Evaluation of Plasma Proteomic Data for Alzheimer Disease State Classification and for the Prediction of Progression From Mild Cognitive Impairment to Alzheimer Disease

Daniel A. Llano, MD, PhD,* Viswanath Devanarayan, PhD,† Adam J. Simon, PhD,‡ and The Alzheimer's Disease Neuroimaging Initiative (ADNI)

Abstract: Previous studies that have examined the potential for plasma markers to serve as biomarkers for Alzheimer disease (AD) have studied single analytes and focused on the amyloid-β and τ isoforms and have failed to yield conclusive results. In this study, we performed a multivariate analysis of 146 plasma analytes (the Human DiscoveryMAP v 1.0 from Rules-Based Medicine) in 527 subjects with AD, mild cognitive impairment (MCI), or cognitively normal elderly subjects from the Alzheimer's Disease Neuro-imaging Initiative database. We identified 4 different proteomic signatures, each using 5 to 14 analytes, that differentiate AD from control patients with sensitivity and specificity ranging from 74% to 85%. Five analytes were common to all 4 signatures: apolipo-protein A-II, apolipoprotein E, serum glutamic oxaloacetic trans-aminase, α-1-microglobulin, and brain natriuretic peptide. None of the signatures adequately predicted progression from MCI to AD over a 12- and 24-month period. A new panel of analytes, opti-mized to predict MCI to AD conversion, was able to provide 55% to 60% predictive accuracy. These data suggest that a simple panel of plasma analytes may provide an adjunctive tool to differentiate AD from controls, may provide mechanistic insights to the etiology of AD, but cannot adequately predict MCI to AD conversion.

Key Words: Alzheimer, mild cognitive impairment, ADNI, Rules-Based Medicine, machine learning, biomarker, random forest, proteomic

(*Alzheimer Dis Assoc Disord* 2012;00:000–000)

Alzheimer disease (AD) is the leading neurodegenerative disease of the elderly. It is characterized by synaptic loss, brain atrophy, loss of cholinergic innervation of the cerebral cortex and hippocampus, extracellular deposition of amyloid-β in the form of neuritic plaques, and intra-neuronal accumulation of hyperphosphorylated tau. It is usually a devastating illness that leads to institutionaliza-tion and death in 10 years after diagnosis. AD is typically diagnosed on the basis of clinical criteria—that is, loss of cognitive function and the absence of competing alternative diagnoses.[1] However, clinical criteria may be quite insen-sitive to the early (and potentially more reversible) stages of the illness, and therefore, there has been great enthusiasm for incorporating biomarker data to achieve an earlier diagnosis.[2] The most recently revised research criteria for AD incorporate biomarker evidence, either cerebrospinal fluid (CSF) analysis, fluorodeoxyglucose positron emission tomographic scanning, or volumetric magnetic resonance imaging (MRI) scanning.[3] However, performing such tests in elderly demented patients, who bring with them a host of comorbidities, may not always be practical or cost-effective. Therefore, we have investigated the potential of a blood-based biomarker to distinguish patients with AD from those with mild cognitive impairment (MCI)[4] and normal controls and whether such markers can predict conversion from MCI to AD. Such markers may provide a convenient and cost-effective way to achieve an early diagnosis of AD, or potentially as an outcome measure that can be repeat-edly assessed in a clinical trial.

Several investigators have studied the potential for peripheral biomarkers to diagnose AD. Previous studies have looked at a host of different blood markers, including various species of amyloid-β,[5–12] apolipoprotein E (APOE) isoforms,[13] cytokines,[14–16] and other proteins.[17–19] How-ever, few of these studies have yielded definitive results. For example, several studies show an increase,[20–22] whereas others show no change in various isoforms of plasma amyloid-β in AD patients.[23–27]

The failure of a clear plasma biomarker to emerge from the numerous previous studies may be because a single plasma marker may not exist that can adequately distinguish AD from MCI or controls. There have been previous attempts to use arrays of peripheral biomarkers, with mixed results.[28–32] For example, the initial promising results of an 18-analyte signature developed on 83 patients[33] did not show good diagnostic accuracy on a subsequent separate set of subjects.[30] Most of the previous studies of arrays of plasma markers were either too small to detect differences or did not use powerful enough computational approaches to detect differences in the populations of proteins. We attempted to circumvent these problems by: (1) using a large and well-characterized population of AD, MCI, and control patients [the Alzheimer's Disease Neuroimaging Initiative (ADNI) database]; (2) by analyzing a large and diverse array (> 140) of potential serum markers; and (3) using a wide variety of statistical machine-learning algorithms to examine the ability of combinations of a large number of markers to distinguish between diagnostic categories.

