



Featured Article

A 22-single nucleotide polymorphism Alzheimer's disease risk score correlates with family history, onset age, and cerebrospinal fluid A β ₄₂

Kristel Slegers^{a,b,*}, Karolien Bettens^{a,b,1}, Arne De Roeck^{a,b}, Caroline Van Cauwenberghe^{a,b}, Elise Cuyvers^{a,b}, Jan Verheijen^{a,b}, Hanne Struyfs^b, Jasper Van Dongen^{a,b}, Steven Vermeulen^{a,b}, Sebastiaan Engelborghs^{b,c}, Mathieu Vandenbulcke^d, Rik Vandenbergh^e, Peter Paul De Deyn^{b,c,f}, Christine Van Broeckhoven^{a,b,*}, on behalf of the BELNEU consortium

^aNeurodegenerative Brain Diseases Group, Department of Molecular Genetics, VIB, Antwerp, Belgium

^bInstitute Born-Bunge, University of Antwerp, Antwerp, Belgium

^cDepartment of Neurology and Memory Clinic, Hospital Network Antwerp Middelheim and Hoge Beuken, Antwerp, Belgium

^dDepartment of Old Age Psychiatry and Memory Clinic, University of Leuven and University Hospitals Leuven Gasthuisberg, Leuven, Belgium

^eLaboratory for Cognitive Neurology, Department of Neurology, University of Leuven and University Hospitals Leuven Gasthuisberg, Leuven, Belgium

^fDepartment of Neurology and Alzheimer Research Center, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Abstract

Introduction: The ability to identify individuals at increased genetic risk for Alzheimer's disease (AD) may streamline biomarker and drug trials and aid clinical and personal decision making.

Methods: We evaluated the discriminative ability of a genetic risk score (GRS) covering 22 published genetic risk loci for AD in 1162 Flanders-Belgian AD patients and 1019 controls and assessed correlations with family history, onset age, and cerebrospinal fluid (CSF) biomarkers (A β ₁₋₄₂, T-Tau, P-Tau_{181P}).

Results: A GRS including all single nucleotide polymorphisms (SNPs) and age-specific *APOE* ϵ 4 weights reached area under the curve (AUC) 0.70, which increased to AUC 0.78 for patients with familial predisposition. Risk of AD increased with GRS (odds ratio, 2.32 (95% confidence interval 2.08–2.58 per unit; $P < 1.0e^{-15}$). Onset age and CSF A β ₁₋₄₂ decreased with increasing GRS ($P_{onset_age} = 9.0e^{-11}$; $P_{A\beta} = 8.9e^{-7}$).

Discussion: The discriminative ability of this 22-SNP GRS is still limited, but these data illustrate that incorporation of age-specific weights improves discriminative ability. GRS-phenotype correlations highlight the feasibility of identifying individuals at highest susceptibility.

© 2015 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords:

Alzheimer's disease; Genetic risk profile; Genotype-phenotype correlation; CSF A β ₁₋₄₂; Onset age; Family history

1. Introduction

Alzheimer's disease (AD) is an incapacitating, incurable disease characterized by a progressive loss of cognitive functions. It has a long disease course and ultimately requires fulltime medical care. Aside from drugs that temporarily relieve symptoms, no treatment exists for AD. With the

growing proportion of people >65 years, and the high costs associated with care for AD patients, the future impact of AD on society and public health is of great concern. Projection models demonstrate that even therapies with a modest effect could decrease the number of AD patients by millions [1].

AD has a long prodromal phase [2], necessitating preclinical therapeutic intervention to successfully avert the pathologic process before widespread neurodegeneration. However, knowledge of the pathophysiological processes during this prodromal phase, and hence the ability to accurately recognize or detect it, is still limited. Owing to technological progress and global collaborative research in the field of AD

The authors have no competing interests.

¹These authors contributed equally.

*Corresponding authors. Tel.: +32-3-265-1102; Fax: +32-3-265-1112.

E-mail address: kristel.slegers@molgen.vib-ua.be (K.S.), christine.vanbroeckhoven@molgen.vib-ua.be (C.V.B.).

<http://dx.doi.org/10.1016/j.jalz.2015.02.013>

1552-5260/© 2015 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

genetics (e.g. the International Genomics of Alzheimer's Project [IGAP]), our understanding of the genetic foundations of AD is rapidly expanding. Genome-wide association studies (GWAS) have resulted in the identification of 20 novel genetic risk loci in addition to *APOE* $\epsilon 4$ [3–8], which have been associated with a range of biological processes including lipid metabolism, immune response, and synaptic processes. The clinical utility of these findings for today's patients, e.g. in risk prediction, is still limited. Most of these risk factors exert only minor effects on susceptibility to AD. Combination of single single nucleotide polymorphisms (SNPs) into a genetic risk score (GRS) has been proposed to improve predictive ability, but in the absence of therapeutic consequences, the clinical application of such a score will be restricted. Nevertheless, there may already be a significant purpose for genetic risk profiling. AD is a multifactorial and heterogeneous condition. Genetic risk profiling can be used for individual molecular subclassification, by determining the pathways enriched for risk alleles within an individual. This may increase the efficiency and outcome of clinical trials and biomarker studies. In the long run, genetic risk profiling could enable tailored treatment of each individual with clinical AD. In addition, if genetic risk profiles are predictive of specific clinical characteristics such as onset age or speed of progression, this may aid decision making at the clinical and personal level. For example, a high score on an 8-SNP weighted genetic risk score (wGRSs) has been associated with a two times more rapid progression from mild cognitive impairment to AD [9].

We investigated the predictive ability of a GRS combining *APOE* $\epsilon 2/\epsilon 3/\epsilon 4$ with the 20 most significant SNPs (*CLU*, *PICALM*, *CRI1*, *ABCA7*, *MS4A6A*, *BIN1*, *CD2AP*, *CD33*, *EPHA1*, *HLA-DRB5*, *PTK2B*, *SORL1*, *SC24A4-RIN3*, *INPP5D*, *MEF2C*, *NME8*, *ZCWPW1*, *CELF1*, *FERMT2*, and *CASS4* loci) identified in recent GWAS and IGAP as well as a rare variant in *TREM2* (p.R47H), which increases risk of AD nearly fourfold [10–12] on a large, well-characterized Flanders-Belgian AD study population, part of which was previously included in the replication phase of GWAS [3–6]. Our aim was twofold. First, we aimed to determine the best discriminating model. Second, we investigated to which extent this model can define subgroups with shared phenotypic characteristics (familial disease history, age at onset, and three cerebrospinal fluid [CSF] biomarkers amyloid- β [$A\beta_{1-42}$], total tau (T-tau), and tau phosphorylated at threonine 181 [P-tau_{181P}]), which may have additional value in clinical practice or drug or biomarker development.

