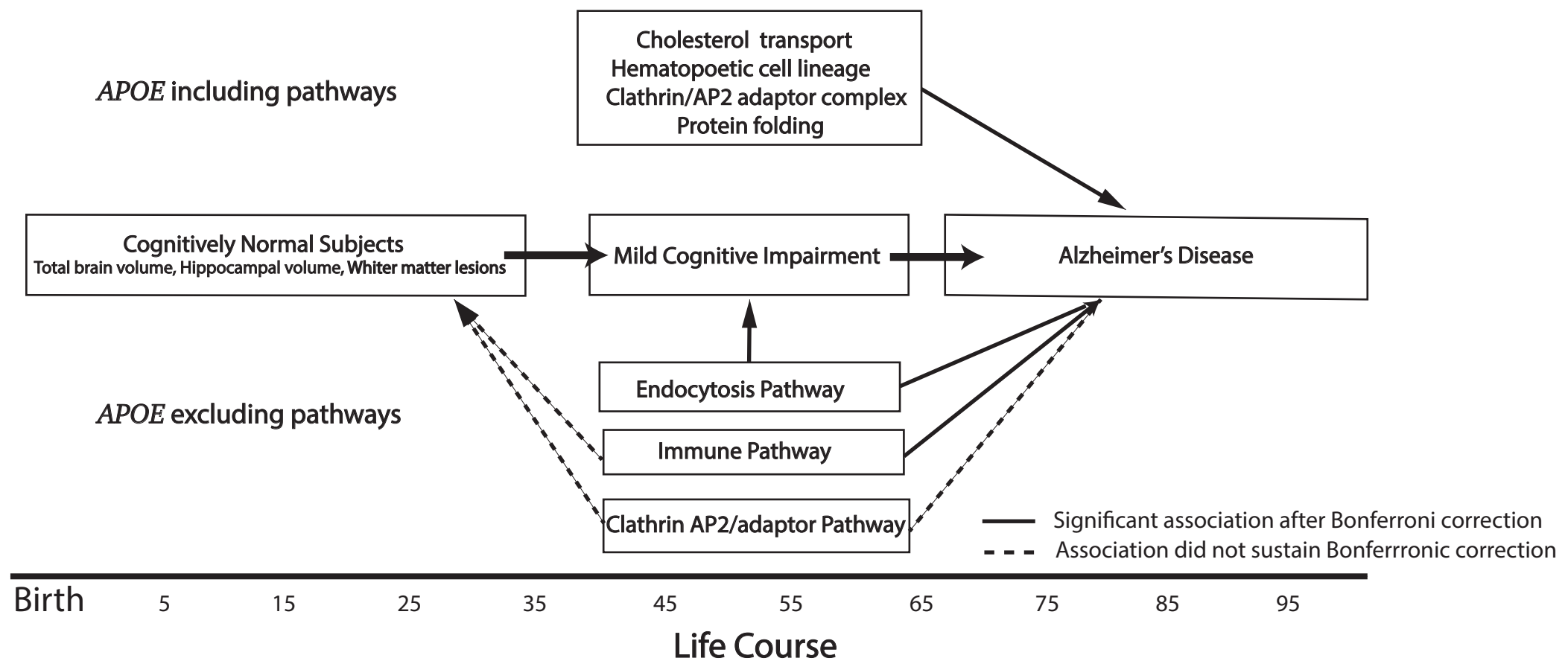
587

588

589 **Figure caption:**

590 **Supplementary Figure 1:** Diagram showing the association results of AD, MCI and WML in
591 cognitively normal subjects with pathways-specific GRS2 both including and excluding
592 *APOE*4*. Bonferroni correction threshold ($P = 9.09 \times 10^{-4}$)

593



Supplementary Material

Supplementary Table 1: List of Genome-wide significant variants associated with AD