METHODS

Data used in the preparation of this article were obtained from the ADNI database (http://adni.loni.u-cla.edu). The ADNI study was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies, and nonprofit organizations, as a $60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography, or other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness and also lessen the time and cost of clinical trials.

The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California—San Francisco, CA. ADNI is the result of the efforts of many coinvestigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the United States and Canada. The initial goal of ADNI was to recruit 800 adults, aged 55 to 90 years, to participate in the research, approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years, and 200 people with early AD to be followed for 2 years. For up-to-date information, see http://www.adni-info.org.

Patient Population

Subjects were recruited from 59 sites across the United States and Canada. Elderly controls, patients with MCI (defined by Petersen criteria[34]) or AD (defined by NINCDS-ADRDA[1]) were recruited and received a series of clinical, neuropsychological, and biomarker assessments. Patients underwent a battery of repeat assessments every 6 months for up to 36 months. Data used for the analyses presented here comprise data from 109 AD, 360 MCI, and 58 normal subjects. Eligible participants were in the age range of 55 to 90 years, fluent in English or Spanish, and had at least 6 years of education. Participants were enrolled

into 1 of 3 groups: cognitively normal, amnestic MCI, or AD. Apart from these latter disorders, participants could have no other significant neurological disease. Normal individuals were free of memory complaints or depression and had a Mini-Mental State Examination (MMSE) score of 24 to 30 and a Clinical Dementia Rating (CDR) score of 0. MCI individuals would have MMSE scores of 24 to 30 and required a CDR of 0.5 and an informant-verified memory complaint substantiated by abnormal education-adjusted scores on the Wechsler Memory Scale Revised—Logical Memory II. Other cognitive domains and everyday functioning were intact. AD patients could have MMSE scores of 20 to 26 and CDR of 0.5 or 1.0.

Subjects Excluded From the Analysis

The 26 MCI and 2 AD subjects that were not diagnosed with moderate to high confidence by the physicians were excluded. These were subjects that were indicated as "uncertain" or "mild" in the ADNI database. The 18 MCI subjects that were not considered as AD-related, and were indicated as "MCI-other" in the ADNI database, were excluded. Among the rest of the subjects, the 2 AD subjects that were diagnosed as "possible AD" instead of "probable AD" were also excluded from the analysis. This left 109 AD, 360 MCI, and 58 normal subjects in the sample. Of the 360 MCI subjects, 253 subjects stayed in the study for at least 24 months. Among these 253, 101 subjects did not progress to AD, and the remaining 152 progressed to AD.

Plasma Samples

Plasma samples were taken after an overnight fast and at a standard time of day (8 AM) and collected in an EDTA-containing purple-top tube. Samples were analyzed on the Luminex xMAP platform by Myriad Rules-Based Medicine (http://www.rulesbasedmedicine.com). A panel of 190 analytes that are potentially related to a diverse array of human diseases (including non-neurological disease, the Human DiscoveryMAP v 1.0) was quantitated. The Luminex assay involves antibody-based detection amplified using a reporter streptavidin-phycoerythrin conjugate. Fluorescent signal is detected using a flow-based laser apparatus to detect fluorescent polystyrene microspheres, which are loaded with different ratios of 2 spectrally distinct fluorochromes. Analytes are analyzed as several panels and multiplexed, 3 to 24 analytes at a time. Details regarding the assay technology and validation can be found at (http://adni.loni.ucla.edu/wp-content/uploads/2010/11/BC_Plasma_Proteomics_Data_Primer.pdf). Of the 190 potential analytes available, only 146 met ADNI quality control criteria and were analyzed further (listed in Supplemental Table 1, http://links.lww.com/WAD/A34). The other 44 analytes were mostly below the assay detection limit.

Analysis

There were 2 broad objectives in our analyses:
1. To identify the markers which differentiate the AD, MCI, and normal groups.
2. To identify the markers which are predictive of 24-month progression from MCI-AD using baseline data from MCI subjects.

To address each of the above objectives, first a univariate analysis was performed to identify markers that on their own significantly differentiate the disease state groups and the MCI-AD progressors from nonprogressors. This was followed by multivariate predictive modeling to

TABLE 1. Baseline Demographics of AD, MCI, and NL Subjects Are Summarized. Age, Education, and Sex Distribution Are Not Significantly Different (*P* > 0.05) Between the 3 Groups

	AD	MCI	NL
No. Subjects (N)	109	360	58
% E4	67.9	55.0	8.6
% F	42.2	35.2	48.2
Age	74.6 (8)	74.9 (7.4)	75.1 (5.8)
Education (y)	15.1 (3.2)	15.7 (3)	15.6 (2.7)
ADAS-11	18.3 (6.3)	11.6 (4.4)	6.2 (2.8)
MMSE	23.6 (1.9)	27 (1.8)	28.9 (1.2)

AD indicates Alzheimer disease; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; NL, normal.

determine optimal combinations of subsets of markers (signatures) that accurately differentiate AD and normal subjects, as well as optimal signatures that predict the 24-month progression from MCI to AD.