2. Methods

2.1. Study population

The Flanders-Belgian patient group consisted of 1328 AD patients (mean age of onset 74.4 ± 8.9 years, % female = 63.0), of which the majority was ascertained at the

Memory Clinic of the ZNA Middelheim and Hoge Beuken, Antwerp, Belgium (S.E. and P.P.D.D.) in the frame of a prospective study of neurodegenerative and vascular dementia in Flanders, the Dutch-speaking region of Belgium [13,14]. Another subset of patients was collected at the Memory Clinic of the University Hospitals of Leuven (UHL), Gasthuisberg, Leuven, Belgium (M.V. and R.V.) as part of a prospective study on the molecular genetics of cognitive impairment which was initiated in October 2006 using the same clinical assessments and biosampling schemes. Each patient underwent a neuropsychological examination and structural and/or functional neuroimaging [15].

For a subset of patients ($n = 338$; mean age at onset 76.2 ± 8.5 years, 63% = women), CSF levels of $A\beta_{1-42}$, T-tau, and P-tau_{181P} were available as part of the diagnostic work-up, determined with commercially available single parameter ELISA kits (Fujirebio Europe, Ghent, Belgium).

The control group ($n = 1123$, mean age at inclusion 64.9 ± 13.7 years, 56% = women) consisted of healthy unrelated individuals, without neurologic or psychiatric antecedents or neurologic complaints or without organic disease involving the central nervous system. Control individuals were examined at the Memory Clinic of ZNA Middelheim and Hoge Beuken, Antwerpen, Belgium and at the memory clinic at the University Hospitals of Leuven, Gasthuisberg, Leuven, Belgium. Additional community control individuals were included after interview concerning medical and family history. Memory impairment was assessed in all AD patients and control individuals by mini mental state examination (MMSE) (control cohort MMSE >26 ; Folstein et al.). Patients and control individuals originated from the same geographical area (Flanders-Belgium). There is no evidence of population substructure (Supplementary Fig. 1).

Part of the Flanders-Belgian Alzheimer ($n = 878$) and control ($n = 661$) population was included in the replication phase of the IGAP [6].

2.2. Ethical assurances

All participants and/or their legal guardian gave written informed consent for participation in clinical and genetic studies. Clinical study protocol and the informed consent forms for patient ascertainment were approved by the Ethics Committee of the respective hospitals at the cohort sampling sites in Flanders-Belgium. The genetic study protocols and informed consent forms were approved by the Ethics Committees of the University of Antwerp and the University Hospital of Antwerp, Belgium and the University Hospitals of Leuven, Belgium. After informed consent, blood samples of patients and control individuals were collected for genetic studies.

2.3. Genotyping

We selected 22 polymorphisms to construct a GRS. This included *APOE* $\epsilon 2/\epsilon 3/\epsilon 4$, *TREM2* p.R47H, and 20 SNPs at the AD risk loci identified in recent GWAS. At each locus,

we selected the most significant SNP from the GWAS in which each locus was first reported. Genotype data were available from previous studies for *APOE* $\epsilon 2/\epsilon 3/\epsilon 4$, *TREM2* p.R47H [10], *CLU* rs11136000 [16], *CR1* rs3818361 [17,18], and *BINI* rs744373 [19] on the full cohort, and for *PICALM* rs3851179, *CD33* rs3865444, *CD2AP* rs9349407, *ABCA7* rs3764650, *MS4A6A* rs610932, *EPHA1* rs11767557, *HLA-DRB5* rs9271192, *PTK2B* rs28834970, *SORL1* rs11218343, *SLC24A4-RIN3* rs10498633, *INPP5D* rs35349669, *MEF2C* rs190982, *NME8* rs2718058, *ZCWPW1* rs1476679, *CELFI* rs10838725, *FERMT2* rs17125944, and *CASS4* rs7274581 on part of the cohort (n = 878 patients, n = 661 control individuals) [6]. Additional genotyping to generate genotype data for the full cohort was for most SNPs performed by MassARRAY using iPLEX Gold chemistry (Sequenom, Hamburg, Germany), followed by MALDI-TOF mass spectrometry. Polymerase chain reaction and extension primers were designed using MassARRAY AssayDesign software v3.0.2.0. Genotypes were called automatically using MassARRAY Typer software v4.0 and were visually inspected by two researchers blinded for disease outcome. SNPs in *PICALM*, *EPHA1*, *HLA-DRB5*, *SORL1*, *ZCWPW1*, and *CELFI* were genotyped by Sanger sequencing. Primers for these SNPs were designed with Primer-BLAST, and sequences were analyzed by two independent researchers using SeqMan or novoSNP [20]. Interplate controls showed 100% concordance for all SNPs. All selected SNPs had a minor allele frequency >5% and were in Hardy-Weinberg equilibrium in the control population ($P > .001$). Inhouse genotyping and IGAP data showed 99.5% concordance. An average genotyping success rate of 93.1% was obtained over the 22 susceptibility genes.

2.4. Statistical analyses

2.4.1. Single SNP analyses

Allele frequencies of AD patients and control individuals were compared using χ^2 statistics. Odds ratios (OR) (calculated relative to the common genotype) and 95% confidence intervals (CIs) were corrected for gender, *APOE* $\epsilon 4$ genotype (presence of one or two *APOE* $\epsilon 4$ alleles vs. the absence of any *APOE* $\epsilon 4$ allele), and age at onset (age at inclusion for control individuals) using a logistic regression model. Single SNP analyses were performed using SPSS 20.0 version for Windows (IBM SPSS Inc, Chicago, IL).

