Chapter 4.1

Genome-wide significant risk factors for Alzheimer's disease: role in progression to dementia due to AD among subjects with mild cognitive impairment

A. Lacour*, A. Espinosa*, **E. Louwersheimer**, S. Heilmann, I. Hernández, S. Wolfsgruber, V. Fernández, H. Wagner, M. Rosende-Roca, A. Mauleón, S. Moreno-Grau, L. Vargas, Y.A.L. Pijnenburg, T. Koene, O. Rodríguez-Gómez, G. Ortega, S. Ruiz, H. Holstege, O. Sotolongo-Grau, J. Kornhuber, O. Peters, L. Frölich, M. Hüll, E. Rüther, J. Wiltfang, M. Scherer, S. Riedel-Heller, M. Alegret, M.M. Nöthen, Ph. Scheltens, M. Wagner, L. Tárraga, F. Jessen, M. Boada, W. Maier, W.M van der Flier, T. Becker*, A. Ramirez*, A. Ruiz*

* contributed equally

Published in: Molecular Psychiatry, 2016

ABSTRACT

Objective Few data are available concerning the role of risk markers for Alzheimer's disease (AD) in progression to AD dementia among subjects with mild cognitive impairment (MCI). We therefore investigated the role of well-known AD-associated single nucleotide polymorphism (SNP) in the progression from MCI to AD dementia.
Methods Four independent MCI datasets were included in the analysis: (a) the German study on Aging, Cognition, and Dementia in primary care patients (n=853); (b) the German Dementia Competence Network (n=812); (c) the Fundació ACE from Barcelona, Spain (n=1245); and (d) the MCI dataset of the Amsterdam Dementia Cohort (n=306). The effects of single markers and combined polygenic scores were measured using Cox proportional hazards models and meta-analyses.
Results The Clusterin (*CLU*) locus was an independent genetic risk factor for MCI to AD progression (*CLU* rs9331888: Hazard ratio (HR)=1.187 [1.054-1.32]; p=0.0035). A polygenic score (PGS1) comprising nine established genome-wide AD risk loci predicted a small effect on the risk of MCI to AD progression in *APOE* ε4 carriers (HR=1.746 [1.029-2.965]; p=0.038). The novel AD loci reported by the International Genomics of Alzheimer's Project (IGAP) were not implicated in MCI to AD dementia progression. SNP-based polygenic risk scores comprising currently available AD genetic markers did not predict MCI to AD progression.
Conclusion We conclude that SNPs in CLU are potential markers for MCI to AD progression.

INTRODUCTION

Alzheimer's disease (AD) is the most common form of neurodegenerative dementia, representing 50 to 60% of all dementia cases. AD pathology commences years, or even decades, before the appearance of clinical symptoms, and current consensus among scientists is that prevention should be started at an early phase in individuals at increased risk. Patients with mild cognitive impairment (MCI) are at increased risk of developing AD dementia. However, the MCI group is heterogeneous, and wide variation in the annual progression to AD dementia rate has been reported, with estimates ranging from 4 to 31%. In a recent study, which involved the follow up of 550 MCI subjects for an average of 26.6 months, the present authors found that the majority (45.5%) of those MCI individuals who subsequently developed dementia displayed the AD dementia phenotype. Thus predicting which MCI cases will actually progress to AD dementia is an important challenge. Several clinical measures and biomarkers have been proposed for this purpose, including neuroimaging, cerebrospinal levels of amyloid-β and, phosphorylated and total tau. However, the predictive value of these biomarkers is low.[2,3] Accordingly, research conducted in recent decades has tended to focus on identifying factors that render MCI patients more susceptible to AD dementia.[4] This research is important, since the early detection of AD will be essential once an efficacious method of preventing or delaying the disease becomes available.

Individual risk for AD is determined by genetic, environmental, and demographic factors, as well as interactions between them. The estimated genetic component of AD, i.e. the so-called heritability, is as high as 79%. Hence in AD, the majority of pathophysiological pathways are likely to be driven by, or include, genetic determinants. Recent genome wide association studies (GWAS) and whole exome sequencing approaches have indeed identified several common and rare low-penetrance risk variants.[5–16]

Within routine clinical practice, the implementation and evaluation of AD risk markers in the prediction of MCI to AD dementia progression is in its inception. To date, the *APOE* locus is the only marker to have shown a consistent association with MCI to AD progression.[17] For other reported AD genetic markers, studies of MCI to AD dementia progression using single nucleotide polymorphisms (SNPs), or combinations of SNPs in polygenic scores (PGS), have generated conflicting results.[18,19]

The aim of the present study was to investigate the role of established AD genetic markers in the progression of MCI to AD using follow-up data from four independent MCI datasets (n=3216 subjects).

METHODS

Patients

The present cohort comprised MCI patients from Germany, Spain, and The Netherlands. These individuals were drawn from the following cohorts: (a) the German study on Aging, Cognition, and Dementia in primary care patients (AgeCoDe; n=853)[20]; (b) the German Dementia Competence Network (DCN; n=812)[21]; (c) the Fundació ACE from Barcelona (ACE, n=1245)[1]; and (d) the MCI datasets of the Amsterdam Dementia Cohort (ADC, n=306).[22] Effective sample size varied depending on phenotype analyses and co-variation matrices (table 1). Clinical characteristics, neuropsychological assessment, behavioral and functional scales, and progression to AD dementia rates for each MCI dataset are shown in table 1, and at http://detritus.fundacioace.com/pgs. The study was approved by the respective ethics committees, and all participants provided written informed consent prior to inclusion.

DNA extraction, SNP selection, and Genotyping

DNA from 3216 MCI samples was extracted using commercial methods. SNP selection was based on a review of the literature. Here, only those SNPs in loci identified by GWAS or meta-GWAS efforts were selected. To avoid missing loci, for all of the loci selected for PGS construction, whenever possible, alternative SNPs in linkage disequilibrium (LD) were also selected (i.e. LD proxies). This additional SNP thus served as a backup in the event that the primary selected SNP failed in the sequenom assay. Further details on the references used to select SNPs, the genotyping procedures, and genotyping quality control are provided in supplementary table 1 and in the genotyping procedures section of the supplementary material. The sequenom technology genotyping methods are described elsewhere.[16]

Statistical Analysis

To investigate the influence of genetic markers, demographic factors, and PGS on MCI to AD dementia progression, methods from survival analysis were used. For the 40 individual SNPs and the three PGS of interest, hazard risks were calculated using the following three models: (i) crude (model 0); (ii) age and gender adjusted (model 1); and (iii) age, gender, *APOE,* and education adjusted (model 2) (for details see statistical analysis in the supplementary material). Unless otherwise specified, the subsequent text refers to model 1 only.

PGS Construction and Evaluation

PGS were calculated in accordance with Purcell et al.[23] (for details, see the supplementary material). PGS were constructed using sets of AD associated loci identified in recent GWAS. Inclusion of SNPs in the PGS was based on definitive evidence of association

in large meta-GWAS reported by the International Genomics of Alzheimer's project (IGAP).[15] Since the established association between *APOE*-ε4 and AD is also present in our four cohorts, the *APOE*-region was excluded from the PGS calculation. Polygenic score 1 (PGS1) comprised the nine established AD-associated SNPs reported prior to publication of the IGAP consortium results (see supplementary table 1.2). Polygenic score 2 (PGS2) comprised nine of the 11 novel AD-associated SNPs identified by IGAP (supplementary table 2).[15] Polygenic score 3 (PGS3) comprised all SNPs from Polygenic Scores 1 and 2. Each of the three calculated PGS was used as a dose, and the proportional hazards model was employed using the three models applied for the analysis of single SNPs. Meta-analysis techniques were used to estimate the global effects of SNPs and PGS. The meta-analysis was conducted using the standard fixed effect approach implemented in the YAMAS software. YAMAS implements standard fixed and random effects meta-analysis, and operates on beta and standard error.[24]

RESULTS

Univariate analyses

The demographic characteristics of the cohorts are summarized in table 1. The results obtained for each analyzed SNP are shown in table 2.

In the meta-analysis, the *APOE*-ε4 allele (rs429358 C allele) showed an association with the rate of MCI to AD dementia progression in all cohorts, with a homogeneous effect being observed across datasets (hazard risk (HR)=1.84 [1.64-2.04], I^2=0, p=1.35 x 10^{-27}) (figure 1a, 2a). Interestingly, the relative risk was approximately 50% of that reported in GWAS.[6–9] Furthermore, the ε4 effect increased with age, reaching its most pronounced effect between 65 and 80 years. In contrast, the *APOE*-ε2 allele conferred a protective effect against MCI to AD dementia progression to other *APOE* genotypes (figure 1a). As with ε4, the effect of *APOE*-ε2 was dose dependent and homogeneous across datasets. The meta-analysis confirmed the protective effect of *APOE*-ε2 (HR=0.69[0·51-0.86], I^2=0, p=0.004; table 2). Six MCI subjects carrying the *APOE*-ε2 allele in a homozygous state did not progress to AD dementia during the observational time period.