Univariate analysis was carried out using an ANCOVA, with age, education, and sex as covariates to the main outcome of AD, MCI, or normal. Although ApoE4$^+$ status is a well-established marker for clinical status and progression in AD, adjustment for ApoE status was not performed, given the small number of ApoE4$^+$ subjects in the healthy control group (5 subjects of 58) and potential confounding with the disease status. Future studies with larger subject pools may permit an appropriate analysis of the impact of ApoE status on the plasma signatures described herein. If the distribution of the markers was not approximately symmetric, appropriate transformation (eg, log10) was used before applying the ANCOVA. The outliers for each analyte were identified using the Tukey outlier criteria on the studentized residuals from the ANCOVA. On the basis of this criteria, samples that fall below $Q1 - 1.5(Q3 - Q1)$ and above $Q3 + 1.5(Q3 - Q1)$ are considered as outliers, where Q1 and Q3 represent the 25th and 75th percentiles, respectively, of the distribution. Statistical significance of the markers between clinical diagnosis groups was reported in terms of both the false-positive rate (*P*-value) and the more stringent false discovery rate (FDR) (*q*-value).[35,36] FDR is the proportion of false positives (discoveries) among those changes declared significant and is the relevant criteria to use when evaluating significance of large number of analytes as in this data set.

Multivariate analyses for identifying optimal diagnostic and prognostic signatures were carried out separately by first filtering out the most significant markers using a robust version of the Students *t* test. The optimal subsets of markers were derived by first evaluating the relative importance of each marker within the framework of random forests,[37] the partial least squares and bagging,[38] and simulated annealing.[39] Then the optimal size of the subsets was determined on the basis of the number of these most important markers that yielded maximal predictive accuracy. In some of these subsets, constraints were imposed on the pairwise correlations between the markers to be not > 80%. These derived subsets were then used in one of the following classification algorithms: (1) diagonal linear discriminant analysis; (2) random forests; (3) support vector machines with radial kernel; (4) neural network; (5) partial least squares; (6) bagging; and (7) k-nearest neighbor.

The predictive performance of these optimal signatures was evaluated with respect to both internal and

TABLE 2. Plasma Proteomic Markers That Are Statistically Significant on Their Own (*P* < 0.05) Between All Pairs of AD, MCI, and Normal Groups Are Summarized

Markers	AD Mean	AD SD	MCI Mean	MCI SD	NL Mean	NL SD	Fold Change AD/MCI	Fold Change MC/NL	Fold Change AD/NL
A1Micro (µg/mL)	13.47	3.23	12.65	3.81	11.21	2.93	1.07	1.13	1.20
HBELGF (pg/mL)	79.49	46.95	98.62	61.13	58.91	35.24	0.81	1.67	1.35
IgM (mL/mL)	1.80	1.10	2.25	3.80	2.76	1.61	0.80	0.82	0.65
MIP-1a (pg/mL)	163.38	31.26	177.86	55.10	152.07	43.25	0.92	1.17	1.07
PAPPA (mIU/mL)	0.02	0.02	0.03	0.02	0.02	0.01	0.58	1.40	0.81
ANG-2 (ng/mL)	4.28	1.25	4.99	2.13	4.23	1.74	0.86	1.18	1.01
ApoA-IV (µg/mL)	19.06	9.31	20.84	10.03	17.90	12.12	0.91	1.16	1.06
C3 (mg/mL)	1.97	0.38	1.73	0.32	1.84	0.32	1.14	0.94	1.07
FasL (pg/mL)	114.83	395.87	75.79	217.13	81.76	41.16	1.51	0.93	1.40
ILGFBP (ng/mL)	83.82	21.95	116.94	79.89	84.79	42.90	0.72	1.38	0.99
PYY (pg/mL)	159.52	149.91	125.66	93.51	88.29	60.69	1.27	1.42	1.81
SGOT (µg/mL)	2.96	0.81	3.67	2.03	3.64	1.02	0.81	1.01	0.81
TTR (mg/dL)	381.05	115.48	355.81	78.66	430.81	102.27	1.07	0.83	0.88