2.4.2. Genetic risk scores

Different GRSs were composed (Table 1). A counted genetic risk score (cGRS) was derived by taking the sum of risk alleles per individual (Model_{ALL-C}). To account for the strength of the genetic association of each allele, wGRSs were derived by multiplying the number of risk alleles for each SNP with the natural logarithm of their respective OR (lnOR; based on discovery ORs [4,6,7,11] [Table 2])

Table 1
GRS models

Model	Algorithm	Number of SNPs	<i>APOE</i> included	<i>APOE</i> weight
APOE-S	wGRS	1	+	Standard
APOE-A	wGRS	1	+	Age-related
ALL-C	cGRS	22	+	NA
ALL-WS	wGRS	22	+	Standard
ALL-WA	wGRS	22	+	Age-related
OTH-W	wGRS	21	-	NA

Abbreviations: GRS, genetic risk score; SNPs, single nucleotide polymorphisms; wGRS, weighted genetic risk score; NA, not applicable.

NOTE. The SNPs were combined by counting (cGRS), or weighted (wGRS) allele algorithms. *APOE* was either included (+) or excluded (-), and *APOE* weighting was done with standard [3] or age-related weights [21].

and taking the sum of these products per sample. Two wGRSs were created, differing in how information about *APOE* $\epsilon 4$ was incorporated. In the first scenario, the contribution of *APOE* to the GRS was weighted by the simple effect estimate of *APOE* $\epsilon 4$ on an independent AD GWAS cohort of European ancestry (Model_{ALL-WS}) [3]. In the second scenario, weighting of *APOE* was based on age-specific ORs for different *APOE* $\epsilon 2/3/4$ genotypes as described by Genin et al. [21] to approximate the known association between *APOE* $\epsilon 4$ and onset age (Model_{ALL-WA}).

After removing individuals with five or more missing genotypes (166 patients and 104 control individuals), a SNP score for missing genotypes was imputed by the product of lnOR and the expected probability of having a risk allele (briefly: $0 \times (1-q)^2 + 1 \times 2(1-q)q + 2 \times q^2 = 2q$, where q equals the risk allele frequency). The final data set available for GRS analyses comprised 1162 patients and 1019 control individuals. To predict the discriminative ability of the cGRS and wGRS for AD, receiver operating characteristic (ROC) curves were generated with pROC in R [22] by plotting the true positive rate against the false positive rate and calculating the area under the curve (AUC). Statistical comparison of ROC curves for different models was performed with bootstrap analysis in a paired or nonpaired approach, depending on the comparison of GRS within or between groups. Stratification analyses were performed for *APOE* and for familial history. For *APOE* stratification, GRSs without *APOE* $\epsilon 4$ (Model_{OTH-W}) were compared between *APOE*- $\epsilon 4$ noncarriers and individuals carrying one or two *APOE*- $\epsilon 4$ alleles. For familial history stratification, GRSs were compared between familial (FAD) and sporadic AD (SAD) patients (Model_{ALL-WA} and Model_{OTH-W}). Known pathogenic mutation carriers in *APP*, *C9orf72*, *PSEN1*, *PSEN2*, and *GRN* were excluded (n = 16).

The relationship between the best GRS model (Model_{ALL-WA}) and risk of AD, FAD, and SAD was modeled using logistic regression in R, using GRS as a continuous predictor. In addition, individuals were partitioned into five quantiles (quintiles) based on Model_{ALL-WA} (with cutoffs

Table 2
Characteristics of the selected SNPs and allelic association in the Flanders-Belgian cohort

Gene	SNP	Published OR	Minor allele	Risk allele	Risk allele frequency			OR (95% CI)	P value
					Controls	Patients	P value		
<i>APOE</i> [3]	rs7412, rs429358	2.53	ε4	ε4	0.14	0.32	1.0 ^{e-7}	3.23 (2.74–3.82)	1.0 ^{e-7}
<i>CRI</i> [4]	rs3818361	1.21	A	A	0.19	0.22	.007	1.15 (0.97–1.36)	.1
<i>BINI</i> [4]	rs744373	1.17	G	G	0.30	0.31	.5	1.24 (1.07–1.45)	.005
<i>CD2AP</i> [7]	rs9349407	1.14	C	C	0.28	0.30	.2	1.10 (0.95–1.28)	.2
<i>ABCA7</i> [4]	rs3764650	1.22	G	G	0.09	0.11	.005	1.37 (1.09–1.72)	.007
<i>HLA-DRB5</i> [6]	rs9271192	1.11	C	C	0.26	0.30	.02	1.13 (0.96–1.33)	.1
<i>PTK2B</i> [6]	rs28834970	1.10	C	C	0.36	0.39	.07	1.08 (0.93–1.26)	.3
<i>INPP5D</i> [6]	rs35349669	1.07	T	T	0.47	0.48	.4	1.03 (0.89–1.18)	.7
<i>CELF1</i> [6]	rs10838725	1.08	C	C	0.32	0.34	.2	1.12 (0.96–1.31)	.1
<i>FERMT2</i> [6]	rs17125944	1.13	C	C	0.09	0.10	.1	1.03 (0.81–1.32)	.8
<i>TREM2</i> [11]	rs75932628	4.50	T	T	0.001	0.004	.08	4.41 (1.08–18.0)	.04
<i>EPHA1</i> [7]	rs11767557	0.85	C	T	0.81	0.82	.6	1.11 (0.93–1.32)	.3
<i>CLU</i> [4]	rs11136000	0.83	T	C	0.60	0.63	.03	0.82 (0.71–0.94)	.006
<i>MS4A6A</i> [4]	rs610932	0.88	T	G	0.58	0.60	.1	0.92 (0.80–1.06)	.3
<i>PICALM</i> [4]	rs3851179	0.87	T	C	0.65	0.68	.06	0.84 (0.72–0.98)	.02
<i>CD33</i> [7]	rs3865444	0.88	A	C	0.69	0.69	.6	0.96 (0.83–1.12)	.6
<i>SORL1</i> [6]	rs11218343	0.76	C	T	0.95	0.96	.006	0.67 (0.47–0.96)	.03
<i>SLC24A4</i> [6]	rs10498633	0.90	T	G	0.80	0.79	.3	1.08 (0.90–1.29)	.4
<i>MEF2C</i> [6]	rs190982	0.92	G	A	0.59	0.61	.4	0.94 (0.81–1.08)	.4
<i>NME8</i> [6]	rs2718058	0.93	G	A	0.62	0.63	.7	1.00 (0.86–1.16)	1.0
<i>ZCWPW1</i> [6]	rs1476679	0.92	C	T	0.66	0.71	.001	0.80 (0.69–0.94)	.005
<i>CASS4</i> [6]	rs7274581	0.87	C	T	0.91	0.92	.2	0.84 (0.64–1.09)	.2

Abbreviations: SNPs, single nucleotide polymorphisms; OR, odds ratio; CI, confidence interval.