Chromosome	Position (BP)	SNP	Gene	Coding Allele	Effect	MAF	RSI_Rsq	RSII_Rsq	RSIII_Rsq	Effect origin
1	207692049	rs6656401	<i>CR1</i>	G	-0.157	0.197	0.953	0.948	0.950	Lambert et al. (2013)[1]
2	127892810	rs6733839	<i>BIN1</i>	T	0.188	0.409	0.960	0.911	0.962	Lambert et al. (2013)
2	234068476	rs35349669	<i>INPP5D</i>	T	0.066	0.488	0.975	0.973	0.976	Lambert et al. (2013)
5	88223420	rs190982	<i>MEF2C</i>	A	0.08	0.408	0.979	0.934	0.978	Lambert et al. (2013)
6	41129252	rs75932628	<i>TREM2</i>	T	0.889	0.0016	0.762	0.726	0.668	Ruiz et al. (2014)[2]
6	32578530	rs9271192	<i>HLA-DRB5-HLA-DRB1[†]</i>	A	-0.108	0.276	0.314	0.312	0.314	Lambert et al. (2013)
6	47487762	rs10948363	<i>CD2AP</i>	G	0.098	0.266	0.998	0.998	0.998	Lambert et al. (2013)
7	100004446	rs1476679	<i>ZCWPW1</i>	T	0.078	0.287	0.995	0.996	0.995	Lambert et al. (2013)
7	143110762	rs11771145	<i>EPHA1</i>	A	-0.102	0.338	0.998	0.998	0.999	Lambert et al. (2013)
7	37841534	rs2718058	<i>NME8</i>	G	-0.07	0.373	1.000	1.000	1.000	Lambert et al. (2013)
8	27195121	rs28834970	<i>PTK2B</i>	C	0.096	0.366	0.993	0.990	0.994	Lambert et al. (2013)
8	27467686	rs9331896	<i>CLU</i>	T	0.146	0.379	0.902	0.974	0.901	Lambert et al. (2013)
11	121435587	rs11218343	<i>SORL1</i>	C	-0.27	0.039	0.998	0.995	0.998	Lambert et al. (2013)
11	47557871	rs10838725	<i>CELF1</i>	C	0.075	0.316	0.998	0.998	0.998	Lambert et al. (2013)
11	59923508	rs983392	<i>MS4A6A</i>	G	-0.108	0.403	0.989	0.990	0.991	Lambert et al. (2013)
11	85867875	rs10792832	<i>PICALM</i>	G	0.13	0.358	0.999	0.999	0.999	Lambert et al. (2013)
14	53400629	rs17125944	<i>FERMT2</i>	C	0.122	0.092	1.000	1.000	1.000	Lambert et al. (2013)
14	92926952	rs10498633	<i>SLC24A4-RIN3</i>	T	-0.104	0.217	0.999	0.999	1.000	Lambert et al. (2013)
19	45411941	rs429358	<i>APOE*4</i>	C	1.3503	0.148	0.949	0.944	0.947	Lambert et al. (2013)
19	1063443	rs4147929	<i>ABCA7</i>	G	-0.135	0.19	0.916	0.917	0.991	Lambert et al. (2013)
20	55018260	rs7274581	<i>CASS4</i>	C	-0.139	0.083	0.990	0.989	0.990	Lambert et al. (2013)

Abbreviations: AD ~ Alzheimer's disease, MAF ~ minor allele frequency, **Rsq = R-squared (HRC imputation quality)**

*Effect (log of odds ratio) is based on Coding Allele column

† Variants have low imputation quality in HRC imputation in our cohorts therefore excluded from genetic risk score calculation .

Supplementary Table 2: Clustering of genome-wide significant variants into their respective 8 biological pathways.

Pathway	Gene*	Assigned SNP	Genes reported in pathway		Constructed GRS(Yes/No)	
			Jones <i>et al</i>	Guerreiro <i>et al</i>	Including APOE	Excluding APOE
Immune Response	<i>CLU</i>	rs9331896	Yes	Yes		
	<i>CR1</i>	rs6656401	Yes	Yes		
	<i>INPP5D</i>	rs35349669	Yes	Yes		
	<i>EPHA1</i>	rs11771145	-	Yes	No	Yes
	<i>MS4A6A</i>	rs983392	-	Yes		
	<i>TREM2</i>	rs75932628	-	Yes		
	<i>MEF2C</i>	rs190982	-	Yes		
Endocytosis	<i>CD2AP</i>	rs10948363		Yes		
	<i>PICALM</i>	rs10792832	Yes	Yes	No	Yes
	<i>BIN1</i>	rs6733839	Yes	Yes		
	<i>SORL1</i>	rs11218343	-	Yes		
Cholesterol transport†	<i>CLU</i>	rs9331896	Yes	Yes		
	<i>ABCA7</i>	rs4147929	Yes	Yes		
	<i>SORL1</i>	rs11218343	-	Yes	Yes	Yes
	<i>APOE*4</i>	rs429358	Yes	Yes		
Hematopoietic cell lineage†	<i>CR1</i>	rs6656401	Yes	-	Yes	No
	<i>APOE*4</i>	rs429358	Yes	-		
Protein ubiquitination	<i>CLU</i>	rs9331896	Yes	-	No	No
Hemostasis	<i>CLU</i>	rs9331896	Yes	-	No	Yes
	<i>INPP5D</i>	rs35349669	Yes	-		
Clathrin/AP2 adaptor complex†	<i>CLU</i>	rs9331896	Yes	-		
	<i>PICALM</i>	rs10792832	Yes	-	Yes	Yes
	<i>APOE*4</i>	rs429358	Yes	-		
Protein folding†	<i>CLU</i>	rs9331896	Yes	-	Yes	No
	<i>APOE*4</i>	rs429358	Yes	-		

