Evaluation of a Genetic Risk Score to Improve Risk Prediction for Alzheimer’s Disease

Vincent Chouraki\textsuperscript{a,b,*}, Christiane Reitz\textsuperscript{c}, Fleur Maury\textsuperscript{d}, Joshua C. Bis\textsuperscript{e}, Celine Bellenguez\textsuperscript{d}, Lei Yu\textsuperscript{f}, Johanna Jakobsdottir\textsuperscript{d}, Shubhabrata Mukherjee\textsuperscript{g}, Hiab H. Adams\textsuperscript{h}, Seung Hoan Choi\textsuperscript{i}, Eric B. Larson\textsuperscript{e,i}, Annette Fitzpatrick\textsuperscript{e}, Andre G. Uitterlinden\textsuperscript{h}, Philip L. de Jager\textsuperscript{i}, Albert Hofman\textsuperscript{h}, Vilmundur Gudnason\textsuperscript{d,l}, Badri Vardarajan\textsuperscript{c}, Carla Ibrahim-Verbaas\textsuperscript{h}, Sven J. van der Lee\textsuperscript{h}, Oscar Lopez\textsuperscript{m}, Jean-François Dartigues\textsuperscript{n}, Claudine Berr\textsuperscript{o}, Sudha Seshadri\textsuperscript{a,b}, and International Genomics of Alzheimer’s Project

\textsuperscript{a}Department of Neurology, Boston University School of Medicine, Boston, MA, USA  \textsuperscript{b}The Framingham Heart Study, Framingham, MA, USA  \textsuperscript{c}Columbia University, New York, NY, USA  \textsuperscript{d}Lille University, Inserm, Lille University Hospital, Institut Pasteur de Lille, U1167 - RID-AGE - Risk factors and molecular determinants of aging-related diseases; Labex Distalz, Lille, France  \textsuperscript{e}University of Washington, Seattle, WA, USA  \textsuperscript{f}Rush Alzheimer’s Disease Center, Chicago, IL, USA  \textsuperscript{g}Icelandic Heart Association, Kopavogur, Iceland  \textsuperscript{h}Erasmus Medical Center, Rotterdam, The Netherlands  \textsuperscript{i}Boston University School of Public Health, Boston, MA, USA  \textsuperscript{j}Group Health Research Institute, Seattle, WA, USA  \textsuperscript{k}Program in Translational Neuropsychiatric Genomics, Departments of Neurology & Psychiatry, Brigham and Women’s Hospital, MA, USA  \textsuperscript{l}Faculty of Medicine, University of Iceland, Reykjavik, Iceland  \textsuperscript{m}Alzheimer’s Disease Research Center, Departments of Neurology, Psychiatry and Psychology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA  \textsuperscript{n}Inserm U1219, Bordeaux Population Health, University of Bordeaux, Bordeaux, France  \textsuperscript{o}Inserm U1061 ‘Neuropsychiatry: epidemiological and clinical research’, Université de Montpellier, Montpellier, France  \textsuperscript{p}National Institute on Aging, Bethesda, MD, USA

Abstract

\textsuperscript{*}Correspondence to: Vincent Chouraki, MD, PhD, Department of Neurology/The Framingham Study, Boston University School of Medicine, 72 East Concord Street B-601, 02118, Boston, MA, USA. Tel.: +1 617 638 8069; vchourak@bu.edu.

Role of the sponsors: The funding agencies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Neurological Disorders and Stroke; the National Heart, Lung and Blood Institute; the National Institute on Aging; or the National Institutes of Health. The information contained in this article does not necessarily reflect the position or the policy of the US government, and no official endorsement should be inferred.

Authors’ disclosures available online (http://j-alz.com/manuscript-disclosures/15-0749r2).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: http://dx.doi.org/10.3233/JAD-150749.
Effective prevention of Alzheimer’s disease (AD) requires the development of risk prediction tools permitting preclinical intervention. We constructed a genetic risk score (GRS) comprising common genetic variants associated with AD, evaluated its association with incident AD and assessed its capacity to improve risk prediction over traditional models based on age, sex, education, and APOE ε4. In eight prospective cohorts included in the International Genomics of Alzheimer’s Project (IGAP), we derived weighted sum of risk alleles from the 19 top SNPs reported by the IGAP GWAS in participants aged 65 and older without prevalent dementia. Hazard ratios (HR) of incident AD were estimated in Cox models. Improvement in risk prediction was measured by the difference in C-index (Δ–C), the integrated discrimination improvement (IDI) and continuous net reclassification improvement (NRI>0). Overall, 19,687 participants at risk were included, of whom 2,782 developed AD. The GRS was associated with a 17% increase in AD risk (pooled HR = 1.17; 95%CI = [1.13–1.21] per standard deviation increase in GRS; p-value = 2.86 × 10−16). This association was stronger among persons with at least one APOE ε4 allele (HRGRS = 1.24; 95%CI = [1.15–1.34]) than in others (HRGRS = 1.13; 95%CI = [1.08–1.18]; pinteraction = 3.45 × 10−2). Risk prediction after seven years of follow-up showed a small improvement when adding the GRS to age, sex, APOE ε4, and education (Δ–Cindex = 0.0043 [0.0019–0.0067]). Similar patterns were observed for IDI and NRI>0. In conclusion, a risk score incorporating common genetic variation outside the APOE ε4 locus improved AD risk prediction and may facilitate risk stratification for prevention trials.