NOTE. Selected SNPs are presented with their published discovery odds ratios (OR), given for the minor allele. For *APOE*, OR was calculated within the Flanders-Belgian population. Frequency of the risk alleles, which are included in GRS models, is presented for Flanders-Belgian patients and controls, along with a nominal *P* value based on χ^2 statistics. The OR and 95% confidence interval (95% CI) obtained in the Flanders-Belgian cohort are given for the minor allele and are adjusted for *APOE* ε4 status (for all SNPs except *APOE* genotype itself), age at onset or inclusion, and gender.

at scores of 2.62, 2.96, 3.43, and 4.31). The distribution of FAD, SAD, and control subjects was evaluated for each fifth. ORs and 95% CIs for disease outcome were calculated per fifth with the third group, containing the 20% average scores in this study, as the reference. Graphical display of results was performed using *plotrix* in R [23].

To evaluate risk profile redundancy, wGRSs comprising subsets of variants were generated. Risk factors were ranked based on their discovery ORs in independent cohorts [4,6,7,11]. Cumulative scores of variants with decreasing impact were then calculated on the Flanders-Belgian cohort and subjected to ROC-curve evaluation. Comparison between these reduced models was done with paired bootstrap analysis.

The relationship between the GRS and AD-related endophenotypes (onset age and CSF biomarker levels Aβ₁₋₄₂, T-tau, and P-tau_{181P}) was assessed in linear regression models in R. For the CSF biomarker analyses, Model_{ALL-WA} was used as a predictor variable. T-tau and P-tau were log₁₀ transformed to approximate normality. Covariates age and disease duration at lumbar puncture were evaluated in a stepwise procedure, resulting in exclusion from the final regression model for all three biomarkers. For onset age analyses, Model_{ALL-WS} without age-specific weights for *APOE* genotypes was used as a predictor variable to allow correlation analysis with onset age.

3. Results

3.1. Single SNP associations in the Flanders-Belgian AD cohort

We included 22 polymorphisms in this study, which were previously implicated in AD risk (Table 2; Table S1). Seven polymorphisms (at *APOE*, *CRI*, *ABCA7*, *SORL1*, *ZCWPW1*, *HLA-DRB5*, and *CLU*) showed nominal evidence of association with AD in the full Flanders-Belgian study cohort and eight polymorphisms remained (*APOE*, *ABCA7*, *SORL1*, *ZCWPW1*, and *CLU*) or became (*BINI*, *TREM2*, and *PICALM*) nominally significantly associated after adjusting for age, gender, and *APOE* ε4 status. Only *APOE* ε4 would be considered significant after Bonferroni correction for 22 tests.

3.2. GRS model comparisons

We composed four different GRS models (Table 1) and compared them against each other as well as with the discriminative ability of two models based on *APOE* genotype alone (Fig. 1A). The AUC based on the weighted alleles of *APOE* ε4 only (Model_{APOE-S}) was 0.65. The discriminative ability of *APOE* improved when taking into account age-specific weights for the different *APOE* genotypes (Model_{APOE-A}; AUC = 0.67, *P* = 5.2e⁻⁹). Model_{ALL-C}, which was derived by the sum of risk alleles at the 22 polymorphisms, had less discriminative ability toward disease outcome in the

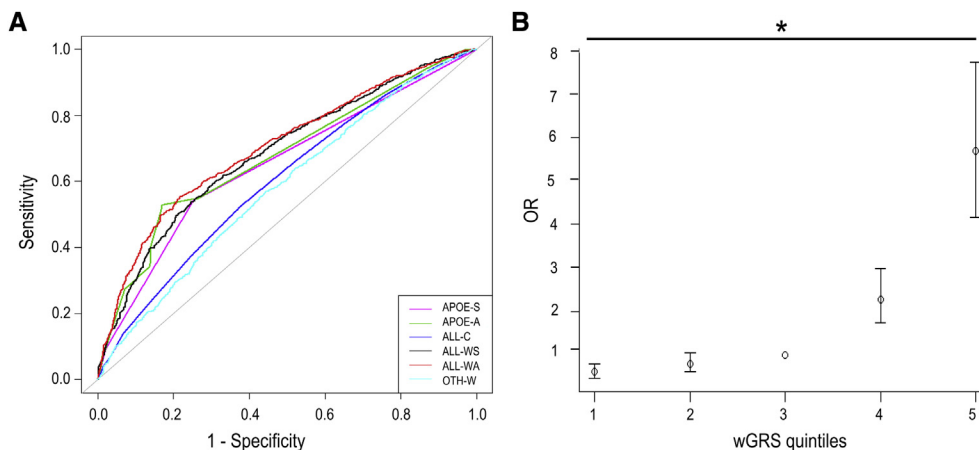


Fig. 1. Comparison of the discriminative ability of different models of GRS for AD. (A) Receiver operating characteristic (ROC) curves for the six GRS models detailed in Table 1, based on the AD and control cohort. Area under the curve (AUC) for APOE-S: 0.65, APOE-A: 0.67; ALL-WA: 0.70; ALL-WS: 0.68; ALL-C: 0.60, and OTH-W: 0.58. (B) Odds ratios (OR) of AD are presented with 95% confidence intervals (CI) for each quintile of Model_{ALL-WA}, with the third (20% average GRS) quintile as reference. The x-axis gives numbers for increasing quintiles. *denotes a P value of $<1.0e^{-15}$ based on Cochran-Armitage test for trend. Abbreviations: GRS, genetic risk score; AD, Alzheimer's disease; wGRS, weighted genetic risk score.