*SNPs classification into pathways is based on the information from Jones *et al.* [3] and Guerreiro *et al*[4].

†APOE*4 variant (rs429358) is grouped under these pathways. *Protein folding* and *Hematopoietic cell lineage* pathways are left with one SNP after excluding APOE*4 variant therefore were not considered for APOE excluding analysis.

Supplementary Table 3: Results of association of AD with risk scores

SNP Cluster*	Including <i>APOE</i>			Excluding <i>APOE</i>		
	β	SE	P-value	β	SE	P-value
GRS1 (Combined)	0.72	0.052	1.80×10^{-43}	0.66	0.135	8.45×10^{-7}
Immune response	-	-	-	0.68	0.221	2.22×10^{-3}
Endocytosis	-	-	-	0.79	0.228	5.37×10^{-4}
Cholesterol Transport	0.70	0.054	3.48×10^{-38}	0.41	0.294	0.159
Hematopoietic cell lineage[†]	0.72	0.055	7.19×10^{-40}	-	-	-
Hemostasis	-	-	-	0.35	0.390	0.364
Clathrin/AP2 Adaptor complex	0.70	0.055	6.47×10^{-38}	0.27	0.314	0.383
Protein folding[†]	0.71	0.055	5.18×10^{-38}	-	-	-

Abbreviations: GRS1 ~ Combined genetic risk score, SNP ~ single nucleotide polymorphism, β ~ regression coefficient, SE ~ Standard error.

*Logistic regression model adjusted for age and sex in RS (N=645) excluding 625 cases included in IGAP meta-analysis

[†]Only one SNP available in excluding *APOE* GRS2

Supplementary Table 4: Results of longitudinal analysis from normal (dementia free) to AD conversion

SNP Cluster*	Including <i>APOE</i>			Excluding <i>APOE</i>		
	HR	95% CI	P-value	HR	95% CI	P-value
GRS1 (Combined)	1.69	1.61-1.79	6.64×10^{-83}	1.27	1.19-1.34	4.88×10^{-15}
Immune response	-	-	-	1.14	1.07-1.21	1.19×10^{-5}
Endocytosis	-	-	-	1.19	1.12-1.26	5.16×10^{-8}
Cholesterol Transport	1.60	1.52-1.68	1.44×10^{-69}	1.07	1.01-1.14	2.45×10^{-2}
Hematopoietic cell lineage[†]	1.60	1.52-1.68	4.29×10^{-71}	-	-	-
Hemostasis	-	-	-	1.08	1.01-1.14	1.62×10^{-2}
Clathrin/AP2 Adaptor complex	1.61	1.52-1.69	4.21×10^{-71}	1.09	1.02-1.15	5.98×10^{-3}
Protein folding[†]	1.60	1.52-1.68	7.66×10^{-70}	-	-	-

Note: Multiple testing correction by Bonferroni $0.05 / (5 \text{ phenotypes} \times 11 \text{ risk scores})$; $P < 9.09 \times 10^{-4}$ was considered significant

Abbreviations: GRS1 ~ Combined genetic risk score, SNP ~ single nucleotide polymorphism,

HR ~ Hazard ratio per 1 standard deviation of risk score, CI ~ Confidence interval

*Cox proportional hazards model adjusted for age at baseline and sex in RS (N = 10370 normal at baseline)