Keywords
Alzheimer’s disease; clinical utility; cohort studies; genetic risk score; IGAP; meta-analysis; risk prediction

INTRODUCTION

With increasing prevalence and no effective treatment, Alzheimer’s disease (AD) has become a global public health priority and a threat to healthcare systems [1]. Failure of recent clinical trials to improve cognitive functions in AD patients [2, 3] and studies suggesting a long preclinical phase of the disease [4] underline the need for earlier interventions, moving from early clinical dementia with mild cognitive complaints to a stage of preclinical AD with pathological changes of AD but no discernible clinical signs.

In this context, building accessible, easy to use prediction tools is of key importance to predict who, or which groups of persons, will develop the pathology of AD. Tools based on several genetic or imaging markers are already being used to improve enrollment in clinical trials by including asymptomatic, at-risk participants [5]. In the Dominantly Inherited Alzheimer Network Trial Units (DIAN-TU) [6] and Alzheimer Prevention Initiative (API) trials [7], risk of developing AD is assessed based on carriage of known autosomal dominant AD mutations, which represents a perfect predictor of early-onset AD, or APOE ε4 status, an important but imperfect predictor of late-onset AD [8–10]. In the Anti-Amyloid Treatment in Asymptomatic Alzheimer’s Disease (A4) trials, risk is assessed using positron emission tomography (PET) imaging of brain amyloid load but such imaging is expensive and exposes the screened population to radiation [11]. Other biomarkers can be measured in
the cerebrospinal fluid (CSF) but lumbar puncture has relatively low acceptance rates in asymptomatic adults.

Recent genome-wide association studies (GWAS) and follow-up genotyping studies have broadened the spectrum of genetic variation known to underlie late-onset AD and provided a promising ground on which to build a prediction tool [8, 12–16]. Since the effect sizes associated with individual genetic variants are small, ranging from a 23% reduction to a 22% increase in risk, an aggregated genetic risk score (GRS) could perform better than any individual variant. Such an approach has been successfully used to replicate association of genetic variants with AD, [17–23] but also to show associations of AD-related genetic variants with mild cognitive impairment [21, 22] and endophenotypes such as brain volumes measured by MRI, [24–27] cognitive tests, [18, 20, 22, 28] age at onset [23, 29], or CSF biomarkers of AD [23, 30]. These results hinted at potential benefits of using genetic information to improve the selection of asymptomatic, at-risk participants in preclinical biomarker or drug trials.

So far, the genetic variations associated with AD outside the APOE locus have shown little improvement in risk prediction over known predictors [8, 18, 19, 21–24, 26]. Moreover, only two of them have used the most recent set of single nucleotide polymorphisms (SNPs) available [21, 23] and only one used a prospective design [21] which might be better suited in a preclinical context.

Our objectives were to extend those results by constructing a GRS based on the most up-to-date results and assess its association with risk of developing incident AD and its utility for risk prediction of AD over age, sex, education, and APOE e4 in eight prospective cohorts.

MATERIALS AND METHODS

Study samples

We used the prospective cohorts from IGAP: the Framingham Heart Study (FHS), the Three City study (3C), the Rotterdam study, the Cardiovascular Health Study (CHS), the Age, Gene/Environment Susceptibility (AGES) study, the Religious Order Study/Memory and Aging Project (ROSMAP), the Adult Changes in Thought (ACT) study, and the Washington Heights-Inwood Community Aging Project (WHICAP). In most samples, baseline was defined as the time of inclusion in the study. In the FHS, we defined baseline as the time of blood draw for DNA extraction (~1990) although participants have been under dementia surveillance since 1974, to avoid survival biases. Participants aged less than 65 years at baseline were excluded as well as persons with prevalent dementia. Local ethics committees at each site approved the study and all participants provided written informed consent at enrollment (Supplementary Methods 1).