Flanders-Belgian population than Model_{APOE-S} or Model_{APOE-A}, with an AUC of 0.60. The inclusion of weights for strength of associations in the full GRS models (wGRSs) resulted in ameliorated prediction as expected (Model_{ALL-C} vs. Model_{ALL-WA}; $P < 1.0e^{-15}$). In line with the results of the two *APOE* models, Model_{ALL-WA}, which takes into account age-specific weights for *APOE*, resulted in a slight but significant improvement of AUC compared with the simple weighted Model_{ALL-WS} (0.70 vs. 0.68, $P = 2.0e^{-6}$). The predictive potential of a model including 21 SNPs without *APOE* (Model_{OTH-W}, AUC = 0.58) was less than that of *APOE* but the SNPs did have a significant added value when combined with *APOE* (Model_{ALL-WA} vs. Model_{APOE-A}; $P = 2.8e^{-6}$). Stratification by *APOE* did not reveal significant differences in ROC curves for Model_{OTH-W} comprising the 21 other variants ($AUC_{APOE \epsilon 4-} = 0.57$ vs. $AUC_{APOE \epsilon 4+} = 0.59$, $P = .4$; Supplementary Figure 2).

Logistic regression analysis using the best GRS (Model_{ALL-WA}) showed an increased risk of AD with increasing risk score (OR, 2.32 [95% CI, 2.08–2.58] per unit increase on the GRS; Z-statistic 15.3, $P < 1.0e^{-15}$). On quintile partitioning, individuals in the highest fifth had an OR of 5.69 (95% CI, 4.17–7.75) of AD compared with the center group (Fig. 1B).

A parsimony analysis indicated that a reduced model, including only 16 variants and age-related *APOE* weights, resulted in similar discriminative ability as a full wGRS model. The AUC of this parsimonious model (0.70) was comparable with Model_{ALL-WA} (AUC = 0.70; $P = .2$; Supplementary Fig. 3).

3.3. GRS and family history of AD

The discriminative ability of GRS Model_{ALL-WA} toward disease outcome was significantly higher in AD patients

with a positive family history ($AUC_{FADvsCon} = 0.78$; $AUC_{SADvsCon} = 0.68$ [$P = 1.7e^{-5}$]; Fig. 2A). Although much of the discriminative ability was attributable to *APOE* (Model_{APOE-A}; $AUC_{FADvsCon} = 0.76$; $AUC_{SADvsCon} = 0.65$ [$P = 5.6e^{-7}$]) and the AUCs for FAD and SAD were not significantly different for Model_{OTH-W}, the 21 genes had a significant added value when combined with *APOE* in FAD patients ($P = .01$) as well as in SAD patients ($P = 5.5e^{-8}$) in contrast to *APOE* only (Fig. 2A).

Risk of AD increased with increasing GRS for both FAD and SAD but the risk increased more strongly for FAD (OR, 3.01 [95% CI, 2.55–3.58] per unit increase on Model_{ALL-WA}; Z-statistic, 12.7; $P < 1.0e^{-15}$) than for SAD (OR, 2.14 [95% CI, 1.92–2.40]; Z-statistic, 13.3; $P < 1.0e^{-15}$). The proportion of familial and sporadic Alzheimer patients was the highest in the highest risk quintiles of Model_{ALL-WA} (Fig. 2B). In total, 41% of all patients with a positive family history had a wGRS >4.31 , as opposed to only 8% of all control subjects (Fig. 2B). Patients in the upper quintile had an OR, 2.03 (95% CI, 1.27–3.27) of having a familial history of AD compared with patients in the center quintile.

3.4. GRS—endophenotype correlations

The GRS was negatively correlated with onset age (Pearson $r = -0.20$, $P = 9.0e^{-11}$; Fig. 3). Onset age decreased with 2.39 years (standard error [SE], 0.36; $P < 1.0e^{-7}$) per unit increase in risk score. Patients in the lower quintile had a median onset age of 80 years compared with 74 years in the upper quintile ($P = 1.3e^{-7}$). To evaluate the influence of *APOE*, additional analysis was performed based on Model_{OTH-W}. The combined score of the 21 genes remained significantly correlated with a decreasing age at onset (Pearson $r = -0.12$, $P = .0001$) after exclusion of *APOE*.

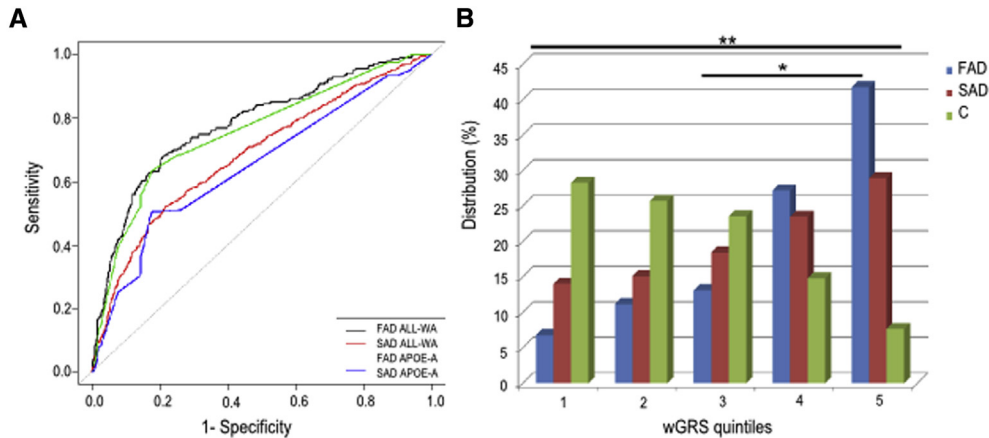


Fig. 2. GRS and family history of AD. (A) ROC curves for familial AD (FAD) patients (defined by at least one first degree relative with dementia, $n = 206$) versus controls and sporadic AD (SAD) patients ($n = 899$) versus controls. ROC curves for FAD based on Model_{ALL-WA} (black curve; AUC = 0.78) and Model_{APOE-A} (green curve; AUC = 0.76); ROC curves for SAD based on Model_{ALL-WA} (red; AUC = 0.68) and Model_{APOE-A} (blue; AUC_{APOE} = 0.65). (B) The percentage of familial AD (blue, $n = 206$), sporadic AD patients (red, $n = 899$), and control subjects (green, $n = 1019$) is shown per quintile. **denotes a P value of $<1.0e^{-15}$ based on Cochran-Armitage test for trend when comparing FAD to SAD patients; *OR, 2.03 (95% CI, 1.27–3.27; P , .004) for positive family history of AD when comparing the upper fifth with the center fifth. Abbreviations: GRS, genetic risk score; AD, Alzheimer's disease; ROC, receiver operating characteristic; AUC, area under the curve; OR, odds ratio; CI, confidence interval.