[†]Only one SNP available in excluding *APOE* GRS2

Supplementary Table 5: Results of single variant association with AD and MCI

Phenotype →		Alzheimer's Disease			Mild Cognitive Impairment		
SNP	Gene	β	SE	P-value	β	SE	P-value
rs429358	<i>APOE*4</i>	0.949	0.057	8.78x10 ⁻⁶³	0.137	0.114	0.229
rs75932628	<i>TREM2</i>	0.816	0.470	0.083	1.293	0.639	0.043
rs6656401	<i>CR1</i>	-0.180	0.055	1.11x10 ⁻³	0.072	0.105	0.496
rs11218343	<i>SORL1</i>	-0.161	0.110	0.142	-0.193	0.206	0.351
rs10838725	<i>CELF1</i>	0.062	0.048	0.192	0.069	0.086	0.424
rs983392	<i>MS4A6A</i>	-0.081	0.045	0.072	-0.127	0.081	0.115
rs10792832	<i>PICALM</i>	0.085	0.045	0.063	0.181	0.083	0.028
rs17125944	<i>FERMT2</i>	0.128	0.070	0.067	-0.016	0.131	0.901
rs10498633	<i>SLC24A4-RIN3</i>	-0.092	0.053	0.081	0.049	0.094	0.601
rs4147929	<i>ABCA7</i>	-0.003	0.061	0.961	-0.064	0.107	0.552
rs6733839	<i>BIN1</i>	0.149	0.045	9.74x10 ⁻⁴	0.262	0.081	1.12x10 ⁻³
rs35349669	<i>INPP5D</i>	0.055	0.044	0.216	-0.081	0.080	0.315
rs7274581	<i>CASS4</i>	-0.120	0.083	0.149	-0.034	0.143	0.810
rs190982	<i>MEF2C</i>	0.010	0.046	0.833	0.061	0.083	0.458
rs10948363	<i>CD2AP</i>	0.079	0.049	0.102	0.030	0.088	0.732
rs1476679	<i>ZCWPW1</i>	0.005	0.047	0.918	-0.037	0.085	0.662
rs11771145	<i>EPHA1</i>	-0.075	0.047	0.107	-0.117	0.085	0.168
rs2718058	<i>NME8</i>	-0.051	0.045	0.258	0.054	0.081	0.507
rs28834970	<i>PTK2B</i>	0.071	0.045	0.111	0.034	0.081	0.679
rs9331896	<i>CLU</i>	0.049	0.047	0.290	0.026	0.084	0.762

Note: Multiple testing correction by Bonferroni 0.05/ (5 phenotypes x 20 variants); $P < 5 \times 10^{-4}$ was considered significant

Abbreviations: SNP ~ Single nucleotide polymorphism, β ~ regression coefficient, SE ~ Standard error