Alzheimer’s disease assessment

Screening procedures for dementia and AD were broadly congruent across studies although they differed in some specifics (Supplementary Methods 1). Most of the studies used standard screening procedures based on history, medical review, screening questions, and cognitive assessments that flagged participants with potential cognitive impairment. These
participants underwent complete neurological and neuropsychological evaluation. An initial decision was made regarding the presence or absence of dementia, using the DSM-IV criteria; [31] a diagnosis of possible, probable, or definite AD was made as a second step using NINCDS-ADRDA (National Institute of Neurological Disorders and Stroke Alzheimer’s Disease and Related Disorders Association) criteria [32].

Genetic data
Each study had genome-wide data acquired on various genotyping platforms that have been described elsewhere [16]. Genetic data were imputed using various versions of the 1000 Genomes reference panel, varying preimputation criteria and imputation software. In each study, genetic data was used as the expected allele count, i.e., allele dosage.

Selection of SNPs and construction of the GRS
To construct the genetic risk score, we selected the top SNP in each locus reaching genome-wide significance in the combined discovery-replication stages of the IGAP meta-analysis, [16] with the exception of APOE. Thus, 19 SNPs were selected: rs6656401 (CR1), rs6733839 (BIN1), rs35349669 (INPP5D), rs190982 (MEF2C), rs9271192 (HLA-DRB5/HLA-DRB1), rs10948363 (CD2AP), rs2718 058 (NME8), rs1476679 (ZCWPW1), rs11771145 (EPHA1), rs28834970 (PTK2B), rs9331896 (CLU), rs10838725 (CELF1), rs983392 (MS4A6A), rs10792 832 (PICALM), rs11218343 (SORL1), rs17125944 (FERMT2), rs10498633 (SLC24A4/RIN3), rs4147 929 (ABCA7), and rs7274581 (CASS4). We then recoded the odds ratios (OR) from the IGAP meta-analysis so they would all indicate an increased risk of AD and used the log(OR) and corresponding alleles to compute a weighted sum of risk allele dosages (Supplementary Methods 2). For descriptive purposes, we applied an additional transformation so that one unit of the GRS would be interpreted as one additional risk allele. Finally as one SNP (rs9271192) was missing in FHS, WHICAP, and Rotterdam because of poor imputation quality, an 18 SNP-based GRS was computed in these cohorts.

Statistical analyses
We used Cox regression models to estimate associations between individual SNPs, GRS, and risk of incident AD. Follow-up time was defined as time from inclusion to (1) first diagnosis of AD in incident cases, (2) last follow-up evaluation for persons who remained alive and did not develop AD, and (3) last time without known dementia for participants who developed dementia due to etiologies other than AD and those who died during follow-up.

We adjusted all models for age at baseline, sex, education levels, and presence or absence of at least one APOE e4 allele (referred below as APOE e4 status). Additional adjustment for study center and familial relationships was performed in multicentric studies and studies with related participants, respectively. We considered the GRS as a continuous variable and as tertiles. To meta-analyze results of studies utilizing the 18- and 19-SNP GRSs, we scaled all GRS so that their mean would be zero and their standard deviation one. Finally, we studied an additional model including an interaction term between APOE e4 status and
GRS. For all models, we verified that proportional hazards assumption was met by examining Schoenfeld residuals graphically.

We assessed improvement in risk prediction in each study by comparing a “base” Cox regression model containing age, sex, education, and APOE ε4 status to a “new” model containing the base model and the GRS as a continuous variable. Improvement in risk prediction was evaluated at fixed follow-up times of 5, 6, 7, and 8 years to allow comparison across studies. Three complementary measures of improvement in discrimination adapted to censored data were estimated: the difference in C-index (Δ-C), the integrated discrimination improvement (IDI), and the continuous net reclassification improvement (NRI>0) (Supplementary Methods 3). To help interpret those results, we used published simulated data to illustrate ranges of values for C-index, IDI, and NRI>0 that correspond to small, medium, and large effects (Supplementary Methods 3).

We used random-effects inverse variance meta-analyses to compute pooled estimations of hazard ratio (HR) and Δ-C. Heterogeneity across studies was quantified using $I^2$.

All analyses were performed using R and various R packages (Supplementary Methods 4). The R scripts used to generate the results presented in this study are freely available on request.