Table 1. Effective sample size and baseline demographics in datasets

	AgeCoDe[a]	DCN[b]	ACE[c]	ADC[d]
patients	853	812	1245	316
detected duplicities	-1	-7	0	0
detected non MCI[e] subjects	-46	0	0	0
no/low genotypes	-3	-6	0	-10
no follow-up data*	-299	-201	-74	0
no age/sex data	0	-1	0	0
efective sample size (model 1)	504	597	1171	306
no APOE/education data	-4	-157	-1	-23
effective sample size (model 2)	500	440	1170	283
AD converters (model 1)	209 (41.4%)	76(12.7%)	395(33.7%)	110(35.9%)
AD converters (model 2)	207(41.4%)	73(16.5%)	395(33.8%)	100(35.3%)
age in years (mean)	81.6	66.2	76.0	66.8
age in years (SD)[f]	4.1	8.9	7.1	7.8
follow up time in months (mean)	43.0	26.9	26.0	27.7
observational Time in months (SD)[e]	25.4	11.2	18.9	17.7
time to conversion (mean)	38.6	19.2	21.1	27.1
time to conversión (SD)	22.6	8.5	16.0	17.4
gender (%) Female	69.4	43.7	64.6	38.9
***APOE* (%) epsilon 4**	25.8	34.2	32.4	52.3
education % of high education (>3 points in harmonized scores)	13.6	5.6	8.1	15

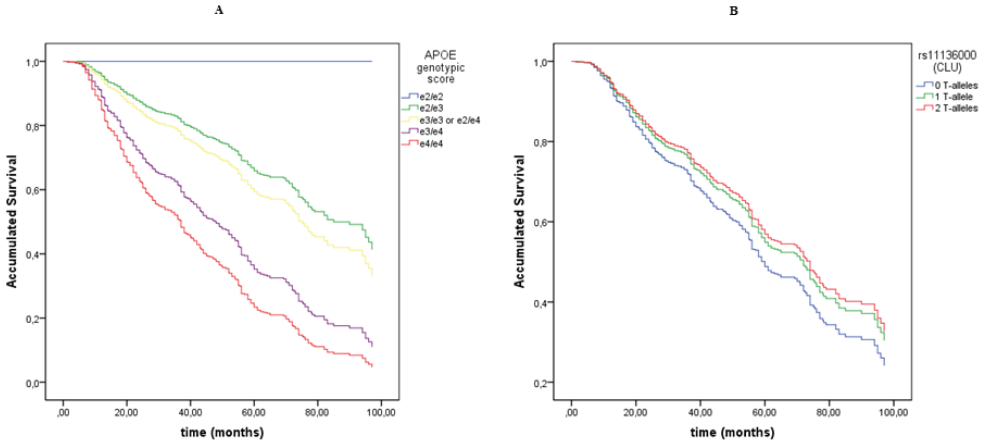
[a] AgeCoDe: German study on Aging, Cognition, and Dementia in primary care patients; [b] DCN: German Dementia Competence Network; [c] the Fundació ACE from Barcelona (ACE, n=1245); [d] ADC: Amsterdam Dementia Cohort; [e] SD: Standard Deviation;

*Subjects genotyped but without follow up

An additional association signal was observed in SNPs at the *CLU* locus (rs9331888, rs11136000). For these variants, a nominally significant result was obtained in the AgeCoDe cohort, and a consistent trend towards association was observed in the DCN, ACE, and ADC cohorts (table 2). The meta-analysis yielded a significant association for both *CLU* SNPs (p=0.003 and p=0.01, respectively). While rs11136000 showed a heterogeneous HR across the series, the HR for rs9331888 was homogeneous across the four cohorts (table 2). Association findings for the *CLU* SNPs withstood all adjustments (table 2; figure 1b and supplementary data files). No major difference in the effect sizes of the *CLU* SNPs was observed following stratification for *APOE* status, gender, or age (table 4, figure 2b; p>0.71). Stratification for these variables confirmed the orthogonality of *CLU* markers with key covariates.

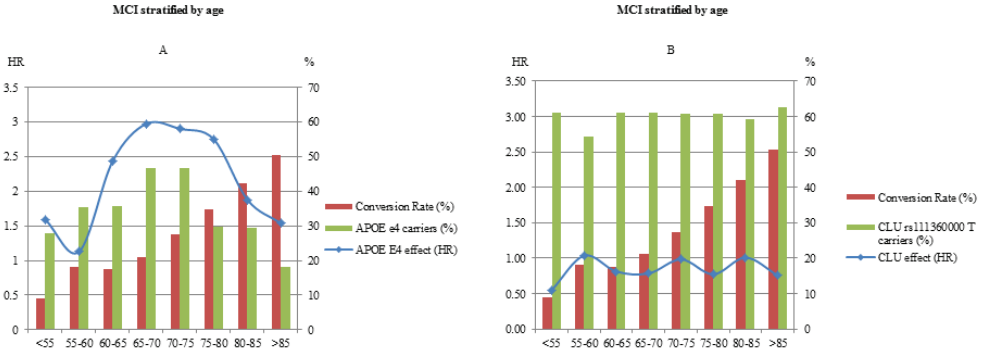
Of the remaining SNPs genotyped in the present study, only rs641120 (located at the *SORL1* locus) showed nominal significance with MCI to AD dementia progression (HR=0.89, p=0.043, model 0). However, this finding did not withstand adjustment.

Figure 1: Cox proportional hazard model multivariate dementia-free survival analyses for *APOE* genotypic score (A) and Clusterin rs111360000 (B).



Hazard Risk meta-analyses were adjusted according to dataset, age, and gender.

Figure 2: Effect size of *APOE* (A) and Clusterin (B) loci in MCI pheno-conversion to AD stratified by age.



Meta-analysis of Hazard Risk for progression to AD dementia in *APOE*-ε4 carriers following stratification for age. The progression rate for each age stratum is 460 shown in the secondary Y2 axis

PGS in MCI to AD dementia progression

The results of the hazard models analysis of PGS are shown in table 3. In the meta-analysis of PGS1, a trend towards association was observed (HR=1.31, p=0.1). Interestingly, stratification according to *APOE* genotype revealed a consistently higher effect size for PGS1 in *APOE*-ε4 carriers (table 4). The meta-analysis of PGS1 showed that the effect in *APOE*-ε4 carriers was nominally significant (HR=1.74[1.03-2.97], p=0.04). However, combined analysis revealed no statistically significant interaction between PGS1 and the *APOE* locus in any of the four datasets (p=0.14). In contrast,

PGS2 did not contribute to MCI to AD dementia progression. The effect size for PGS2 observed in the meta-analyses indicated a non-significant protective effect. This suggests that the accumulation of risk alleles was implicated in protection from MCI to AD dementia progression in the present series.

The analysis of PGS3 yielded an intermediate and non-significant result (HR=1.03, p=0.96). The PGS3 results reflect the findings of PGS1, as biased by the noise from PGS2. No significant interaction was found between PGS3 and age, gender, *APOE*-ε4 status, or cohort (tables 3, 4).

Table 2a. Effect of candidate SNPs on conversion of mild cognitive impairment to Alzheimer's disease*: meta-analysis, AgeCoDe and DCN samples.