Supplementary Table 6: Single variant association with MRI phenotypes

SNP	Gene	White matter lesions			Hippocampal volume			Total Brain volume		
		β	SE	P-value	β	SE	P-value	β	SE	P-value
rs429358	<i>APOE*4</i>	0.002	0.023	0.919	-0.001	0.023	0.973	0.004	0.010	0.676
rs75932628	<i>TREM2</i>	0.276	0.221	0.212	-0.016	0.221	0.942	-0.050	0.093	0.591
rs6656401	<i>CR1</i>	-0.003	0.022	0.893	-0.004	0.022	0.846	-0.012	0.009	0.187
rs11218343	<i>SORL1</i>	-0.001	0.040	0.978	-0.045	0.040	0.265	-0.021	0.017	0.208
rs10838725	<i>CELF1</i>	-0.005	0.018	0.765	0.036	0.018	0.041	0.008	0.007	0.275
rs983392	<i>MS4A6A</i>	-0.004	0.017	0.815	0.012	0.017	0.460	0.002	0.007	0.803
rs10792832	<i>PICALM</i>	0.036	0.017	0.030	-0.007	0.017	0.667	-0.003	0.007	0.668
rs17125944	<i>FERMT2</i>	0.009	0.027	0.723	-0.047	0.027	0.084	0.001	0.011	0.901
rs10498633	<i>SLC24A4-RIN3</i>	-0.016	0.019	0.399	0.006	0.020	0.774	0.004	0.008	0.636
rs4147929	<i>ABCA7</i>	0.023	0.022	0.285	0.010	0.022	0.664	0.006	0.009	0.488
rs6733839	<i>BIN1</i>	0.008	0.017	0.621	-0.039	0.017	0.022	-0.002	0.007	0.814
rs35349669	<i>INPP5D</i>	0.015	0.017	0.380	-0.016	0.017	0.328	-0.010	0.007	0.132
rs7274581	<i>CASS4</i>	-0.012	0.030	0.676	-0.012	0.030	0.702	0.021	0.012	0.098
rs190982	<i>MEF2C</i>	0.013	0.017	0.431	1.67x10 ⁻⁴	0.017	0.992	-0.011	0.007	0.127
rs10948363	<i>CD2AP</i>	-0.013	0.018	0.472	0.031	0.018	0.094	1.87x10 ⁻⁴	0.008	0.980
rs1476679	<i>ZCWPW1</i>	-0.012	0.017	0.487	0.001	0.018	0.948	-0.005	0.007	0.513
rs11771145	<i>EPHA1</i>	-0.024	0.017	0.157	-0.012	0.017	0.502	0.003	0.007	0.638
rs2718058	<i>NME8</i>	0.009	0.017	0.608	-0.029	0.017	0.089	-0.001	0.007	0.912
rs28834970	<i>PTK2B</i>	-0.029	0.017	0.082	0.018	0.017	0.274	0.007	0.007	0.288
rs9331896	<i>CLU</i>	0.034	0.017	0.051	-0.006	0.018	0.749	0.004	0.007	0.597

Note: Multiple testing correction by Bonferroni 0.05/ (5 phenotypes x 20 variants); $P < 5 \times 10^{-4}$ was considered significant

Abbreviations: MRI ~ Magnetic resonance imaging, SNP ~ Single nucleotide polymorphism, β ~ regression coefficient, SE ~ Standard error

References

- [1] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet.* 2013;45:1452-8.
- [2] Ruiz A, Dols-Icardo O, Bullido MJ, Pastor P, Rodriguez-Rodriguez E, Lopez de Munain A, et al. Assessing the role of the TREM2 p.R47H variant as a risk factor for Alzheimer's disease and frontotemporal dementia. *Neurobiol Aging.* 2014;35:444 e1-4.
- [3] International Genomics of Alzheimer's Disease C. Convergent genetic and expression data implicate immunity in Alzheimer's disease. *Alzheimers Dement.* 2015;11:658-71.
- [4] Guerreiro R, Bras J, Hardy J. SnapShot: genetics of Alzheimer's disease. *Cell.* 2013;155:968- e1.

Table 1: Cohort characteristics

Characteristics	Rotterdam Study
AD data set	
Total	8893
Late-onset AD	1270
AD free controls	84.30 (6.8)
Age-of-onset (SD)	5228 (59%)
Females (%)	
MCI data set	
Total	3605
MCI cases	360
Controls	71.9 (7.2)
Age (SD)	2063 (57%)
Females (%)	
MRI data set	
Total	4527
Age (SD)	64.74 (10.8)
Females (%)	2516 (56%)

Abbreviation: SD ~ Standard deviation, AD ~ Alzheimer's disease,
MCI ~ Mild cognitive impairment, MRI ~ Magnetic resonance imaging

Table 2: Results of association of AD with risk scores-GRS

SNP Cluster*	Including APOE			Excluding APOE		
	β	SE	P-value	β	SE	P-value
GRS1 (Combined)	0.73	0.040	6.53×10^{-74}	0.69	0.101	1.12×10^{-11}
Immune response	-	-	-	0.69	0.166	3.20×10^{-5}
Endocytosis	-	-	-	0.75	0.171	1.28×10^{-5}
Cholesterol Transport	0.71	0.042	3.22×10^{-64}	0.39	0.219	0.077
Hematopoietic cell lineage†	0.73	0.042	5.16×10^{-66}	-	-	-
Hemostasis	-	-	-	0.50	0.292	0.090
Clathrin/AP2 Adaptor complex	0.72	0.042	4.68×10^{-65}	0.50	0.236	0.036
Protein folding†	0.72	0.042	2.96×10^{-64}	-	-	-

Note: Multiple testing correction by Bonferroni (0.05/ (5 phenotypes x 11 risk scores); $P < 9.09 \times 10^{-8}$ was considered significant)

Abbreviations: GRS1 ~ Combined genetic risk score, SNP ~ Single nucleotide polymorphism, β ~ Regression coefficient, SE ~ Standard error.