**RESULTS**

Overall, 19,687 participants at risk were included in this project, of whom 2,782 developed AD after times of follow-up ranging from a mean of 5.9 years in AGES to 10.9 years in Rotterdam. Other demographic characteristics of participants in the individual studies are presented in Table 1. After a fixed period of 7 years of follow-up, observed risk of AD ranged from 0.068 [0.047–0.097] in WHICAP to 0.214 [0.190–0.240] in ROSMAP (Supplementary Table 1). These varying risks reflect the varying ages and baseline cognitive status of the cohorts. As expected, observed risk of AD increased with time in all cohorts.

We first considered associations between individual SNPs and risk of incident AD. Most of the associations were not statistically significant (Table 2). Effect sizes were globally small, ranging from 0.92 [0.85–1.01] for CR1 to 1.07 [0.93–1.22] for CASS4 and were different from the estimations drawn from the IGAP meta-analysis (Table 2). As expected, having at least one copy of APOE ε4 was strongly associated with risk of incident AD (HR = 2.08; 95%CI = [1.92–2.26]; $p$-value = $2.09 \times 10^{-72}$; $I^2 = 0\%$). As the samples used in this project were partially overlapping with the ones used in the original IGAP study, we ran an additional IGAP meta-analysis after excluding those and did not find significant changes in the estimations of HRs for the SNPs considered (Supplementary Table 2).

Distributions of the GRS across studies are provided in Supplementary Figure 1. In nearly all studies, association of the GRS reached or was close to statistical significance, using a $p$-value threshold of 0.05 (Fig. 1). Effect sizes were remarkably consistent across studies, with HRs ranging from 1.12 in CHS to 1.24 in ROSMAP. When meta-analyzed, those results yielded a pooled HR of 1.17 (95%CI = [1.13–1.21]) per standard deviation (SD) increase in GRS; $p$-value = $2.86 \times 10^{-16}$) with no evidence of important heterogeneity ($I^2 = 0\%$). When
considering tertiles of the GRS, being in the third tertile was associated with a HR of 1.42 (95%CI = [1.29–1.55]; p-value = 1.09 × 10^{-13}) when compared to the first tertile (Supplementary Figure 2).

We observed a significant interaction between APOE ε4 status and the GRS (p-value = 3.45 × 10^{-2}), with a higher effect size in APOE ε4 carriers (HR = 1.24; 95%CI = [1.15–1.34] per SD increase in GRS; p-value = 7.63 × 10^{-8}; I^2 = 31%) than in non-carriers (HR = 1.13; 95%CI = [1.08–1.18] per SD increase in GRS; p-value = 2.40 × 10^{-7}; I^2 = 0%) (Fig. 1). APOE ε4 carriers that were in the third tertile of GRS had an increased risk of incident AD of 62% (HR = 1.62; 95%CI = [1.26–2.07]; p-value = 1.33 × 10^{-4}; I^2 = 55%) compared to APOE ε4 carriers in the first tertile of GRS. In APOE ε4 non-carriers, this risk was increased by 33% (HR = 1.33; 95%CI = [1.19–1.49]; p-value = 8.18 × 10^{-7}; I^2 = 0%) in the third tertile of GRS compared to the first tertile (Supplementary Figure 2).

Measures of improvement in risk prediction after 7 years of follow-up are presented in Fig. 2 and Supplementary Table 3. The “base” model, comprised of age, sex, education, and APOE ε4 status had good discrimination abilities, as illustrated by C-index ranging from 0.688 [0.641–0.735] in AGES to 0.803 [0.776–0.830] in FHS (Supplementary Table 3). Calibration of the models at 7 years of follow-up was good across studies (Supplementary Figure 3).

When adding the GRS in this model, the difference in C-index was small, ranging from 0.002 in FHS to 0.009 in WHICAP, and was not statistically different from zero, except in the Rotterdam study (Fig. 2). Meta-analysis of the difference in C-index yielded a pooled estimation of 0.0043 [0.0019–0.0067], homogeneous across studies (I^2 = 0%) and statistically different from zero (p-value = 4.92 × 10^{-4}). As for hazard ratio, difference in C-index was higher in APOE ε4 carriers (Δ-C = 0.0112 [0.0015–0.0208]; p-value = 2.32 × 10^{-2}; I^2 = 42%) than in non-carriers (Δ-C = 0.0018 [−0.0003–0.0039]; p-value = 9.83 × 10^{-2}; I^2 = 0%) (Fig. 2). Similar patterns were observed for IDI and NRI>0 across studies (Supplementary Table 3). Using fixed follow-up periods of 5, 6, and 8 years did not modify those figures, except in APOE ε4 carriers where a slight improvement was observed with increasing follow-up time (Supplementary Table 3 and Supplementary Figure 4).