We observed a significant correlation between the GRS (Model_{ALL-WA}) and CSF $A\beta_{1-42}$ levels in patients but not T-tau or P-tau_{181P} (Fig. 4A–C). The GRS was negatively correlated with $A\beta_{1-42}$ concentrations in CSF (Pearson $r = -0.28$, $P = 8.9e^{-7}$). $A\beta_{1-42}$ decreased with 49.9 pg/mL (SE 10.6) per unit increase on the GRS ($P = 4.0e^{-6}$) and had a median value of 531 pg/mL in the lower quintile and 399.5 pg/mL in the upper quintile ($P = .001$). After exclusion of *APOE* from the GRS, however, this correlation was no longer significant (Pearson $r = -0.071$, $P = .1$). In contrast, the 21-SNP GRS excluding *APOE* (Model_{OTH-W}) appeared correlated with \log_{10} T-tau (Pearson $r = 0.11$, $P = .03$) and \log_{10} P-tau_{181P} (Pearson $r = 0.11$, $P = .04$).

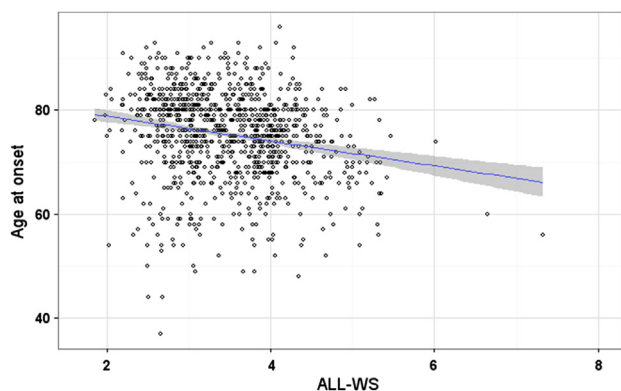


Fig. 3. Negative correlation between GRS and onset age in patients. Scatterplot of onset age (y-axis) and GRS (x-axis). A regression line is presented with 95% CI. The association between onset age and GRS was determined using Model_{ALL-WS} without incorporated age weights. Abbreviations: GRS, genetic risk score; CI, confidence interval.

4. Discussion

In this study, 22 AD susceptibility loci were combined to explore the potential of GRSs in AD. Except for *APOE*, these individual variants only have a marginal contribution toward disease prediction, but when combined, they result in a better risk profile than based on *APOE* alone. When combining genotype data on all 22 variants, best discrimination between patients and control individuals was achieved by a wGRS with age-related weighting of *APOE* genotypes (Model_{ALL-WA}). As expected, wGRS outperformed the cGRS, given that the latter model does not take into account the stronger risks conveyed by *APOE* $\epsilon 4$ and *TREM2* p.R47H [24,25]. There were no notable differences in genetic risk profile in the presence or absence of *APOE* $\epsilon 4$ in our study population.

Parsimony analysis demonstrated the potential of reduced models to have equal predictive value of disease outcome. This can be especially interesting in a clinical context where the aim is to capture as much information as possible, while limiting the cost and labor of genotyping.

It should be noted that determination of parsimonious models may be population-specific, because of differences in allele frequencies, patterns of linkage disequilibrium, and interactions. In addition, in the context of personalized medicine and given the diversity of pathophysiological pathways contributing to AD susceptibility, parsimonious profiles might omit relevant information. The findings of the parsimony analysis are not fully unexpected, given the diminishing effect sizes of loci detected due to increasingly large study cohorts and suggest that there may be limited gain in future performance of the GRS by inclusion of new indirectly associated SNPs with small ORs.

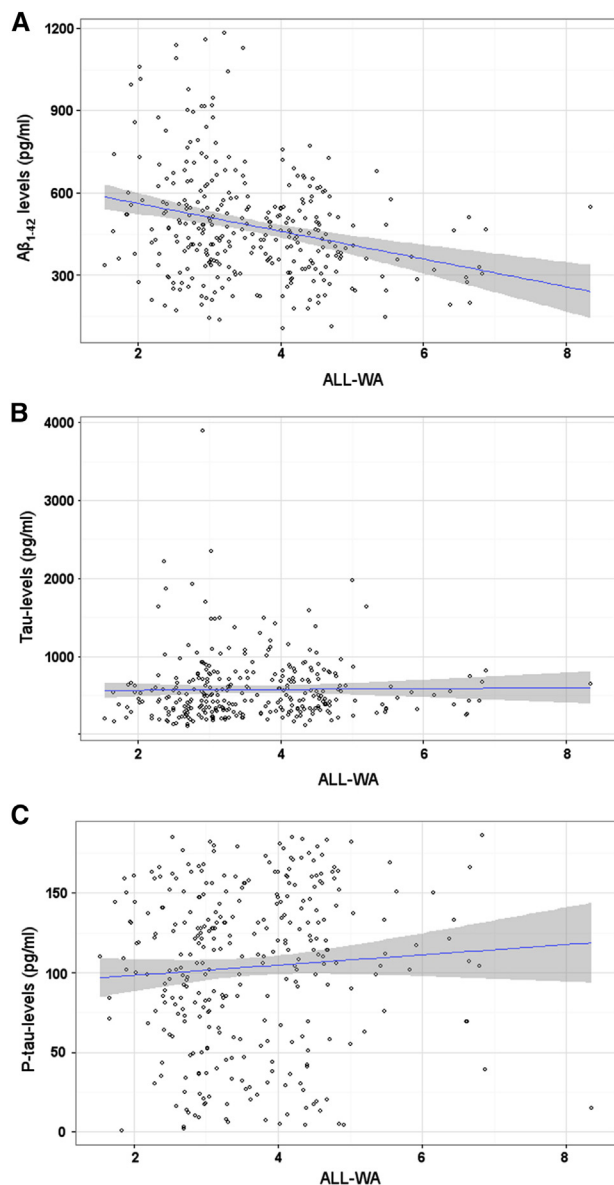


Fig. 4. GRS-CSF biomarker correlations. Scatterplots of GRS Model_{ALL-WA} with patient CSF concentrations of (A) A β ₁₋₄₂, (B) T-tau, and (C) P-tau_{181P}. Regression lines are presented with 95% CI. Abbreviations: GRS, genetic risk score; CSF, cerebrospinal fluid; CI, confidence interval.