q < 0.05   *P* < 0.05

Those that are significant at the more stringent criteria of *q* < 0.05 are italicized. Results for all 146 analytes for each pair of AD, MCI, and normal groups are listed in Supplemental Tables 2A–C, respectively. A1Micro indicates α-1-microglobulin; AD, Alzheimer disease; ANG-2, angiopoietin-2; ApoA-IV, apolipoprotein A-IV; C3, complement C3; FasL, Fas ligand; IgM, immunoglobulin M; ILGFBP, insulin-like growth factor–binding protein; MCI, mild cognitive impairment; NL, normal; PAPPA, pregnancy-associated plasma protein A; PYY, peptide YY; SGOT, serum glutamic oxaloacetic transaminase; TTR, transthyretin.
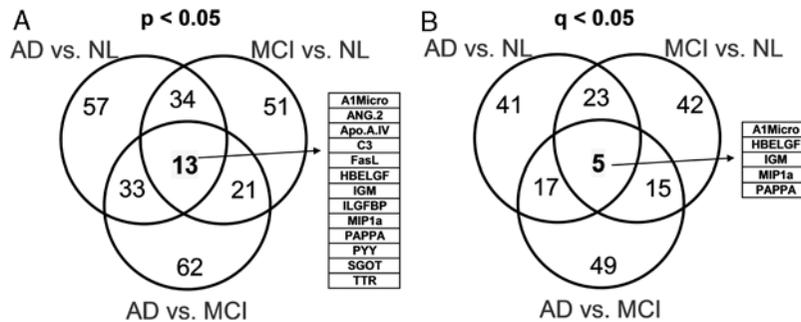
FIGURE 1. A and B, The number of plasma proteomic markers that are significant on their own for different pairs of Alzheimer disease (AD), mild cognitive impairment (MCI), and normal (NL) groups are summarized using Venn diagrams for false-positive rate $P < 0.05$ (A) and false discovery rate $q < 0.05$ (B). The list of markers that are significant for each pair with respect to the P-value and q-value criteria are provided in Supplemental Tables 2A–C (http://links.lww.com/WAD/A35). Those that are significant for all 3 pairs at $q < 0.05$ are summarized in Table 2 and Figure 2. A1Micro indicates α-1-microglobulin; ANG-2, angiopoietin-2; ApoA-IV, apolipoprotein A-IV; C3, complement C3; FasL, Fas ligand; IgM, immunoglobulin M; ILGFBP, insulin-like growth factor–binding protein; PAPPA, pregnancy-associated plasma protein A; PYY, peptide YY; SGOT, serum glutamic oxaloacetic transaminase; TTR, transthyretin.

pseudoexternal validation. First, the original data set were randomly divided into 2 subsets of two-third and one-third size, stratified to ensure similar distribution of the disease groups within these 2 sets. The first two-third training set was used to identify the predictive signatures, with performance estimated using an internal cross-validation procedure, and the second one-third test set (not used to build a model) was used to further test the performance of these signatures (this is referred as "pseudoexternal," as it is not an independent cohort of samples).

The internal cross-validation within the two-third training set to derive and evaluate the optimal signatures was carried out using 10 iterations of a fully embedded 5-fold–stratified cross-validation procedure. This was carried out by first dividing the full two-third training data subset randomly into 5 equal parts, stratified to ensure that each of these parts had the same distribution of the groups (AD and normal, or MCI-AD progressors and nonprogressors) as was found in the original data set. Then each part was left out one at a time (that iteration's test set), and the remaining 4 parts were used as a training set for that iteration to filter the markers, derive the optimal signature, and fit the classification model as described above. The models constructed on the training sets were then used to predict the test sets, and the predictions from all the 5 test sets were pooled together to estimate the performance misclassification measures, sensitivity, and specificity. This entire procedure was iterated 10 times to yield robust estimates of sensitivity and specificity. It is to be noted that all of the steps in the analyses for deriving the optimal signature and fitting the predictive classification model were embedded within this cross-validation procedure. Such a rigorous derivation of optimal signatures is recommended also for high throughput data sets such as genomics,[40] and it provides a more accurate reflection of the true performance in a future cohort. For example, when the cross-validation is done only in the final model-fitting step after the signatures are derived from the entire data set, estimates of sensitivity and specificity will be significantly biased upward.[41,42] Therefore, performance measures reported by different publications should be compared with caution. Please note that the above procedure is the best one can do with only a single data set available, as is the case for the ADNI plasma data.

The performance of the optimal signatures from the internal cross-validation on the training set and the pseudoexternal validation on the test set were summarized in terms of the area under the receiver-operating characteristic curve (AUC), sensitivity, and specificity. Sensitivity refers to the ability to correctly identify AD (for AD vs. normal analysis) or MCI-AD progression (for MCI-AD progression analysis). Specificity refers to the ability to correctly identify normal subjects or subjects who would not progress to dementia of the AD type.