	SNP	Chr.	Position	Minor/major	Meta-analysis				AgeCoDe sample			DCN sample		
Gene				Allele	P	HR	σHR	I	P	HR	σHR	P	HR	σHR
ABCA7	rs3764650	19	1046520	G/T	0.2350	0.90	0.08	0.0	0.8280	0.96	0.17	0.3147	0.72	0.23
ABCA7	rs3752246	19	1056492	G/C	0.2265	0.90	0.08	27.5	0.4360	0.90	0.12	0.0799	0.64	0.16
ADAMST20	rs7295246	12	43967677	G/T	0.4310	1.04	0.05	0.0	0.8770	0.99	0.09	0.2523	1.20	0.19
BIN1	rs7561528	2	127889637	A/G	0.5507	1.03	0.06	0.0	0.7590	0.97	0.10	0.1905	1.25	0.22
BIN1	rs744373	2	127894615	C/T	0.4857	1.04	0.06	4.0	0.7340	1.04	0.11	0.1277	1.31	0.23
CASS4	rs7274581	20	55018260	C/T	0.6657	0.96	0.08	0.0	0.9540	1.01	0.17	0.3610	1.27	0.33
CD2AP	rs10948363	6	47487762	G/A	0.6454	0.97	0.06	0.0	0.5560	0.93	0.11	0.9713	0.99	0.18
CD33	rs3865444	19	51727962	A/G	0.3575	1.05	0.06	0.0	0.4270	1.09	0.11	0.3488	1.17	0.19
CLU	rs11136000	8	27464519	T/C	0.0111	0.87	0.05	0.0	0.0184	0.78	0.08	0.2912	0.84	0.14
CLU	rs9331888	8	27468862	C/G	0.0035	1.19	0.07	0.0	0.1380	1.17	0.13	0.7383	1.06	0.18
CR1	rs6656401	1	207692049	A/G	0.6741	0.95	0.11	58.1	0.6520	0.95	0.12	0.1433	0.72	0.16
CR1	rs3818361	1	207784968	C/T	0.9333	0.99	0.08	19.3	0.6020	0.94	0.12	0.2388	0.78	0.16
ECHDC3	rs7920721	10	11720308	G/A	0.4273	1.04	0.05	0.0	0.5920	1.06	0.11	0.4970	0.89	0.15
EPHA1	rs10808026	7	143099133	A/C	0.7468	0.98	0.06	0.0	0.7650	0.97	0.11	0.9316	0.98	0.21
FERMT2	rs17125944	14	53400629	C/T	0.3620	0.92	0.09	0.0	0.1420	0.77	0.14	0.9712	0.99	0.30
FRMD4A	rs7081208	10	13991865	A/G	0.2568	0.90	0.09	53.2	0.3570	1.11	0.12	0.0276	0.65	0.13
FRMD4A	rs17314229	10	14016159	T/C	0.7526	1.04	0.11	0.0	0.6330	0.91	0.18	0.9726	0.99	0.30
GAB2	rs2373115	11	78091150	T/G	0.5383	0.96	0.07	0.0	0.4200	0.90	0.12	0.4543	0.84	0.20
HS3ST1	rs6448799	4	11630049	T/C	0.5308	0.96	0.06	20.1	0.3880	1.10	0.12	0.3730	0.85	0.15
INPP5D	rs35349669	2	234068476	T/C	0.9073	1.01	0.07	29.9	0.5260	1.06	0.10	0.2404	0.82	0.14
MEF2C	rs190982	5	88223420	G/A	0.1918	1.10	0.08	44.2	0.6150	0.95	0.10	0.4999	1.12	0.19
MS4A	rs4938933	11	60034429	C/T	0.3051	0.93	0.06	26.9	0.6660	1.04	0.11	0.6523	1.08	0.18
MTHFD1L	rs11754661	6	151207078	A/G	0.8502	0.98	0.11	0.0	0.8420	1.05	0.24	0.9704	0.99	0.29
NDUFAF6	rs7818382	8	96054000	T/C	0.1804	1.07	0.05	0.0	0.5010	1.07	0.11	0.5276	1.10	0.16
NME8	rs2718058	7	37841534	G/A	0.3797	1.09	0.11	69.0	0.6790	0.96	0.09	0.0196	1.49	0.25
none	rs6678275	1	193625233	C/G	0.9538	1.00	0.06	0.0	0.3890	1.11	0.14	0.9359	0.98	0.20
PICALM	rs561655	11	85800279	G/A	0.3934	0.95	0.05	0.0	0.3840	0.91	0.10	0.4566	1.13	0.19
PICALM	rs3851179	11	85868640	A/G	0.5097	0.96	0.05	0.0	0.8660	0.98	0.10	0.5659	1.10	0.19
PILRA	rs2405442	7	99971313	A/G	0.6871	0.97	0.07	27.0	0.2180	0.87	0.10	0.1933	0.79	0.14
PILRA	rs34995835	7	99990364	T/G	0.6823	0.98	0.06	0.0	0.3520	0.90	0.10	0.3331	0.84	0.15
PTK2B	rs28834970	8	27195121	C/T	0.9757	1.00	0.06	8.6	0.8990	1.01	0.11	0.1338	0.77	0.13
SCIMP	rs7225151	17	5137047	A/G	0.1055	1.13	0.08	0.0	0.9920	1.00	0.15	0.9477	0.98	0.25
SLC24A4	rs10498633	14	92926952	T/G	0.3628	0.89	0.11	64.5	0.4030	1.10	0.13	0.1515	0.73	0.16
SORL1	rs641120	11	121380965	T/C	0.0774	0.91	0.05	0.0	0.3140	0.90	0.09	0.9097	0.98	0.16
SORL1	rs11218343	11	121435587	C/T	0.9564	0.99	0.17	37.0	0.2380	1.30	0.29	0.3904	1.31	0.41
SORL1	rs2070045	11	121448090	G/T	0.7424	0.98	0.06	0.0	0.6830	0.95	0.11	0.9858	1.00	0.20
SPPL2A	rs8035452	15	51040798	C/T	0.6115	0.97	0.06	27.1	0.7820	1.03	0.10	0.2125	1.23	0.20
TOMM40	rs2075650	19	45395619	G/A	1.19e-14	1.62	0.10	0.0	1.02e-04	1.56	0.18	0.0032	1.67	0.29
TREML2	rs9381040	6	41154650	T/C	0.7648	0.98	0.08	40.6	0.9700	1.00	0.11	0.0512	0.70	0.13
CWPW1	rs1476679	7	100004446	C/T	0.7958	0.99	0.06	0.0	0.4000	0.91	0.10	0.4148	0.86	0.16

Table 2b. Effect of candidate SNPs on conversion of mild cognitive impairment to Alzheimer's disease: ACE and ADC samples

	SNP	Chr.	Position	Minor/major	ACE sample			ADC sample		
Gene				Allele	P	HR	σHR	P	HR	σHR
ABCA7	rs3764650	19	1046520	G/T	0.6034	0.94	0.11	0.2485	0.76	0.18
ABCA7	rs3752246	19	1056492	G/C	0.7572	1.03	0.10	0.1996	0.79	0.15
ADAMST20	rs7295246	12	43967677	G/T	0.7416	1.02	0.07	0.4502	1.11	0.15
BIN1	rs7561528	2	127889637	A/G	0.9193	1.01	0.09	0.4831	1.11	0.16
BIN1	rs744373	2	127894615	C/T	0.5729	0.95	0.09	0.3606	1.15	0.18
CASS4	rs7274581	20	55018260	C/T	0.4805	0.92	0.11	0.4553	0.82	0.21
CD2AP	rs10948363	6	47487762	G/A	0.7333	0.97	0.09	0.8185	1.04	0.16
CD33	rs3865444	19	51727962	A/G	0.9048	0.99	0.09	0.5796	1.08	0.15
CLU	rs11136000	8	27464519	T/C	0.1411	0.89	0.07	0.8962	1.02	0.15
CLU	rs9331888	8	27468862	C/G	0.0975	1.16	0.10	0.0210	1.41	0.21
CR1	rs6656401	1	207692049	A/G	0.0560	1.21	0.12	0.2749	0.82	0.15
CR1	rs3818361	1	207784968	C/T	0.1693	1.15	0.12	0.6728	0.93	0.16
ECHDC3	rs7920721	10	11720308	G/A	0.5159	1.05	0.08	0.4677	1.11	0.16
EPHA1	rs10808026	7	143099133	A/C	0.9940	1.00	0.10	0.7267	0.94	0.17
FERMT2	rs17125944	14	53400629	C/T	0.9970	1.00	0.15	0.7627	0.94	0.18
FRMD4A	rs7081208	10	13991865	A/G	0.3280	0.90	0.09	0.2491	0.84	0.12
FRMD4A	rs17314229	10	14016159	T/C	0.9779	1.00	0.18	0.1650	1.44	0.38
GAB2	rs2373115	11	78091150	T/G	0.8222	1.02	0.11	0.8089	0.96	0.17
HS3ST1	rs6448799	4	11630049	T/C	0.0933	0.88	0.07	0.6739	1.07	0.18
INPP5D	rs35349669	2	234068476	T/C	0.5088	0.95	0.07	0.1509	1.23	0.18
MEF2C	rs190982	5	88223420	G/A	0.0018	1.26	0.09	0.7074	1.06	0.15
MS4A	rs4938933	11	60034429	C/T	0.0230	0.83	0.07	0.4229	0.89	0.13
MTHFD1L	rs11754661	6	151207078	A/G	0.2979	0.84	0.14	0.2580	1.38	0.40
NDUFAF6	rs7818382	8	96054000	T/C	0.4346	1.06	0.07	0.5178	1.09	0.15
NME8	rs2718058	7	37841534	G/A	0.2620	0.92	0.07	0.0901	1.29	0.19
none	rs6678275	1	193625233	C/G	0.2446	0.90	0.09	0.4404	1.13	0.18
PICALM	rs561655	11	85800279	G/A	0.6238	0.96	0.08	0.4172	0.89	0.13
PICALM	rs3851179	11	85868640	A/G	0.6135	0.96	0.08	0.2728	0.85	0.13
PILRA	rs2405442	7	99971313	A/G	0.6401	1.04	0.08	0.3719	1.14	0.16
PILRA	rs34995835	7	99990364	T/G	0.7368	1.03	0.08	0.7445	1.05	0.16
PTK2B	rs28834970	8	27195121	C/T	0.9222	1.01	0.07	0.3142	1.16	0.17
SCIMP	rs7225151	17	5137047	A/G	0.0813	1.19	0.12	0.3362	1.23	0.27
SLC24A4	rs10498633	14	92926952	T/G	0.7479	1.03	0.10	0.0151	0.62	0.12
SORL1	rs641120	11	121380965	T/C	0.1913	0.90	0.07	0.4198	0.89	0.13
SORL1	rs11218343	11	121435587	C/T	0.2513	0.78	0.17	0.2515	0.64	0.25
SORL1	rs2070045	11	121448090	G/T	0.9881	1.00	0.09	0.7583	0.95	0.15
SPPL2A	rs8035452	15	51040798	C/T	0.1252	0.89	0.07	0.3613	0.87	0.13
TOMM40	rs2075650	19	45395619	G/A	1.53e-07	1.76	0.19	0.0022	1.49	0.19
TREML2	rs9381040	6	41154650	T/C	0.2735	1.09	0.08	0.7930	0.96	0.14
CWPW1	rs1476679	7	100004446	C/T	0.7532	1.03	0.08	0.6037	1.08	0.16