* Logistic regression model adjusted for age and sex in RS (N=1270 cases)

† Only one SNP available in excluding APOE GRS2

Formatted: Highlight

Formatted: Font: 8 pt

Table 3: Results of association of MCI with risk scores GRS

SNP Cluster*	Including APOE			Excluding APOE		
	β	SE	P-value	β	SE	P-value
GRS1 (combined)	0.19	0.075	0.012	0.59	0.179	9.51x10 ⁻⁴
Immune response	-	-	-	0.46	0.295	0.116
Endocytosis	-	-	-	1.16	0.305	1.44x10 ⁻⁴
Cholesterol Transport	0.11	0.082	0.164	0.39	0.392	0.322
Hematopoietic cell lineage†	0.09	0.084	0.269	-	-	-
Hemostasis	-	-	-	-0.08	0.524	0.872
Clathrin/AP2 Adaptor complex	0.12	0.082	0.128	0.72	0.423	0.089
Protein folding†	0.10	0.083	0.218	-	-	-

Note: Multiple testing correction by Bonferroni 0.05/ (5 phenotypes x 11 risk scores); P < 9.09x10⁻⁴ was considered significant

Abbreviations: GRS1 ~ Combined genetic risk score, SNP ~ Single nucleotide polymorphism, β ~ Regression coefficient, SE ~ Standard error.

* Logistic regression model adjusted for age and sex in RS (N=360 cases)

† Only one SNP available in excluding APOE pathway based GRS2

Formatted: Highlight

Formatted: Font: 8 pt

Formatted: Highlight

Table 4: Results for association of risk scores-GRS with MRI phenotypes

SNP cluster [†]	Including APOE									Excluding APOE								
	White matter lesions			Hippocampal volume			Brain volume			White matter lesions			Hippocampal volume			Brain volume		
	β	SE	P	β	SE	P	β	SE	P	β	SE	P	β	SE	P	β	SE	P
GRS1 (combined)	0.012	0.016	0.448	-0.001	0.016	0.929	0.002	0.007	0.806	0.059	0.037	0.114	-0.009	0.038	0.810	-0.006	0.016	0.724
Immune response	-	-	-	-	-	-	-	-	-	0.149	0.062	0.016	-0.024	0.062	0.706	-0.010	0.026	0.692
Endocytosis	-	-	-	-	-	-	-	-	-	0.071	0.062	0.254	-0.046	0.063	0.462	0.004	0.026	0.865
Cholesterol Transport	0.005	0.017	0.785	0.001	0.017	0.964	0.004	0.007	0.574	0.063	0.080	0.434	0.013	0.080	0.875	0.023	0.033	0.497
Hematopoietic cell lineage [†]	0.002	0.017	0.901	0.001	0.017	0.976	0.004	0.007	0.556	-	-	-	-	-	-	-	-	-
Hemostasis	-	-	-	-	-	-	-	-	-	0.228	0.108	0.034	-0.077	0.109	0.479	-0.009	0.045	0.835
Clathrin/AP2 Adaptor complex	0.011	0.017	0.507	-0.002	0.017	0.924	0.003	0.007	0.658	0.258	0.088	3.55x10⁻³	-0.077	0.109	0.479	-0.009	0.045	0.835
Protein folding [†]	0.007	0.017	0.700	-0.001	0.017	0.970	0.004	0.007	0.619	-	-	-	-	-	-	-	-	-

Note: Multiple testing correction by Bonferroni 0.05/ (5 phenotypes x 11 risk scores); P < 9.09x10⁻⁴ was considered significant

Abbreviations: GRS1 ~ Combined genetic risk score, MRI ~ Magnetic resonance imaging, SNP ~ Single nucleotide polymorphism, β ~ Regression coefficient, SE ~ Standard error, P ~ P-value

* Linear regression model with MRI phenotype as outcome and risk score-GRS as predictor, adjusted for age at MRI scan, sex in RS (N=4527)

† Only one SNP available in excluding APOE pathway based GRS2

Formatted: Highlight

Formatted: Font: 8 pt

Formatted: Highlight

Formatted: Highlight