A consideration was given to include more well-studied biomarkers for AD in our predictive models, such as amyloid-based positron emission tomography, fluorodeoxyglucose-based positron emission tomography, CSF amyloid, or tau. However, not all of these markers were tested on all the subjects, and using these as comparators to plasma markers would necessarily truncate the data set. The power of the plasma signatures is that they are derived from the large number of samples. Therefore, this comparison would likely not provide an accurate representation of the utility of these plasma signatures, although this could be of interest for subsequent research.

Analyses were performed using R,[43] version 2.12, and the contributed libraries for the different machine-learning methods were used in our analyses.[44,45]

RESULTS

The key baseline characteristics of the AD, MCI, and controls are summarized in Table 1. The 109 AD, 360 MCI, and 58 normal subjects were not significantly different with respect to age and length of education ($P > 0.05$). The ADAS-Cog and MMSE scores for the AD showed progressive worsening of cognitive function across the spectrum of normal, MCI, and AD. Sex distribution was uneven between the groups ($P < 0.05$). As expected, over 50% of the MCI and AD subjects had 1 or more ApoE4 alleles, and < 9% of the control subjects had at least ApoE4 allele (5 of 58 subjects).

Disease State Comparison (AD vs. MCI vs. Normal): Diagnostic Biomarker

For disease state classification, both univariate and multivariate analyses were performed. The distribution of 141 of 146 analytes was approximately lognormal, that is, these analytes required log transformation before the univariate ANCOVA analysis. The distribution of other
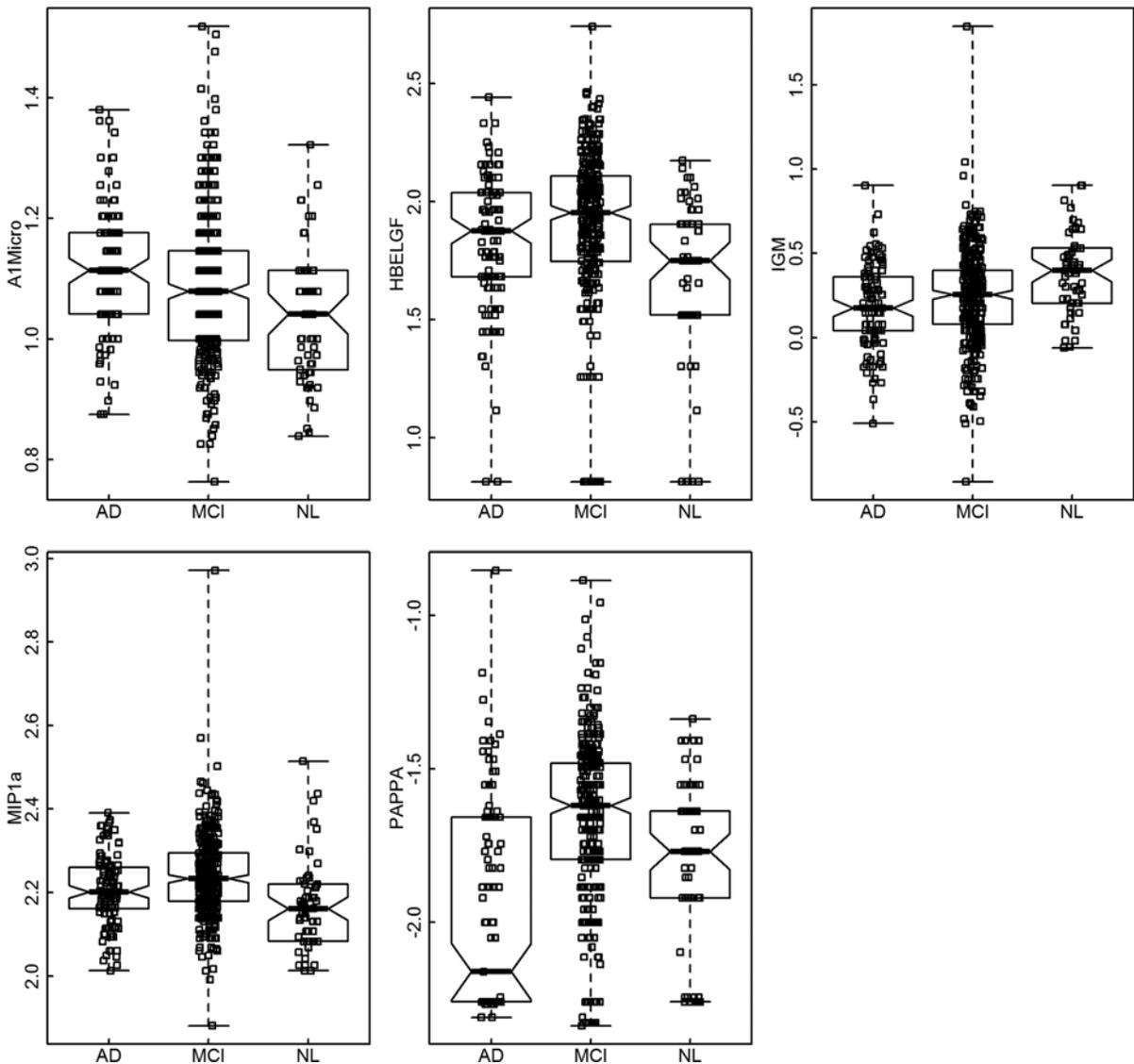
FIGURE 2. Plasma proteomic markers that are statistically significant on their own at the stringent $q < 0.05$ criteria between all pairs of AD, MCI, and normal (NL) groups are graphed here with the individual subject results overlaid on the box plots. Although these markers are statistically significant, none of these show impressive diagnostic accuracy even between AD versus NL. However, a clever optimal combination of some of these and/or other markers may have high predictive value (see text and additional figures).