AgeCoDe: German study on Aging, Cognition, and Dementia in primary care patients; DCN: German Dementia Competence Network; the Fundació ACE from Barcelona; ADC: Amsterdam Dementia Cohort; SNP: single nucleotide polymorphism; Chr.: Chromosome; P: p-value; HR: Hazard risk; σHR: Hazard Risk standard deviation; I2: Heterogeneity index. *Hazard Risks were calculated with univariate Cox proportional hazard model with adjustment for age and gender (model 1).

Table 3. Effect of polygenic scores on conversion from mild cognitive impairment to Alzheimer's disease*.

	Meta-analysis				AgeCoDe[b] sample			DCN[c] sample			ACE[d] sample			ADC[e] sample		
PGS[a]	P[f]	HR[g]	C.I. 95%[h]	I^{2} [i]	P[f]	HR[g]	C.I. 95%[h]	P[f]	HR[g]	C.I. 95%[h]	P[f]	HR[g]	C.I. 95%[h]	P[f]	HR[g]	C.I. 95%[h]
PGS1	0.139	1.29	[0.86;1.72]	0.0	0.678	1.15	[0.57;2.21]	0.504	0.70	[0.25;1.99]	0.041	1.64	[1.02;2.44]	0.92	1.05	[0.43;2.56]
PGS2	0.669	0.89	[0.41;1.37]	33.1	0.690	0.85	[0.40;1.84]	0.100	0.34	[0.09;1.23]	0.667	0.88	[0.47;1.66]	0.201	2.02	[0.68;6.00]
PGS3	0.625	1.18	[0.37;2.00]	31.6	0.895	1.07	[0.40;2.88]	0.130	0.27	[0.05;1.46]	0.136	1.81	[0.83;3.94]	0.489	1.66	[0.40;6.92]

a; PGS: polygenic score
b; AgeCoDe: German study on Aging, Cognition, and Dementia in primary care patients
c; DCN: German Dementia Competence Network
d; the Fundació ACE from Barcelona
e; ADC: Amsterdam Dementia Cohort
f; P: p-value
g; HR: Hazard Risk
h; C.I. 95%: Confidence interval 95%
I; I2: Heterogeneity index.
*Hazard Risks were calculated in a univariate Cox proportional hazard model with adjustment for age and gender (model 1)

Table 4. Stratification analysis of candidate SNPs or polygenic scores by the presence of *APOE*-ε4 allele.

Marker or SNP polygenic score	APOE-ε4 carriers HR[C.I. 95%]; p-value	APOE-ε4 non-carriers HR[C.I. 95%]; p-value	Overall HR[C.I. 95%]; p-value
CLU rs9331888	1.206 [0.95-1.46]; p=0.081	1.138 [0.96-1.32]; p=0.112	1.187 [1.054-1.32]; p=0.0035
PGS 1	1.746 [1.029-2.965]; p=0.038	1.026 [0.650-1.620]; p=0.912	1.288 [0.86-1.72]; p=0.139
PGS 2 (new IGAP loci)	0.943 [0.496-1.790]; p=0.857	0.790 [0.441-1.417];p=0.428	0.889 [0.41-1.37];p=0.668
PGS 3 (all loci)	1.824 [0.805-4.132]; p=0.149	0.903 [0.433-1.883]; p=0.785	1.186 [0.37-2.00]; p=0.625

Note: effects sizes were calculated using cox proportional models adjusted by cohort, age and gender.
CLU rs933188 effect size was calculated per each T allele.
PGS: Polygenic scores. Hazard risks for PGSs were calculated per each score point.
PGS1 comprises nine genome-wide significant AD loci reported in advance of IGAP.
PGS2 comprised nine confirmed loci reported by IGAP initiative. PGS3 included 18 genome-wide loci for AD (PGS1+PGS2).
For details on PGS construction see supplementary material.

DISCUSSION

For several years, intensive research has attempted to identify the role of genetic factors in the progression of MCI to AD dementia. To date, however, only the *APOE* locus has shown a consistent association. Elias-Sonnenschein et al. performed a meta-analysis of 35 prospective MCI studies, which comprised a total of 6095 subjects.[17] Of these, 1236 individuals progressed to AD dementia within a 2.9 year period of follow up. For MCI subjects carrying the *APOE*-ε4 allele, the authors reported an odds ratio (OR) of 2.29 [1.88-2.80] for progression to AD dementia. The present findings support the hypothesis that the *APOE*-ε4 allele is implicated in MCI to AD dementia progression (HR=2.20 [1.88-2.53]) for subjects carrying *APOE-ε4* allele). However, we cannot exclude the possibility that additional loci around *APOE* may also modulate the age at onset for AD, as has been suggested for *TOMM40*, a gene adjacent to *APOE*.[25,26] Detailed mapping data of the LD region around *APOE* are now available. These have identified a poly-T length polymorphism in an intron of *TOMM40*. Interestingly, research has demonstrated that the allele distribution of the poly-T polymorphism explains a larger proportion of the observed survival curves of age at onset in AD than is the case for *APOE*-ε4 containing haplotypes alone.[25,26] To confirm the role of *TOMM40* poly-T in AD progression, genotyping of this poly-T is currently being scheduled in our large MCI dataset.

In the present study, the MCI to AD dementia progression rate increased continuously with age, whereas the effect of the allele *APOE*-ε4 on AD dementia progression decreased after the age of 80 years (figure 2a). However, previous research has shown that both the incidence of AD, and the AD risk effect of *APOE*-ε4, decrease in the elderly.[27,28] The observation of a reduced association between *APOE*-ε4 and MCI to AD dementia progression is consistent with the survivor effect, since *APOE*-ε4 is a risk factor for both a shorter lifespan and dementia.[29] A plausible hypothesis therefore is that most *APOE*-ε4-carrying MCI patients from the present cohorts had converted to dementia or died at an earlier age, thereby causing an enrichment of survivor *APOE*-ε4 MCI carriers among our elderly MCI subjects. This latter group is protected against the progression risk effect conferred by *APOE*-ε4, and this may have led to the observed reduction in the association between *APOE*-ε4 and progression to AD dementia in the present study. This hypothesis is supported by the fact that a reduced *APOE*-ε4 allele frequency was found within this age group compared to younger individuals (figure 2a).

Besides *APOE*, no other SNP or PGS combination reached study-wise statistical significance (Bonferroni corrected p-value=0.00125). However, for some of these markers (i.e. SNPs contributing to PGS1), definitive evidence of association with AD has been reported. Hence, the application of Bonferroni correction in this context

could be considered over-conservative, since our study was based on validated AD susceptibility loci.