**DISCUSSION**

In this large set of prospective cohorts from IGAP, we studied the association of an aggregate measure of the genetic variants identified in a previous work [16] with incident AD. We observed that, compared to SNPs considered separately, the aggregate measure showed a significant association with incident AD, even without inclusion of APOE ε4. Persons in the top tertile of this genetic risk score had a 42% higher risk of AD compared to persons in the bottom tertile, with higher risk among APOE ε4 carriers (62%) than in non-carriers (33%). When assessing the utility of this measure for individual risk prediction at 7 years, we observed small improvement over age, sex, education, and APOE ε4 status, with stronger effect in APOE ε4 carriers.
In this study, an increase in one standard deviation of the GRS, corresponding approximately to 2.5 risk alleles, was associated with a 17% increase in risk of developing AD. GRSs have been previously used to increase statistical power when studying associations with either AD status or AD-related endophenotypes. Results have often shown significant associations, depending greatly on whether \textit{APOE} \( \varepsilon \text{4} \) was included or not in the GRS [17–30]. Previous work from the Rotterdam Study, which was included in this study, reported similar increases of 14% (\( p = 0.010 \)) [18] and 15% (\( p = 0.058 \)) [21] in AD risk for an increase of one standard deviation of the GRS. These results confirm the small increase in risk carried on average by common genetic variants identified by GWAS and legitimate the use of GRSs to increase statistical power to identify such associations.

Persons in the top tertile of our GRS had a 42% higher risk of incident AD compared to persons in the bottom tertile. A recent study by Sleegers et al. reported an odds ratio of AD of 5.69 [4.17–7.75] when comparing the fifth quintile of a GRS to the third quintile [23]. Although their GRS included \textit{APOE} \( \varepsilon \text{4} \) and had different units which makes the comparison to our GRS difficult, these results highlight the benefit of using a sufficient number of variants outside \textit{APOE} \( \varepsilon \text{4} \) to improve stratification of participants at higher risk of AD in the context of preclinical biomarkers or drug trials.

C-indexes for the base model were relatively high in all the studies indicating good predictive value. Our decision to include or exclude variables in this model was based on literature, expected availability and expected heterogeneity in the measure across studies. In terms of clinical utility for individual risk prediction, the estimations of differences in C-index, NRI>0 and IDI were small across all of the studies, indicating limited clinical utility over age, sex, education, and \textit{APOE} \( \varepsilon \text{4} \) status. Those results are in agreement with previous studies that reported little clinical utility of genetic data outside \textit{APOE} \( \text{s} \) in improving risk prediction of AD over other predictors such as age and education level [8, 18, 19, 21–23]. Given our results, the GRS seems currently more useful for risk stratification in a population than for individual risk prediction.

The significant interaction of the GRS with \textit{APOE} \( \varepsilon \text{4} \) status hinted at potentially useful strategies. Such results could be the basis for a genetic risk prediction decision tree. One could assess the \textit{APOE} genotype, and according to it, follow-up with more complex genetic analysis. The GRS could help refine risk and age-at-onset predictions in the \textit{APOE} \( \varepsilon \text{4} \) positive subgroup that is currently the focus of several prevention efforts. Further, this measure could be a useful adjunct for risk stratification in clinical trials among the 75% of the population who are \textit{APOE} \( \varepsilon \text{4} \) negative, as 50% of the population burden of AD will manifest in this subgroup.

All the samples we used came from population studies and it would be interesting to see if the predictive performances of this GRS are improved in selected population, either persons consulting for memory complaints, mild cognitive impairment, or persons with other risk factors such as diabetes or hypertension. The timing is also of importance in assessing the utility of such a genetic risk prediction tool. We restricted our samples to persons over 65 years of age but the GRS could work best in young to middle-aged populations where gene-
environment interactions leading to AD have not had the chance to express themselves. Such questions will require an even larger sample size to answer reliably.

The main limitation of this study is the use of study populations that were already part of the initial IGAP study where the variants used in the GRS were discovered [16]. As such, it is currently not feasible to obtain a large sample outside IGAP for independent validation. Nevertheless, the eight cohorts studied here were a small fraction of the larger sample used in the discovery of the SNPs constituting the GRS. We could have used the ORs derived from the independent replication stage of IGAP but as these were very close to the ones we used to construct our GRS, this should not have dramatically affected our results. The encouraging results we obtained should therefore be replicated in an independent cohort study.