5 analytes was approximately symmetric in their original scale; hence, these did not require any transformation. Most analytes have a few outliers; hence, these were excluded from the univariate ANCOVA analysis using the criteria stated under the statistical methods.

Univariate analysis demonstrated that 57 of 146 proteins were significantly different between AD versus normal at $P < 0.05$ false-positive rate, out of which 41 were significant at the more stringent $q < 0.05$ FDR criteria. Of the 146 proteins, 51 were significantly different between MCI versus normal at $P < 0.05$, out of which 42 were significant at the more stringent $q < 0.05$ criteria. Of the 146 proteins, 62 were significantly different between AD versus normal at $P < 0.05$, out of which 49 were significant at the more stringent $q < 0.05$ criteria.

Because of the high clinical utility of distinguishing between AD, MCI, and normal, markers were identified

TABLE 3. Four of the optimal signatures that differentiate Alzheimer disease and normal groups are summarized. For each signature, statistical details such as the number of markers prefiltered, the subset derivation method, the subset size, and the classification model

Signature #	Model	No. Prefiltered	Subset Derivation	Signature Size
1	RF	25	PLS.imp	8
2	RF	50	RF.imp	5
3	RF	0	PLS.imp	14
4	RF	25	Bagging.imp	14

RF indicates random forest; PLS.imp, RF.imp, and Bagging.imp refer to the relative importance of the markers in the subset determined with respect to partial least squares, random forests, and bagging methods, respectively.

TABLE 4. Four of the optimal signatures that differentiate Alzheimer disease and normal groups are summarized. The list of markers in each signature.

Sig. #1	Sig. #2	Sig. #3	Sig. #4
A1Micro	A1Micro	A1Micro	A1Micro
ApoA-II	ApoA-II	ApoA-II	ApoA-II
ApoE	ApoE	ApoE	ApoE
BNP	BNP	BNP	BNP
SGOT	SGOT	SGOT	SGOT
Eotaxin-3		Eotaxin-3	A2Macro
IgM		HBELGF	BTC
PLGF		IgM	CRP
		IL-16	Eotaxin-3
		PLGF	IgM
		PYY	IL-16
		TN-C	MPO
		TTR	PLGF
		Vitronectin	RAGE

A1Micro indicates α-1-microglobulin; A2Micro, α-2-macroglobulin; ApoA-II, apolipoprotein A-II; ApoE, apolipoprotein E; BNP, brain natriuretic peptide; BTC, betacellulin; C3, complement C3; CRP, c-reactive protein; IgM, immunoglobulin M; IL-16, interlukin-16; ILGFBP, insulin-like growth factor–binding protein; MPO, myeloperoxidase; PAPPA, pregnancy-associated plasma protein A; PLGF, placenta growth factor; PYY, peptide YY; RAGE, receptor for advanced glycosylation end products; SGOT, serum glutamic oxaloacetic transaminase; TN-C, tenascin-C; TTR, transthyretin.

that differentiated these 3 states. Thirteen proteins were significant for all pairs at $P < 0.05$, of which 5 were significant at the more stringent $q < 0.05$ criteria within a FDR framework (Table 2 and Fig. 1). These 5 markers are: (1) α-1-microglobulin (A1Micro); (2) heparin-binding EGF-like growth factor (HBELGF); (3) immunoglobulin M (IgM); (4) macrophage inflammatory protein-1 α (MIP-1a); and (5) pregnancy-associated plasma protein A (PAPPA). The log-transformed individual data points are shown in Figure 2 for the 5 analytes described above.