The univariate analyses identified a consistent effect on MCI to AD dementia progression for two SNPs (rs11136000 and rs9331888) in the *CLU* gene (p=0.0035). For both SNPs, a small but consistent effect was observed in all four series, as well as in the meta-analysis. The effect sizes and allele directions of both SNPs are consistent with those reported in previous AD case control GWAS.[7] Rodriguez-Rodriguez et al. also obtained a significant result for rs11136000 allele T in MCI to AD dementia progression in a small dataset.[19] The effect size observed in the Rodriguez-Rodriguez series was inflated compared to both the present data and previous results on the role of *CLU* markers in AD risk.[6,7,15,19] Nonetheless, the reported confidence interval overlaps with our evaluation. Therefore, the present *CLU* results represent an independent replication of a previous report, and have confirmed, in a much larger sample size, the involvement of *CLU* in MCI to AD dementia progression.

Interestingly, other research has shown that the rate of cognitive decline among individuals who were cognitively normal at study baseline, but who subsequently developed MCI or AD, was significantly faster in those carrying the C allele of rs11136000 compared to noncarriers.[30] Furthermore, cognitively normal carriers of the risk allele C of rs11136000 have been reported to show a significant increase in regional cerebral blood flow in brain areas intrinsic to memory processes.[30] Overall, the genetic evidence supports the hypothesis that the *CLU* locus makes an independent contribution to MCI to AD dementia progression.

Along the same lines, the gene product of the *CLU* gene, clusterin / apolipoprotein J, has been proposed as a potential biomarker for AD. In this regard, the plasma concentration of apolipoprotein J has been associated with the severity and speed of disease progression in AD patients, as well as with atrophy of the entorhinal cortex and the hippocampus in AD.[31,32] In the prodromal stages of AD, e.g. MCI, elevated plasma levels of apoliprotein J have also been associated with lower rate of brain atrophy.[33] This atrophy involved the hippocampus and the entorhinal cortex, i.e. brain regions affected in the early stages of AD pathogenesis. Together, these findings suggest that clusterin levels respond in a selective manner along the cascade of events occurring in AD, and that this commences during the prodromal stages. This protective plasma response may modulate, at least in part, the progression of MCI to AD dementia. Our data provide additional support for this hypothesis, since they demonstrate an association between genetic variability in *CLU* and MCI to AD dementia progression. Although the precise molecular mechanism through which genetic variability in *CLU* modulates plasma clusterin levels remains unclear, research suggests the potential involvement of genetic variability in *CLU* in the modulation of gene expression. Hence, *CLU* appears a promising therapeutic target for AD.

The lack of association for most of the investigated SNPs in the present study may suggest that AD susceptibility loci have only small effects in terms of MCI to AD dementia progression risk. If this is the case, the present MCI datasets would have limited statistical power to detect them. Another power-reducing factor may have been the inclusion in MCI subjects who will never develop AD dementia or who will convert to other unrelated forms of dementia. In support of this, the effect sizes of true AD susceptibility genes in the present MCI series were low compared to conventional AD case control datasets (OR=3.5 versus HR=2.2 for *APOE*), and the progression rate for elderly MCI subjects was higher than in the case-control context. An alternative explanation is that the relative risk in GWAS studies was obtained from analysis of progression from healthy control status to AD dementia. In this case, only part of this relative risk was examined in the present study, as our series comprised individuals who were already diagnosed with MCI, many of whom had not yet converted to AD dementia but would do so in the near future.[1] Hence, "missing" relative risk in MCI studies may be found when analyzing progression of healthy individuals to MCI or by extending the period of follow-up. Alternatively, the lack of association observed for most SNPs in the present study may suggest that many genuine genetic risk factors for AD exert their pathological effects earlier, i.e. during pathological processes that occur during the pre-MCI stages of the disease. This hypothesis would imply that differing genetic factors contribute to AD susceptibility as compared to AD progression. In fact, selecting intermediate phenotypes such as MCI, which are more proximal to a specific event along the causal chain of AD, may capture more variations in the underlying heritable traits and further enhance the statistical power of the study. Interestingly, a number of previous studies selected intermediate AD phenotypes for genetic association analyses, which included neuropsychiatric test measures,[34] magnetic resonance imaging data,[35,36] biomarkers from blood and cerebrospinal fluid,[37,38] and direct measurements of AD pathology.[39] Most of the association signals identified by these studies do not overlap with known genetic susceptibility genes.

On the other hand, genetic studies based on longitudinal samples provide new insight into the pathways related to disease progression. A recent GWAS (based on 18F-florbetapir PET) of time-dependent amyloid accumulation in AD implicated the microglial activation-associated gene, *IL1RAP*.[40] Furthermore, the authors also found that *APOE* and *CLU* affect amyloid accumulation, which is consistent with the known effects of these molecules on disease susceptibility. The *IL1RAP* gene was also associated with a greater likelihood of progression from MCI to AD dementia. Interestingly, the IL1 pro-inflammatory pathway, to which *IL1RAP* belongs, is involved in plaque-associated activation of microglia and amyloid burden.[41] This inflammatory pathway is shared with *CLU* because clusterin modulates neuroinflammation by inhibiting the inflammatory response associated with complement activation.[42] In the case of *APOE*, research has

linked the gene product, apoE, to innate inflammatory responses induced via the TLR4 and IL-4R receptor pathways.[43] Furthermore, apoE and clusterin cooperate to regulate the clearance and deposition of amyloid-β in brain.[44] Notably, clusterin and apoE promote the clearance of amyloid-β by interacting with several receptors located on microglia cells, including TREM2.[42,45,46] These findings, together with our own, suggest that immune responses and microglial clearance of amyloid-β play a role in disease progression from MCI to AD dementia. Interestingly, a recent study on AD identified a significant association with signals in SNPs located in genes involved in immune response pathways,[47] suggesting a partial overlap between disease progression pathways and those that increase susceptibility to AD.

Previous AD studies have investigated the predictive value of PGS constructed using the effects sizes of multiple SNPs. For example, Verhaaren et al. constructed a genetic risk score (GRS) that was similar to the present PGS1.[48] By using this GRS in 5171 non-dementia cases from Rotterdam, the authors demonstrated that although the GRS without *APOE* was associated with the development of AD ($p=0.010$), it provided only a marginal improvement in the prediction of AD dementia beyond that provided by age, sex, and *APOE* status (area under the curve: 0.8159 vs. 0.8148, respectively). Using a similar strategy, Rodriguez-Rodriguez et al. used a GRS based on eight non-*APOE* genetic AD risk variants in order to study its effect on MCI to AD dementia progression, and on rapid progression from MCI to AD dementia.[19] Although the authors observed no association between GRS and progression risk, they found that AD-converters harboring six or more risk alleles progressed twice as rapidly to AD compared to individuals with less than six risk alleles. Thus the present findings for PGS1 are consistent with these previous studies, and support the hypothesis that the first identified AD susceptibility locus plays only a limited role in MCI to AD dementia progression. Interestingly, whereas PGS1 achieved nominal significance in *APOE*-ε4 carriers in the present study, this was not observed by Rodriguez-Rodriguez et al. (Table 4).[19] This may have been due to an enrichment of truly prodromal AD within *APOE*-ε4 carriers, who were therefore likely to progress to AD dementia within our observational time. Nevertheless, this observation with PGS1 suggests that AD susceptibility genes other than APOE also contribute to disease progression. However, the predictive value of the PGS1 composite effect for diagnosis is too small to improve prediction, and this precludes its use in routine clinical practice.

The markers included in PGS1 are the best AD-associated SNP set reported to date, since - with the exception of *CD33* - all were re-confirmed in the large replication dataset included in the IGAP effort. Many of the SNPs discovered by IGAP only reached GWAS significance during the last round of replication. Consequently, many of these SNPs still await an extensive independent replication effort to confirm genuine

loci and remove false positives.[16] The existence of some false positives among the IGAP results cannot be excluded, and this would affect PGS results.

Strength and limitations

The present study had several limitations. First, the sample size may have been too small to detect certain associations. Unexpectedly, we observed a worsening of PGS risk prediction following the addition of the novel SNPs identified by the IGAP. In fact, a non-significant protective effect was observed for PGS2 in our meta-analyses (p=0.25). A possible explanation for this finding is that novel loci included in PGS2 have even smaller effect sizes than the SNPs included in PGS1, in which case our sample would have been too small to detect association. Alternatively, the effect of the IGAP-SNPs may be restricted to very late or early onset AD, or to an undetermined and very specific sub-group of AD patients that was poorly represented in our MCI datasets. Second, only 9 of 11 novel risk loci found by IGAP were represented in PGS2 and PGS3. Unfortunately, the genotyping method failed for rs9271192 at *HLA-DRB5–HLA-DRB1,* and for rs10838725 at *CELF1,* and no additional backup SNPs were available for either locus. Thus, the conclusions drawn for PGS2 and PGS3 should be viewed with caution. Notwithstanding, the small effect sizes of the two markers are unlikely to have made a strong contribution to the overall effect of PGS2 or, more particularly, PGS3. Nevertheless, future efforts are necessary to investigate the potential implications of these missing markers (either by themselves or in combination with other loci) in terms of the progression of MCI to AD dementia.