There are other decisions we made in the construction of the GRS that could be argued but in this initial attempt to construct and validate a GRS, we chose a conservative approach. Thus our selection of the SNPs used in the GRS was based purely on statistical grounds. By selecting SNPs based on GWAS p-value threshold, we limited the number of false positive signals in our GRS, but we also excluded signals that could be replicated in future studies. For example, a follow-up study has already reported replication of the TRIP4 locus that was suggestive in the IGAP meta-analysis [33]. We also excluded some signals that failed to reach genome-wide association threshold in IGAP but have persuasive biological evidence in favor of a real implication in the pathophysiology of AD, such as CD33[34]. Future work on alternative versions of a GRS could expand the list of SNPs included using a less stringent threshold of p-values and other criteria, e.g. predicted functionality or biological evidence.

Moreover, inclusion of rare variants could also improve the GRS. It is unlikely that a unique rare variant could dramatically improve risk stratification due to its low prevalence in any population, but they could improve risk prediction in selected individuals and inclusion of genetic information in genes known to harbor multiple rare variants that alter risk for the disease could represent an interesting addition.

The SNPs used to create the GRS were based on studies using a case-control design. Thus, it is possible that the genetic risk factors identified so far only act very late in the pathophysiology of AD. Nevertheless, designing a GWAS of incident AD with a number of incident cases similar to IGAP would require prospective follow-up of hundreds of thousands of participants, which is currently not feasible.

Finally our GRS did not consider any gene-gene or gene-environment interactions which could be of importance [19].

Our study has several strengths. First it presents data from the largest set of prospective studies of AD to date. The large sample sizes ensure sufficient statistical power and fairly precise estimations of risk while the prospective design ensure more accurate assessment of AD onset date. Second, we accounted for the differences in follow-up time by choosing a fixed time for estimating measures of improvement in risk prediction, ensuring comparability across studies. Finally, all studies used homogeneous definitions for the
variables used as confounders and the statistical analyses were performed using the same statistical script.

In conclusion, we have shown that evaluation of the clinical utility of a GRS requires more than study of its association with incidence of a disease. In the context of AD, a risk prediction tool build using newly identified genetic information outside $\textit{APOE} \varepsilon4$ showed a statistically significant, small improvement in individual AD risk prediction after 7 years of follow-up over age, sex, education, and $\textit{APOE} \varepsilon4$ status. Nevertheless, this GRS may be a useful tool for improving risk stratification in clinical trials aimed at preventing the onset of clinical AD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

All cohorts thank their participants and staff without whom these data could not have been collected.

The Three City study was supported by the National Foundation for Alzheimer’s disease and related disorders, the Institut Pasteur de Lille, the Centre National de Génotypage, Inserm, FRC (fondate pour la recherche sur le cerveau), Rotary and the LABEX (laboratory of excellence program investment for the future) DISTALZ grant (Development of Innovative Strategies for a Transdisciplinary approach to ALzheimer’s disease).

The Adult Changes in Thought study was supported by the following grants: U01AG006781 (Larson, PI); eMERGE genotyping 1U01HG006375 and U01HG004610 (Larson and Jarvik, multiple PIs).

The Aging Gene-Environment Susceptibility-Reykjavik Study has been funded by NIA Contract N01-AG-12100 with contributions from NEI, NIDCD and NHLBI, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association) and the Althingi (the Icelandic Parliament).

This Cardiovascular Health Study research was supported by NHLBI contracts HHSN2682012000 36C, HHSN268200800007C, N01HC55222, N01 HC85079, N01HC85080, N01HC85081, N01HC 85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL1 05756, R01HL103612, and R01HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629, R01AG15928, R01AG20098, and R01AG027058 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The Framingham Heart Study was supported by contracts with the NHLBI (N01HC25105, and HHSN268201500001I) and grants from the NIA (R01AG008122, R01AG033193, and U01AG 049505).

The MAP and ROS data in the analysis is supported by National Institute on Aging grants P30AG10161, R01AG17917, R01AG15819, R01AG30146, the Illinois Department of Public Health, and the Translational Genomics Research Institute.

The Rotterdam Study is supported by the Erasmus MC University Medical Center and Erasmus University Rotterdam; The Netherlands Organisation for Scientific Research (NWO); The Netherlands Organisation for Health Research and Development (ZonMw); the research Institute for Diseases in the Elderly (RIDE); The Netherlands Genomics Initiative (NGI); the Ministry of Education, Culture and Science; the Ministry of Health, Welfare and Sports; the European Commission (DG XII); and the Municipality of Rotterdam.