Although the group means show significant differences across disease state, there is significant overlap across the diagnostic categories. To assess the ability of each of these markers to distinguish between disease states and to facilitate comparison with other markers, we computed receiver-operating characteristic curve parameters for the comparison between AD and control for each marker. The AUC values for the AD versus control subjects were all < 72%, with reasonably good specificity (∼85%) but poor sensitivity (∼50%). The results for all the 146 analytes for each pair of AD, MCI, and normal groups are listed in

Supplemental Tables 2A–C (http://links.lww.com/WAD/A35), respectively.

The multivariate analysis revealed several optimal diagnostic signatures with good predictive ability to differentiate AD and normal control subjects using the various methods described under the Methods section. The details around 4 of these optimal signatures are summarized in Tables 3–5. The number of markers in these 4 signatures varied from 5 to 14, and the performances of these signatures were examined using both internal and pseudoexternal validation as described in the Methods section. As signatures with fewer markers are generally easier to implement in practice, preference was given to signatures with fewer markers as long as their performance was not significantly inferior to signatures with many markers. Sensitivity, specificity, and AUC of these signatures in the internal and pseudoexternal validation varied mostly between 74% and 85%. The optimal subsets of markers in these signatures were derived on the basis of their relative importance scores with respect to partial least squares, random forests, and bagging algorithms.[44] Random forests algorithm was the optimal classification model for the subsets of markers in the 4 reported signatures. Each of the 19 markers present in one or more of the four signatures graphed with individual subject samples overlaid on the box plots for the AD and the normal groups (Fig. 3). A1Micro, apolipoprotein A-II (ApoA-II), ApoE, brain natriuretic peptide (BNP), and serum glutamic oxaloacetic transaminase (SGOT) are present in all the 4 signatures. Eotaxin-3, IgM, and placenta growth factor are present in 3 of the signatures, and interleukin-16 is present in 2 of the signatures. The relative importance of the markers within the subsets for each of the 4 signatures with respect to the random forests algorithm is summarized in Figures 4A–D. A1Micro, ApoA-II, ApoE, BNP, SGOT, and eotaxin-3 are consistently among the most important markers in these 4 signatures. To further investigate whether SGOT differences might be related to body weight, we examined the correlation between body weight and SGOT levels and found no significant correlation (Spearman rank correlation = −5%, $P = 0.30$).

It should be noted again that these signatures were developed using two thirds of the AD and normal control subjects from this data set. In addition to the internal cross-validation and pseudoexternal validation (on the remaining one third) for evaluating their performance to classify AD versus normal controls (summarized above), these signatures were tested further for their ability to accurately identify the 360 MCI patients not used in this multivariate

TABLE 5. Four of the optimal signatures that differentiate Alzheimer disease and normal groups are summarized. The performance of these signatures in internal validation on training set and external validation on test set

	Internal Cross-Validation						External Validation (Test Set)		
	Sensitivity		Specificity		AUC				
Signature #	Mean (%)	SD (%)	Mean (%)	SD (%)	Mean (%)	SD (%)	Sensitivity (%)	Specificity (%)	AUC (%)
1	80.1	1.6	84.1	3.8	81.5	1.8	78.4	78.9	78.7
2	73.3	3.6	80.5	4.0	75.9	2.7	75.7	73.7	74.7
3	80.8	2.7	86.7	2.9	83.8	2.0	83.8	78.9	81.4
4	78.9	3.1	85.9	2.2	82.4	1.8	86.5	84.2	85.3