In summary, the present data support the hypothesis that *CLU* plays an independent role in MCI to AD progression. As in previous studies, the data also confirm the role of *APOE* in this process. Furthermore, our longitudinal data suggest that the genetic effect of AD risk factors on MCI progression may be age-dependent. Finally, our findings confirm the poor predictive value of the current genome-wide AD risk loci for MCI to AD dementia progression. Further studies in larger longitudinal MCI samples are now warranted to replicate these data, and to disentangle the genetic factors which influence the progression of MCI to AD dementia. Information on loci acting in the prodromal stages of AD, i.e. in patients with MCI, will be of relevance for drug target selection in secondary prevention trials.

ACKNOWLEDGMENTS

We thank all patients for their participation in this project. We are obliged to Trinitat Port-Carbó and her family for their support of the Fundació ACE research programs. Fundació ACE collaborates with the *Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas* (CIBERNED, Spain), and is one of the participating centers of the Dementia Genetics Spanish Consortium (DEGESCO). CIBERNED is an Instituto de Salud Carlos III ISCIII Project. Agustín Ruiz is supported by grant PI13/02434 (Acción Estratégica en Salud. Instituto de Salud Carlos III (ISCIII). Ministerio de Economía y Competitividad,

Spain), and Obra Social "La Caixa" (Barcelona, Spain). Part of this work will be included in the doctoral thesis of Ana Espinosa from the University of Barcelona. Tim Becker and Markus M Nöthen are members of the German Research Council (DFG)-funded Excellence Cluster ImmunoSensation. This publication was funded in part by the German Federal Ministry of Education and Research (grants KND: 01GI0102, 01GI0420, 01GI0422, 01GI0423, 01GI0429, 01GI0431, 01GI0433, 01GI0434; grants KNDD: 01GI0710, 01GI0711, 01GI0712, 01GI0713, 01GI0714, 01GI0715, 01GI0716, 01ET1006B). Research performed at the VUmc Alzheimer center is part of the neurodegeneration research program of the Neuroscience Campus Amsterdam, the Netherlands. The VUmc Alzheimer Center is supported by Stichting Alzheimer Nederland and Stichting VUmc Fonds. The clinical database structure was developed using funding from Stichting Dioraphte. E. Louwersheimer was supported by a research fellowship from Alzheimer Nederland (WE 15-2014-04). This article is dedicated to Prof. Hanns Hippius, M.D.

REFERENCES

1. Espinosa A, Alegret M, Valero S, et al. A longitudinal follow-up of 550 mild cognitive impairment patients: evidence for large conversion to dementia rates and detection of major risk factors involved. J Alzheimers Dis. 2013;34(3):769–780.
2. Gainotti G. Origins, controversies and recent developments of the MCI construct. Curr Alzheimer Res. 2010;7:271–279.
3. Drago V, Babiloni C, Bartrés-Faz D, et al. Disease tracking markers for Alzheimer's disease at the prodromal (MCI) stage. Adv. Alzheimer's Dis. 2011. p. 331–371.
4. Weiner MW, Veitch DP, Aisen PS, et al. The Alzheimer's Disease Neuroimaging Initiative: A review of papers published since its inception. Alzheimer's Dement. 2013;9:e1–e120.
5. Strittmatter WJ, Saunders a M, Schmechel D, et al. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. Proc Natl Acad Sci U S A. 1993 Mar 1;90(5):1977–1981.
6. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet. Nature Publishing Group; 2009 Oct;41(10):1088–1093.
7. Seshadri S, Fitzpatrick AL, Schrijvers EMC, Ramirez- R, van Duijn CM, Breteler MMB. Genome-wide Analysis of Genetic Loci associated with Alzheimer disease. Jama. 2010;303(18):1832–1840.
8. Hollingworth P, Harold D, Sims R, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nat Genet. 2011 May;43(5):429–435.
9. Naj AC, Jun G, Beecham GW, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet. 2011;43(5):436–441.
10. Antúnez C, Boada M, González-Pérez A, et al. The membrane-spanning 4-domains, subfamily A (MS4A) gene cluster contains a common variant associated with Alzheimer's disease. Genome Med. 2011;3:33.
11. Jonsson T, Stefansson H, Steinberg S, et al. Variant of TREM2 associated with the risk of Alzheimer's disease. N Engl J Med. 2013 Jan 10;368(2):107–116.
12. Guerreiro R, Wojtas A, Bras J, et al. TREM2 variants in Alzheimer's disease. N Engl J Med. 2013 Jan 10;368(2):117–127.
13. Boada M, Antúnez C, Ramírez-Lorca R, et al. ATP5H/KCTD2 locus is associated with Alzheimer's disease risk. Mol Psychiatry. 2014;19(May 2013):682–687.
14. Cruchaga C, Kauwe JSK, Harari O, et al. GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. Neuron. Elsevier Inc.; 2013 Apr 24;78(2):256–268.
15. Lambert JC, Ibrahim-Verbaas C a, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013 Dec;45(12):1452–1458.
16. Ruiz a, Heilmann S, Becker T, et al. Follow-up of loci from the International Genomics of Alzheimer's Disease Project identifies TRIP4 as a novel susceptibility gene. Transl Psychiatry. 2014 Jan;4(January):e358.
17. Elias-Sonnenschein LS, Viechtbauer W, Ramakers IHGB, Verhey FRJ, Visser PJ. Predictive value of APOE-{varepsilon}4 allele for progression from MCI to AD-type dementia: a meta-analysis. J Neurol Neurosurg Psychiatry. 2011;82(Dsm Iv):1149–1156.
18. Hu X, Pickering E, Liu YC, et al. Meta-analysis for genome-wide association study identifies multiple variants at the BIN1 locus associated with late-onset Alzheimer's disease. PLoS One. 2011;6(2):e16616.
19. Rodríguez-Rodríguez E, Sánchez-Juan P, Vázquez-Higuera JL, et al. Genetic risk score predicting accelerated progression from mild cognitive impairment to Alzheimer's disease. J Neural Transm. 2013 May;120(5):807–812.

20. Kornhuber J, Schmidtke K, Frölich L, et al. Early and differential diagnosis of dementia and mild cognitive impairment: Dement Geriatr Cogn Disord. 2009;

21. Jessen F, Wiese B, Bickel H, et al. Prediction of dementia in primary care patients. PLoS One. 2011 Jan;6(2):e16852.

22. Van der Flier WM, Pijnenburg Y a L, Prins N, et al. Optimizing patient care and research: the Amsterdam Dementia Cohort. J Alzheimers Dis. 2014 Jan;41(1):313–327.

23. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009 Aug 6;460(7256):748–752.

24. Meesters C, Leber M, Herold C, et al. Quick, "Imputation-free" meta-analysis with proxy-SNPs. BMC Bioinformatics. 2012;13:231.

25. Roses AD. An inherited variable poly-T repeat genotype in TOMM40 in Alzheimer disease. Arch Neurol. 2010;67:536–541.

26. Roses a D, Lutz MW, Amrine-Madsen H, et al. A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease. Pharmacogenomics J. 2010;10:375–384.

27. Miech R a, Breitner JCS, Zandi PP, Khachaturian a S, Anthony JC, Mayer L. Incidence of AD may decline in the early 90s for men, later for women: The Cache County study. Neurology. 2002;58:209–218.

28. Valerio D, Raventos H, Schmeidler J, et al. Association of Apolipoprotein E-e4 and Dementia Declines with Age. Am J Geriatr Psychiatry. 2014;1–4.

29. Kulminski AM, Arbeev KG, Culminskaya I, et al. Age, gender, and cancer but not neurodegenerative and cardiovascular diseases strongly modulate systemic effect of the Apolipoprotein E4 allele on lifespan. PLoS Genet. 2014;10:e1004141.

30. Thambisetty M, Beason-Held LL, An Y, et al. Alzheimer risk variant CLU and brain function during aging. Biol Psychiatry. 2013;73:399–405.

31. Thambisetty M, Simmons A, Velayudhan L, et al. Association of plasma clusterin concentration with severity, pathology, and progression in Alzheimer disease. Arch Gen Psychiatry. 2010;67:739–748.

32. Hardy J, Guerreiro R, Lovestone S. Clusterin as an Alzheimer biomarker. Arch Neurol. 2011;68:1459–1460.

33. Thambisetty M, An Y, Kinsey A, et al. Plasma clusterin concentration is associated with longitudinal brain atrophy in mild cognitive impairment. Neuroimage. 2012;59:212–217.