The Washington Heights-Inwood Community Aging Project was supported in part by grants from National Institutes of Health (AG037212).
REFERENCES


Fig. 1.
Association of a genetic risk score based on IGAP with incident AD. The hazard ratios are estimated for an increase in one standard deviation of the genetic risk score.
Fig. 2.
Improvement in risk prediction of a genetic risk score based on IGAP after a fixed follow-up period of 7 years. In each study, Δ-C represents the difference in C-indexes between a model adjusted for age, sex, education and APOE ε4 status (base) and a model with additional adjustment for the genetic risk score (base+GRS).
Table 1

Baseline characteristics of the participants across studies

<table>
<thead>
<tr>
<th></th>
<th>3C</th>
<th>ACT</th>
<th>AGES</th>
<th>CHS</th>
<th>FHS</th>
<th>ROSMAP</th>
<th>Rotterdam</th>
<th>WHICAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>N at risk</td>
<td>6079</td>
<td>2110</td>
<td>2553</td>
<td>1998</td>
<td>1757</td>
<td>1262</td>
<td>3334</td>
<td>594</td>
</tr>
<tr>
<td>Incident AD % (N)</td>
<td>7.7% (468)</td>
<td>19.8% (418)</td>
<td>11.0% (280)</td>
<td>18.4% (367)</td>
<td>13.1% (231)</td>
<td>28.4% (359)</td>
<td>18.2% (607)</td>
<td>8.8% (52)</td>
</tr>
<tr>
<td>Follow-up time (years)</td>
<td>6.31 (3.13)</td>
<td>8.04 (4.99)</td>
<td>5.85 (2.16)</td>
<td>6.52 (3.49)</td>
<td>7.61 (4.49)</td>
<td>7.93 (4.72)</td>
<td>10.85 (5.77)</td>
<td>6.20 (4.56)</td>
</tr>
<tr>
<td>GRS</td>
<td>19.64 (2.58)</td>
<td>19.81 (2.61)</td>
<td>19.95 (2.53)</td>
<td>19.61 (2.48)</td>
<td>19.15 (2.28)</td>
<td>19.68 (2.52)</td>
<td>19.20 (2.48)</td>
<td>18.52 (2.49)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>74.23 (5.46)</td>
<td>75.63 (6.53)</td>
<td>75.52 (5.12)</td>
<td>74.81 (4.72)</td>
<td>76.05 (7.35)</td>
<td>78.37 (7.05)</td>
<td>73.99 (6.52)</td>
<td>76.71 (6.75)</td>
</tr>
<tr>
<td>Male % (N)</td>
<td>39.2% (2384)</td>
<td>43.7% (922)</td>
<td>40.4% (1031)</td>
<td>37.8% (755)</td>
<td>42.7% (751)</td>
<td>30.5% (385)</td>
<td>39.7% (1325)</td>
<td>38.4% (228)</td>
</tr>
<tr>
<td>APOE e4% (N)</td>
<td>20.6% (1250)</td>
<td>25.2% (531)</td>
<td>27.6% (704)</td>
<td>22.7% (454)</td>
<td>20.5% (360)</td>
<td>24.2% (305)</td>
<td>27.7% (922)</td>
<td>20.9% (124)</td>
</tr>
</tbody>
</table>

Results are presented as mean (standard deviation) unless stated otherwise.

- High school: 63.8% (3881) | 10.1% (214) | 20.5% (523) | 21.5% (430) | 17.0% (298) | 4.2% (53) | 43.7% (1456) | 22.1% (131) |
- GED degree: 12.6% (765) | 24.7% (522) | 50.5% (1290) | 30.0% (600) | 37.8% (664) | 14.5% (183) | 25.8% (860) | 28.3% (168) |
- College: 23.6% (1433) | 25.5% (538) | 16.1% (412) | 24.4% (488) | 23.7% (416) | 16.1% (203) | 24.9% (831) | 35.5% (211) |
- College degree: 39.6% (836) | 12.8% (328) | 24.0% (480) | 21.6% (379) | 65.2% (823) | 5.6% (187) | 14.1% (84) |
Table 2

Association of individual SNPs with risk of incident Alzheimer’s disease and comparison with IGAP