The best model had an AUC of 0.70 in our study population, with sensitivity of 55% and specificity of 78% at its point of balanced accuracy [26] and maximum potential effectiveness (Youden index [27]). In comparison, a CSF biomarker panel comprising A β ₁₋₄₂ and T-tau and P-tau_{181P} can discriminate autopsy-confirmed AD from cognitively healthy elderly with a sensitivity of 86% and a specificity of 89% [28]. Particularly in light of the current lack of preventive or therapeutic consequences, these GRSs have limited utility in risk prediction. Nevertheless, the marked increase in risk of AD with increasing GRS, as well as the GRS-endophenotype correlations, identify a potential toward application at the shorter term. Individuals

with the 20% highest scores (score >4.31) were at a substantial, fivefold, increased risk to develop AD compared with individuals with the 20% average scores. Moreover, this group had a significantly lower onset age, even when excluding the known effect of *APOE* ϵ 4 on onset age. The ability to genetically determine this high risk profile may be of value to expedite drug and biomarker trials, which are shifting their focus toward prodromal phases of AD.

CSF A β ₁₋₄₂ decreased with increasing GRS, but this appeared to be driven by *APOE* ϵ 4. This suggests that there may be added benefit from a combined genetic and CSF risk profile, as the GWAS loci appear to capture pathophysiological processes other than A β pathology. The value of more complex risk profiling incorporating additional nongenetic information is demonstrated by the fact that the model including age-specific effects for *APOE* performed better than a model ignoring this information.

Interestingly, the discriminative ability of the GRS was stronger when limiting the analysis to patients with a positive family history. As expected, much of the predictive power came from *APOE*. Nonetheless, adding the 21 susceptibility genes to the risk model significantly improved its discriminative ability. The risk of FAD increased even more strongly with increasing GRS than risk of SAD. This suggests that part of the familial aggregation of AD can be explained by a high combined SNP risk score. We cannot exclude that de novo mutations or recessive causes of disease exist in the SAD cohort which may have diluted the discriminative ability of the GRS in SAD; however, we do not expect this to have a large effect given the most likely rare events of de novo mutation and recessive AD. Future studies are warranted to address the ability of the GRS to predict—in individuals with a positive family history of AD—who eventually becomes affected.

Although not ready for clinical application, GRSs are a promising method to identify individuals at high risk which may facilitate therapeutic or biomarker trials. Future directions include improvement of GRSs with novel functional variants and comparing and combining genetic profiles with CSF or plasma-based biomarkers.

Acknowledgments

Research is funded in part by the Belgian Science Policy Office Interuniversity Attraction Poles program P7/16 (BELSPO, www.belspo.be), the Alzheimer's Disease Research Foundation (SAO-FRA S#13023, <http://alz.org>), a Methusalem Excellence Program initiative of the Flemish Government (EWI, www.ewi-vlaanderen.be), the Research Foundation Flanders (FWO G026710N, www.fwo.be), the Agency for Innovation by Science and Technology, Flanders (IWT 120835, www.iwt.be), the University of Antwerp Research Fund (<http://www.uantwerpen.be/>), the Medical Research Foundation Antwerp, Belgium, the EU FP7 project AgedBrainSYSBIO (<http://ec.europa.eu/research/fp7>), the

Innovative Medicines Initiative (IMI) EMIF 115372 (<http://www.imi.europa.eu/>). E.C. is a PhD fellow of IWT (111477). K.B. is a postdoctoral fellow of the FWO. R.V. is a senior clinical investigator of the FWO.

The authors thank the VIB DMG Genetic Service Facility (<http://www.vibgeneticservicefacility.be/>) and Jill Luyckx for their technical assistance.

BELNEU consortium: Patrick Cras (Institute Born-Bunge, University of Antwerp, Antwerp, Belgium; Antwerp University Hospital, Edegem, Belgium); Dirk Nuytten (Hospital Network Antwerp Stuivenberg, Antwerp, Belgium); Patrick Santens, Jan De Bleecker, Anne Sieben, Bart Dermaut (University Hospital Ghent, Ghent, Belgium); Jan Versijpt, Alex Michotte (University Hospital Brussels, Brussels, Belgium); Olivier Deryck, Bruno Bergmans (AZ Sint-Jan Brugge, Bruges, Belgium); Christiana Willems (Jessa Hospital, Hasselt, Belgium); Adrian Ivanoiu (Saint-Luc University Hospital, Université Catholique de Louvain, Louvain-la-Neuve, Belgium); and Eric Salmon (University of Liege and Memory Clinic, CHU Liege, Liege, Belgium).

Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.jalz.2015.02.013>.

RESEARCH IN CONTEXT

1. Systematic review: We included all independently replicated genetic risk loci from large-scale genome-wide association studies and next generation sequencing studies on AD identified through systematic search of PubMed.
2. Interpretation: A GRS for AD based on current knowledge has limited utility to predict AD, but this study illustrates that incorporation of nongenetic information, such as age-specific effect estimates for *APOE* $\epsilon 4$, can improve the discriminative ability. This study highlights the feasibility to identify individuals at highest susceptibility. This enables genetic preselection to streamline biomarker and drug trials and may aid clinical and personal decision making.
3. Future directions: Identification of novel loci and variants will further increase the discriminative ability, but it may be important to combine a GRS with environmental and biomarker data. For biomarker and drug trials, a genetic risk profile ideally should be pathway-based. This will require pinpointing the culprit genes underlying extended genome-wide association signals.