34. McQueen MB, Bertram L, Lange C, et al. Exploring candidate gene associations with neuropsychological performance. Am J Med Genet B Neuropsychiatr Genet. 2007;144B:987–991.

35. Potkin SG, Guffanti G, Lakatos A, et al. Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. PLoS One. 2009 Jan;4(8):e6501.

36. Seshadri S, DeStefano AL, Au R, et al. Genetic correlates of brain aging on MRI and cognitive test measures: a genome-wide association and linkage analysis in the Framingham Study. BMC Med Genet. 2007;8 Suppl 1:S15.

37. Papassotiropoulos A, Streffer JR, Tsolaki M, et al. Increased brain beta-amyloid load, phosphorylated tau, and risk of Alzheimer disease associated with an intronic CYP46 polymorphism. Arch Neurol. 2003;60:29–35.

38. Peskind ER, Li G, Shofer J, et al. Age and apolipoprotein E*4 allele effects on cerebrospinal fluid beta-amyloid 42 in adults with normal cognition. Arch Neurol. 2006;63:936–939.

39. Bennett D a, De Jager PL, Leurgans SE, Schneider J a. Neuropathologic intermediate phenotypes enhance association to Alzheimer susceptibility alleles. Neurology. 2009;72:1495–1503.

40. Ramanan VK, Risacher SL, Nho K, et al. GWAS of longitudinal amyloid accumulation on F-florbetapir PET in Alzheimer ' s disease implicates microglial activation gene IL1RAP. 2015;1–13.

41. Prinz M, Priller J, Sisodia SS, Ransohoff RM. Heterogeneity of CNS myeloid cells and their roles in neurodegeneration. Nat Neurosci. 2011;13:1227–1235.

42. Nuutinen T, Suuronen T, Kauppinen A, Salminen A. Clusterin: A forgotten player in Alzheimer's disease. Brain Res Rev. 2009;61:89–104.

43. Tai LM, Ghura S, Koster KP, et al. APOE -modulated Aβ-induced neuroinflammation in Alzheimer's disease: current landscape, novel data and future perspective. J Neurochem. 2015;133:n/a – n/a.

44. DeMattos RB, Cirrito JR, Parsadanian M, et al. ApoE and Clusterin Cooperatively Suppress A?? Levels and Deposition: Evidence that ApoE Regulates Extracellular A?? Metabolism In Vivo. Neuron. 2004;41:193–202.

45. Atagi Y, Liu C-C, Painter MM, et al. Apolipoprotein E Is a Ligand for Triggering Receptor Expressed on Myeloid Cells 2 (TREM2). J Biol Chem. 2015;290(43):26043–26050.

46. Bailey CC, DeVaux LB, Farzan M. The Triggering Receptor Expressed on Myeloid Cells 2 Binds Apolipoprotein E. J Biol Chem. 2015;290(43):jbc.M115.677286.

47. Jones L, Lambert J-C, Wang L-S, et al. Convergent genetic and expression data implicate immunity in Alzheimer's disease. Alzheimers Dement. 2014;11:658–671.

48. Verhaaren BFJ, Vernooij MW, Koudstaal PJ, et al. Alzheimer's disease genes and cognition in the nondemented general population. Biol Psychiatry. Elsevier; 2013 Mar 1;73(5):429–434.

SUPPLEMENTARY MATERIAL

Genotyping Procedures

DNA Extraction

DNA was extracted using standard procedures. This was performed for a total of 3226 MCI samples from the three datasets: the German Series (DCN, n=812 and AgeCoDe, n=853); Fundació ACE (ACE, n=1245); and the MCI dataset of the Amsterdam Dementia Cohort (ADC, n=316).

Single Nucleotide Polymorphism (SNP) Selection Criteria

A total of 51 SNPs with reported evidence of association with AD risk were selected for replication and polygenic score construction (supplementary table 1). Nine of these SNPs were discarded due to technical problems. Thus 42 SNPs located in 31 different AD associated loci were investigated in the present study. Nine of these SNPs were identified in the recent International Genomics of Alzheimer's Project (IGAP) consortium study.[1] Other SNPs that had been identified by IGAP but which did not reach genome wide significance were chosen for univariate analysis but not for polygenic score analysis. The SNP rs2075650, which is located at *TOMM40*, was included in the analysis as a proxy of *APOE* effect. The SNPs rs8412 and rs429358, which determine the *APOE*-ε2/-ε3/-ε4 alleles,[2] were encoded as individual SNPs, and *APOE*-ε2/-ε3/-ε4 diplotype conventional nomenclature.

Genotyping

The primer molecules for the multiplex reaction were selected using the Assay Design Suite tool (www.mysequenom.com, Sequenom, San Diego, California, USA). Of the 51 SNPs, three SNPs (rs9349407, rs9271192, rs610932, supplementary table 1) were excluded due to problems during the design phase. The primer sequences and assay conditions for the genotyped SNPs are available upon request. In total, 48 SNPs were genotyped in all four series, representing 46 non-APOE-related SNPs and two SNPs that define the *APOE*-ε2/-ε3/-ε4 haplotypes, i.e., rs7412 and rs429358 (supplementary table 1). *APOE*-ε2/-ε3/-ε4 haplotypes were available for the majority of individuals. The SNP rs429358, whose alleles correspond to ε4 and non-ε4 carriers, was also extracted from the *APOE* haplotypes.

Quality Control

Only SNPs with a call rate of ≥ 95% and a Hardy-Weinberg Equilibrium (HWE) p-value of >0.001 in all datasets were used in the subsequent analyses (table 1). The SNPs major and minor alleles, and allelic frequencies, were consistent with previous reports. The overall genotype conversion rate during the genotyping process was 96.7%.

Of the 1101 individuals from the German datasets (DCN and AgeCoDe cohorts), 285 individuals converted to AD or MD during follow up. Subjects who converted to non-AD forms of dementia were censored at that time-point. Forty-six AD patients were excluded due to a lack of follow-up data, and one individual was excluded due to the absence of a diagnosis of either MCI or AD. Eight MCI patients had been recruited twice and were therefore removed once. Two AD and seven MCI patients were excluded due to missing genotypes.

Of the 48 genotyped SNPs, six were excluded due to genotyping failure (rs11767557 and rs34919929), HWE violations (rs10751667, rs10838725, and rs74615166), or a very low minor allele frequency (MAF) (rs8093731; MAF <0.01). Since the German sample was collected from two different sources, systematic differences in the demographics (gender and age) were tested. For gender, no significant correlation was found with either the F-test ($p=0.97$), or a t-test ($p=0.81$). In contrast, both the F-test ($p=2.2 \times 10^{-16}$) and the Welch test ($p=2.2 \times 10^{-16}$) revealed systematic deviation for age. This was attributed to differences in recruitment methodology and demographics, since the AgeCoDe cohort comprised patients with a very late onset disorder. These systematic differences were considered in all subsequent analyses. Accordingly, after quality control analyses and internal discussion, the two German series were analyzed separately.

The ACE dataset comprised 1245 MCI patients, of whom 443 converted to AD or MD. No duplicated samples were detected. Six SNPs were removed due to genotyping failure (rs11767557, rs34919929); HWE violation (rs10838725, rs74615166); high missing rate (rs10751667); or very low MAF (rs8093731).

The ADC dataset comprised 316 MCI patients, of whom 113 converted to AD. No duplicated samples were detected. A total of 10 subjects were removed due to low genotype quality. Three SNPs were removed due to either genotyping failure (rs11767557, rs34919929), or high missing rate (rs10751667). No major HWE violations were detected in the ADC series.

Overall, 42 genotyped markers fulfilled the quality control criteria (supplementary table 1.1). Two of these correspond to SNPs used to genotype APOE (rs7412 and rs429358). Main table 2 therefore displays univariate results for the remaining markers (40 SNP markers genotyped in all datasets).

Statistical Analysis

Methods from survival analysis were used to investigate the influence of genetic markers and demographic factors on the rate of progression from MCI to AD dementia.