<table>
<thead>
<tr>
<th>Locus</th>
<th>SNP</th>
<th>Chr.</th>
<th>Position</th>
<th>EAF</th>
<th>Alleles</th>
<th>Current Study</th>
<th>IGAP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Studies</td>
<td>HR</td>
</tr>
<tr>
<td>CR1</td>
<td>rs6656401</td>
<td>1</td>
<td>207692049</td>
<td>0.20</td>
<td>A/G</td>
<td>8</td>
<td>0.92</td>
</tr>
<tr>
<td>BIN1</td>
<td>rs6733839</td>
<td>2</td>
<td>127892810</td>
<td>0.41</td>
<td>T/C</td>
<td>8</td>
<td>0.96</td>
</tr>
<tr>
<td>INPP5D</td>
<td>rs35349669</td>
<td>2</td>
<td>234068476</td>
<td>0.49</td>
<td>T/C</td>
<td>8</td>
<td>0.97</td>
</tr>
<tr>
<td>MEF2C</td>
<td>rs190982</td>
<td>5</td>
<td>88223420</td>
<td>0.59</td>
<td>A/G</td>
<td>8</td>
<td>1.02</td>
</tr>
<tr>
<td>HLA-DRB5/HLA-DRB1</td>
<td>rs9271192</td>
<td>6</td>
<td>32578530</td>
<td>0.28</td>
<td>C/A</td>
<td>5</td>
<td>0.99</td>
</tr>
<tr>
<td>CD2AP</td>
<td>rs10948363</td>
<td>7</td>
<td>47487762</td>
<td>0.27</td>
<td>G/A</td>
<td>8</td>
<td>0.97</td>
</tr>
<tr>
<td>NME8</td>
<td>rs2718058</td>
<td>7</td>
<td>37841534</td>
<td>0.63</td>
<td>A/G</td>
<td>8</td>
<td>0.98</td>
</tr>
<tr>
<td>ZCWPW1</td>
<td>rs1476679</td>
<td>7</td>
<td>100004466</td>
<td>0.71</td>
<td>T/C</td>
<td>8</td>
<td>1.01</td>
</tr>
<tr>
<td>EPHA1</td>
<td>rs11771145</td>
<td>7</td>
<td>143110762</td>
<td>0.66</td>
<td>G/A</td>
<td>8</td>
<td>1.02</td>
</tr>
<tr>
<td>PTK2B</td>
<td>rs28834970</td>
<td>8</td>
<td>27195121</td>
<td>0.37</td>
<td>C/T</td>
<td>8</td>
<td>1.04</td>
</tr>
<tr>
<td>CLU</td>
<td>rs9331896</td>
<td>8</td>
<td>27467686</td>
<td>0.62</td>
<td>T/C</td>
<td>8</td>
<td>1.03</td>
</tr>
<tr>
<td>CELF1</td>
<td>rs10838725</td>
<td>11</td>
<td>47557871</td>
<td>0.32</td>
<td>C/T</td>
<td>8</td>
<td>1.01</td>
</tr>
<tr>
<td>MS4A6A</td>
<td>rs983392</td>
<td>11</td>
<td>59923508</td>
<td>0.60</td>
<td>A/G</td>
<td>8</td>
<td>1.00</td>
</tr>
<tr>
<td>PICALM</td>
<td>rs10792832</td>
<td>11</td>
<td>85867875</td>
<td>0.64</td>
<td>G/A</td>
<td>8</td>
<td>1.03</td>
</tr>
<tr>
<td>SORL1</td>
<td>rs11218343</td>
<td>11</td>
<td>121435587</td>
<td>0.96</td>
<td>T/C</td>
<td>8</td>
<td>1.00</td>
</tr>
<tr>
<td>FERMT2</td>
<td>rs17125944</td>
<td>14</td>
<td>53400629</td>
<td>0.09</td>
<td>C/T</td>
<td>8</td>
<td>1.01</td>
</tr>
<tr>
<td>SLC24A4/RIN3</td>
<td>rs10498633</td>
<td>14</td>
<td>92926952</td>
<td>0.78</td>
<td>G/T</td>
<td>8</td>
<td>1.03</td>
</tr>
<tr>
<td>ABCA7</td>
<td>rs4147929</td>
<td>19</td>
<td>1064443</td>
<td>0.19</td>
<td>A/G</td>
<td>8</td>
<td>0.94</td>
</tr>
<tr>
<td>CASS4</td>
<td>rs7274581</td>
<td>20</td>
<td>55018260</td>
<td>0.92</td>
<td>T/C</td>
<td>8</td>
<td>1.07</td>
</tr>
</tbody>
</table>

Positions on the genome are from hg19; Alleles are displayed as effect / other alleles; IGAP columns: results from the stage1+2 meta-analysis published by Lambert et al, 2013 [16]. SNP, single nucleotide polymorphism; Chr, chromosome number; EAF, effect allele frequency; HR, hazard ratio; 95% CI, 95% confidence interval; OR, odds ratio.