References

- [1] Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement* 2007;3:186–91.
- [2] Jack CR Jr, Holtzman DM. Biomarker modeling of Alzheimer's disease. *Neuron* 2013;80:1347–58.
- [3] Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-wide association study identifies variants at *CLU* and *PICALM* associated with Alzheimer's disease. *Nat Genet* 2009;41:1088–93.
- [4] Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM, et al. Common variants at *ABCA7*, *MS4A6A/MS4A4E*, *EPHA1*, *CD33* and *CD2AP* are associated with Alzheimer's disease. *Nat Genet* 2011;43:429–35.
- [5] Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, et al. Genome-wide association study identifies variants at *CLU* and *CR1* associated with Alzheimer's disease. *Nat Genet* 2009;41:1094–9.
- [6] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* 2013;45:1452–8.
- [7] Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buross J, et al. Common variants at *MS4A4/MS4A6E*, *CD2AP*, *CD33* and *EPHA1* are associated with late-onset Alzheimer's disease. *Nat Genet* 2011;43:436–41.
- [8] Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* 2010;303:1832–40.
- [9] Rodriguez-Rodriguez E, Sanchez-Juan P, Vazquez-Higuera JL, Mateo I, Pozueta A, Berciano J, et al. Genetic risk score predicting accelerated progression from mild cognitive impairment to Alzheimer's disease. *J Neural Transm* 2013;120:807–12.
- [10] Cuyvers E, Bettens K, Philtjens S, Van Langenhove T, Gijssels I, van der Zee J, et al. Investigating the role of rare heterozygous *TREM2* variants in Alzheimer's disease and frontotemporal dementia. *Neurobiol Aging* 2013;35:726.e11–19.
- [11] Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, et al. *TREM2* variants in Alzheimer's disease. *N Engl J Med* 2013;368:117–27.
- [12] Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, et al. Variant of *TREM2* associated with the risk of Alzheimer's disease. *N Engl J Med* 2013;368:107–16.
- [13] Engelborghs S, Dermaut B, Goeman J, Saerens J, Marien P, Pickut BA, et al. Prospective Belgian study of neurodegenerative and vascular dementia: *APOE* genotype effects. *J Neurol Neurosurg Psychiatry* 2003;74:1148–51.
- [14] Engelborghs S, Dermaut B, Marien P, Symons A, Vloeberghs E, Maertens K, et al. Dose dependent effect of *APOE* epsilon4 on behavioral symptoms in frontal lobe dementia. *Neurobiol Aging* 2006;27:285–92.
- [15] Bettens K, Brouwers N, van Miegroet H, Gil A, Engelborghs S, De Deyn PP, et al. Follow-up study of susceptibility loci for Alzheimer's Disease and onset age identified by genome-wide association. *J Alzheimers Dis* 2009;19:1169–75.
- [16] Bettens K, Brouwers N, Engelborghs S, Lambert J-C, Rogaeva E, Vandenberghe R, et al. Both common variations and rare non-synonymous substitutions and small insertion/deletions in *CLU* are associated with increased Alzheimer risk. *Mol Neurodegener* 2012;7:3.
- [17] Brouwers N, Van Cauwenbergh C, Engelborghs S, Lambert JC, Bettens K, Le Bastard N, et al. Alzheimer risk associated with a copy number variation in the complement receptor 1 increasing C3b/C4b binding sites. *Mol Psychiatry* 2011;17:223–33.

- [18] Van Cauwenberghe C, Bettens K, Engelborghs S, Vandenbulcke M, Van Dongen J, Vermeulen S, et al. CR1 coding variant p.Ser1610Thr in Alzheimer disease and related endophenotypes. *Neurobiol Aging* 2013;34: 2235.e1–6.
- [19] Chapuis J, Hansmannel F, Gistelincq M, Mounier A, Van Cauwenberghe C, Kolen KV, et al. Increased expression of BIN1 mediates Alzheimer genetic risk by modulating tau pathology. *Mol Psychiatry* 2013;18:1225–34.
- [20] Weckx S, Del-Favero J, Rademakers R, Claes L, Cruts M, De Jonghe P, et al. novoSNP, a novel computational tool for sequence variation discovery. *Genome Res* 2005;15:436–42.
- [21] Genin E, Hannequin D, Wallon D, Sleegers K, Hiltunen M, Combarros O, et al. APOE and Alzheimer disease: a major gene with semi-dominant inheritance. *Mol Psychiatry* 2011;16:903–7.
- [22] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;12:77.
- [23] Lemon J. Plotrix: A package in the red light district of R. *R-News* 2006;6:8–12.
- [24] Chen H, Poon A, Yeung C, Helms C, Pons J, Bowcock AM, et al. A genetic risk score combining ten psoriasis risk loci improves disease prediction. *PLoS One* 2011;6:e19454.
- [25] De Jager PL, Chibnik LB, Cui J, Reischl J, Lehr S, Simon KC, et al. Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: A weighted genetic risk score. *Lancet Neurol* 2009;8:1111–9.
- [26] Brodersen KH, Ong CS, Stephan KE, Buhmann JM. The balanced accuracy and its posterior distribution. Istanbul, Turkey: Proceedings of the 20th International Conference on Pattern Recognition; 2010. p. 3121–4.
- [27] Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32–5.
- [28] Engelborghs S, De Vreese K, Van de Castele T, Vanderstichele H, Van Everbroeck B, Cras P, et al. Diagnostic performance of a CSF-biomarker panel in autopsy-confirmed dementia. *Neurobiol Aging* 2008;29:1143–59.

Did you know?

The screenshot shows the homepage of the journal *Alzheimer's & Dementia*. At the top, there is a search bar with a dropdown menu for "This Period" and a "Search" button. Below the search bar, there are links for "Advanced Search", "MEDLINE", "My Account", and "My Saved Searches". A black arrow points to the search bar, and another black arrow points to the "MEDLINE" link. The main content area features a "Current Issue" section for November 2009, Vol. 5, No. 6, with a "Now Included on MEDLINE" badge. Below this, there are sections for "Featured Articles" and "Journal Access". The "Journal Access" section states that full-text articles are available from July 2006 to the present, with access to abstracts being complimentary. A "STAAFT" logo is visible in the bottom left corner, and a "JOIN" button is at the bottom center.

You can save your online searches and get the results by email.

Visit www.alzheimersanddementia.org today!