Univariate analysis using Cox proportional hazards models

Proportional hazards models relate the time taken to the occurrence of a given event to a set of covariates that may be related to the elapsed time-period. The applied hazards model is given by

$$h(t,z) = h_0(t)\, e^{\beta_{snp} z_{snp} + \beta_{age} z_{age} + \beta_{sex} z_{sex} + \beta_{\varepsilon 4} z_{\varepsilon 4} + \beta_{cht} z_{edu}} . \qquad (1)$$

The hazard rate function h(t,z) is calculated according to accumulated information concerning the individuals who have converted during the observation period, the minimum period of time to AD dementia, and the entire period of observation. If an individual has not converted during the observation period, this period is called "censor time". Similarly, individuals who have not converted during that time-period are classified as "censored". The unknown function $h_0(t) = \lim_{\Delta t \to 0} P(t \le \tau \le t + \Delta t | t \le \tau)/\Delta t$ describes the hazard rate in the absence of covariates. The regression coefficients β_i of a proportional hazards model can be calculated from the logarithm of the hazard ratio, i.e. the ratio between hazard rate functions whose corresponding covariate states differ by one. This hazard ratio, $\exp(\beta_i)$, can be interpreted as a per-unit relative risk increase of the parameter. The unique effect of an increase of a parameter by one unit is multiplicative with respect to the hazard rate.

In the unadjusted hazards model, The parameters were evaluated in the following order: (i) the number of minor alleles $z_{snp} \in \{0,1,2\}$ in the individual for the investigated SNP (model 0, unadjusted, or crude model); (ii) the age at study recruitment z_{age} (continuous); (iii) gender $z_{sex} \in \{1,2\} \widehat{=} \{\text{male,female}\}$ (model 1); (iv) *APOE*-ε4 status, as defined by $z_{\varepsilon 4} \in \{0,1\} \widehat{=} \{\text{non-carrier,carrier}\}$; and (v) education as additional covariates (model 2). The analysis was performed using the R (The R project for scientific computing, http://www.r-project.org/) package "survival", where the ties were handled using the Efron method.[3] The response (the left hand side of the equation) was calculated in a non-parametric manner using R's function *Surv()*. P-values for the covariates were obtained using the Wald test by testing the full model against the model without the covariate of interest

Polygenic scores

Polygenic analysis is performed to test a polygenetic model, in which multiple common SNPs, which show no apparent association with the trait individually, show a collective effect on the phenotype in aggregate. Polygenic scores were calculated according to Purcell et al.[4] with the inclusion of a normalization factor.

Polygenic Score Construction

Polygenic scores were constructed using sets of AD associated loci identified in recent genome wide association studies (GWAS). For polygenic score construction, we selected only a single SNP per locus. Proxy SNPs are not independent, and therefore including both SNPs in polygenic score (PGS) may provide null information due to Linkage disequilibrium (LD) between SNPs. The inclusion of SNPs in the polygenic score was based on definitive evidence of association in large meta-GWAS reported by four large

AD genetics consortia (EADI, ADGC, GERAD, and CHARGE). Since the well-known association of *APOE*-ε4 with AD is present in our study cohorts, the *APOE*-region was excluded from our calculation of PGS. Consequently, polygenic score 1 (PGS1) represented the nine (see supplementary table 2) well-known AD-associated SNPs reported prior to publication of the IGAP consortium results (see upper part of table 2 in Lambert et al.). Polygenic score 2 (PGS2) comprised nine (see supplementary table 2) of the 11 novel AD-associated SNPs identified by IGAP (remainder of table 2 in Lambert et al.). Polygenic score 3 (PGS3) comprised the full list of 18 associated SNPs. The reasons for designing three different PGS are four-fold. First, the markers of PGS1 are the best AD-associated SNP set reported to date. Second, all SNPs of PGS1 have larger effect sizes than IGAP identified signals, since they were identified using less than half of the sample used in the IGAP effort. Third, with the exception of CD33, these SNPs have been re-confirmed in the large replication dataset included in the IGAP effort, as well as in many other independent studies. Fourth, the PGS2 SNPs, which were discovered by IGAP, still await a large independent replication effort.
The polygenic score z_{pgs} for a given individual was calculated as the sum of manifest risk alleles of all considered SNPs, with each one being weighted with the logarithm of its odds ratio.

$$z_{pgs} = \frac{\sum_{i\epsilon\{snps\}} z_i \ln OR_i}{\sum_{i\epsilon\{snps\}} \ln OR_i}, \qquad (2)$$

Here, a risk allele was characterized as one that increases the odds of AD susceptibility in a case-control analysis. For this purpose, allele information and odds ratios were gathered from Lambert et al. To ensure that interpretation of the resulting hazard ratio remained possible, the score was divided by the sum of the odds ratios' logarithms, such that the resulting score was a number between zero and two. An advantage of this approach is that it creates comparable PGSs, since otherwise hazard ratios would strongly depend on the weights. As suggested by previous authors,[4] missing genotypes were imputed with their expected values, which are calculated as two times the risk allele frequency. The latter was taken from the CEU population.[5] Individuals with a significant number (>1/3) of missing contributing SNPs were excluded from subsequent analysis.

Polygenic Score Evaluation

For each individual, the aforementioned polygenic scores were calculated. The polygenic score was used as a dose, and the proportional hazards model was employed in correspondence to the model applied for analyzing single SNPs

$$h(t,z) = h_0(t)\, e^{\beta_{pgs} z_{pgs} + \beta_{miss} z_{miss} + \beta_{age} z_{age} + \beta_{sex} z_{sex} + \beta_{\varepsilon 4} z_{\varepsilon 4} + \beta_{cht} z_{edu}}, \qquad (3)$$

where $z_{pgs} \in [0,2]$ as mentioned above. z_{miss} (continuous) is given by the sum of the odds ratios' logarithms from the missing genotypes only ($z_{miss} = \sum_{i\epsilon\{snps|missing\}} \ln OR_i$). This covariate symbolizes a cross check for correlation between the hazard and the missingness, and is not substantial.

Interaction Analysis and Plots

The effect of rs11136000, polygenic score 1, and polygenic score 2 were also analyzed after stratification for the presence of the *APOE*-ε4 allele. Rs11136000 *CLU* and polygenic score 1, which exhibited significant effects during meta-analysis, were further explored by introducing interaction terms in Cox proportional hazard models with all co-variables (*APOE*-ε4 status, education, gender, and age). Stratified analyses, interaction calculations, and graphic representations were conducted using the IBM Statistical Package for the Social Sciences (SPSS) software v19.0.0.

Meta-Analysis

Meta-analysis techniques were used to estimate global effects of SNPs and the polygenic score. These studies were conducted using the standard fixed effect approach of the Yet Another Meta-Analysis Software (YAMAS) software v.931.68.[6]

Effect of age and gender on MCI to AD dementia progression

The demographic characteristics of the four cohorts are summarized in table 1. Final effective sample size for the meta-analysis was 2578 subjects for models 0 (without covariates) and 1 (age and gender adjusted), and 2393 for model 2 (age, gender APOE, and education adjusted). In the AgeCoDe and ACE series, the average age at MCI diagnosis was significantly higher compared to the DCN and ADC series.

In all four cohorts, age significantly increased the annual chance of MCI to AD dementia progression (ACE $p=1.00 \times 10^{-11}$; AgeCoDe $p=2.02 \times 10^{-6}$; DCN $p=6.06 \times 10^{-6}$; and ADC $p=1.0 \times 10^{-3}$). The p-value differences reflected differences in the statistical power of each dataset rather than between series heterogeneity ($I^2=4.9$). Meta-analysis of age revealed a statistically significant effect in MCI to AD dementia progression. The estimated average annual increase in MCI to AD dementia progression risk was 5.8% (Hazard risk (HR)=1.058[1.05-1.07], p=2.16 x 10-21).

Compared to age, the effect of gender on MCI to AD dementia progression was less clear. Statistically significant effects were observed in the DCN, ACE, and ADC series, with the HR ranging from 1.38 ($p=4.0 \times 10^{-3}$, ACE) to 1.71 (p=0.02, DCN). However, no significant effects were found for AgeCoDe (HR=0.83, p=0.243). For gender, heterogeneous effects were found across the four study cohorts ($I^2=72.5$) and the meta-analysis identified a non-significant moderately higher risk for females (HR=1.292, p=0.096, model 1).

REFERENCES:

1. Lambert JC, Ibrahim-Verbaas C a, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013 Dec;45(12):1452–1458.
2. Fullerton SM, Clark AG, Weiss KM, et al. Apolipoprotein E variation at the sequence haplotype level: implications for the origin and maintenance of a major human polymorphism. Am J Hum Genet. 2000;67:881–900.
3. Efron B. The Efficiency of Cox's Likelihood Function for Censored Data. J Am Stat Assoc. 1977;72:557–565.
4. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009 Aug 6;460(7256):748–752.
5. http://www.ncbi.nlm.nih.gov/SNP/. Biotechnology TNC for dbSNP Short Genetic Variations.
6. Meesters C, Leber M, Herold C, et al. Quick, "Imputation-free" meta-analysis with proxy-SNPs. BMC Bioinformatics. 2012;13:231.