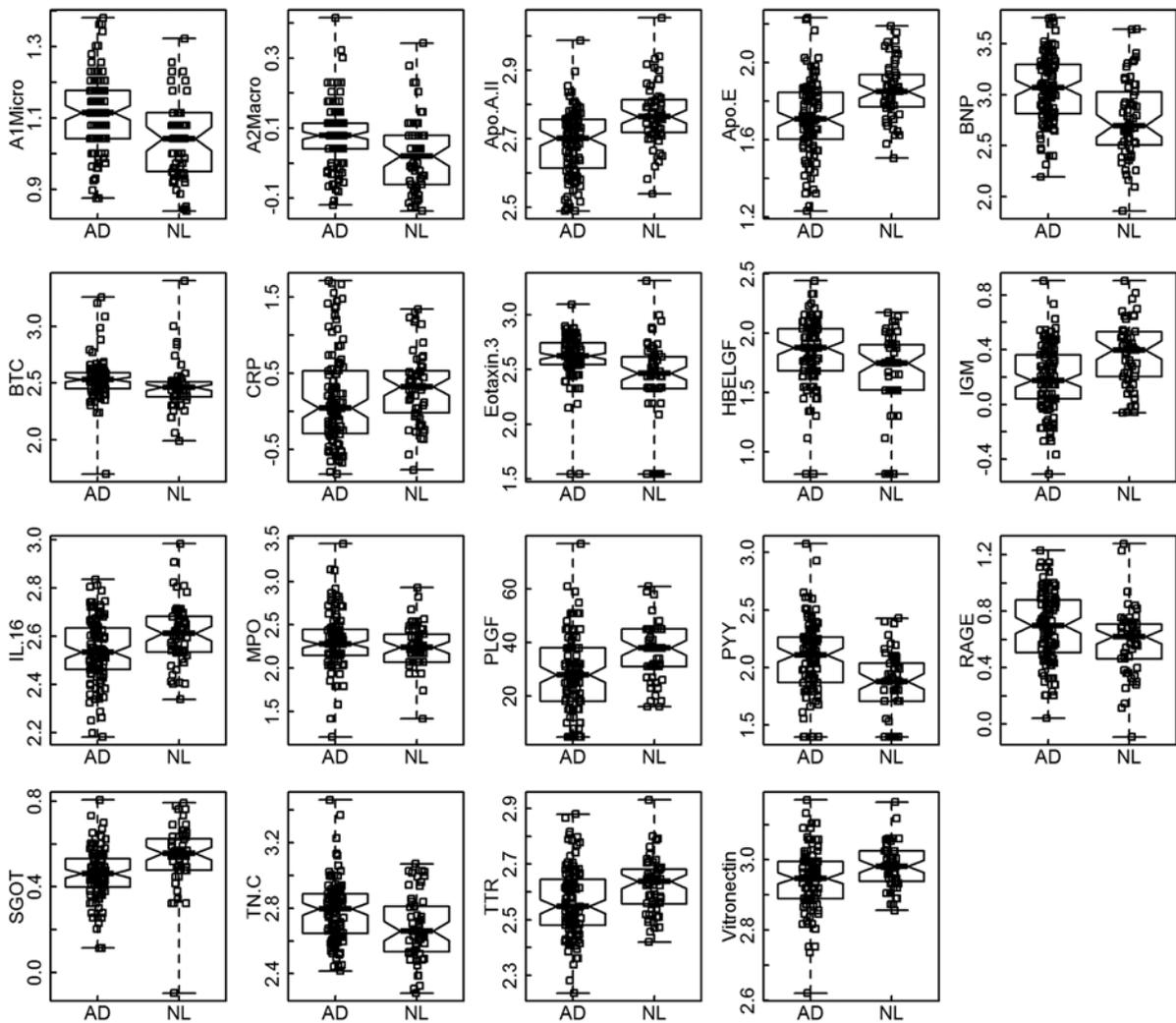AUC indicates area under the receiver-operating characteristic curve.

FIGURE 3. Box plots of each of the 19 markers (log transformed) present in 1 or more of the 4 signatures are provided. Individual subject sample results are overlaid. These graphs reflect that each marker on its own is not a strong diagnostic, but as a collection in one of the 4 signatures proposed here, provides 75% to 85% classification accuracy.

analysis. These 4 signatures correctly identified 65.6%, 73.3%, 68.9%, and 68.1%, respectively, of the 360 MCI patients.

MCI to AD Disease Progression: Prognostic Biomarker

The baseline levels of markers such as plasminogen activator inhibitor-1 (PAI-1) and PAPPA were significantly associated ($P < 0.05$) with 12-month progression, and markers such as ApoA-II and transthyretin (TTR) even appeared significant within a false-positive rate *P*-value approach with 24-month clinical follow-up. However, none of these were significant at the more stringent $q < 0.05$ criteria within a FDR framework. Multivariate analysis yielded signatures with only ∼55% accuracy for predicting a 12- and a 24-month disease progression. This prognostic biomarker performance improved to ∼62% predictive accuracy for patients who were ApoE4 positive (possible as there were sufficient MCI progressor and nonprogressor sample sizes to conduct this subgroup analysis that was not possible with the disease classification data). Canonical

plots for 1 such poor signature in Figure 5 illustrate poor performance in training and test sets.

DISCUSSION

In the current study, we used a large and well-characterized population of AD, MCI, and control participants (the ADNI database) to study the potential of an array of 146 plasma biomarkers to distinguish between these 3 diagnostic categories. A multivariate approach was used to find a "signature" of biomarkers that optimally separate these categories. A univariate analysis showed that several markers showed statistically significant differences between the diagnostic categories. However, none of them showed particularly high sensitivity or specificity when checked for either internal or pseudoexternal cross-validation. In contrast, an array of 5 to 14 markers showed very good diagnostic performance in cross-validation studies. Four independent signatures were identified in different iterations, and 5 of the markers: APOA-II, APOE, BNP, SGOT, and A1micro appeared in all the 4 signatures.
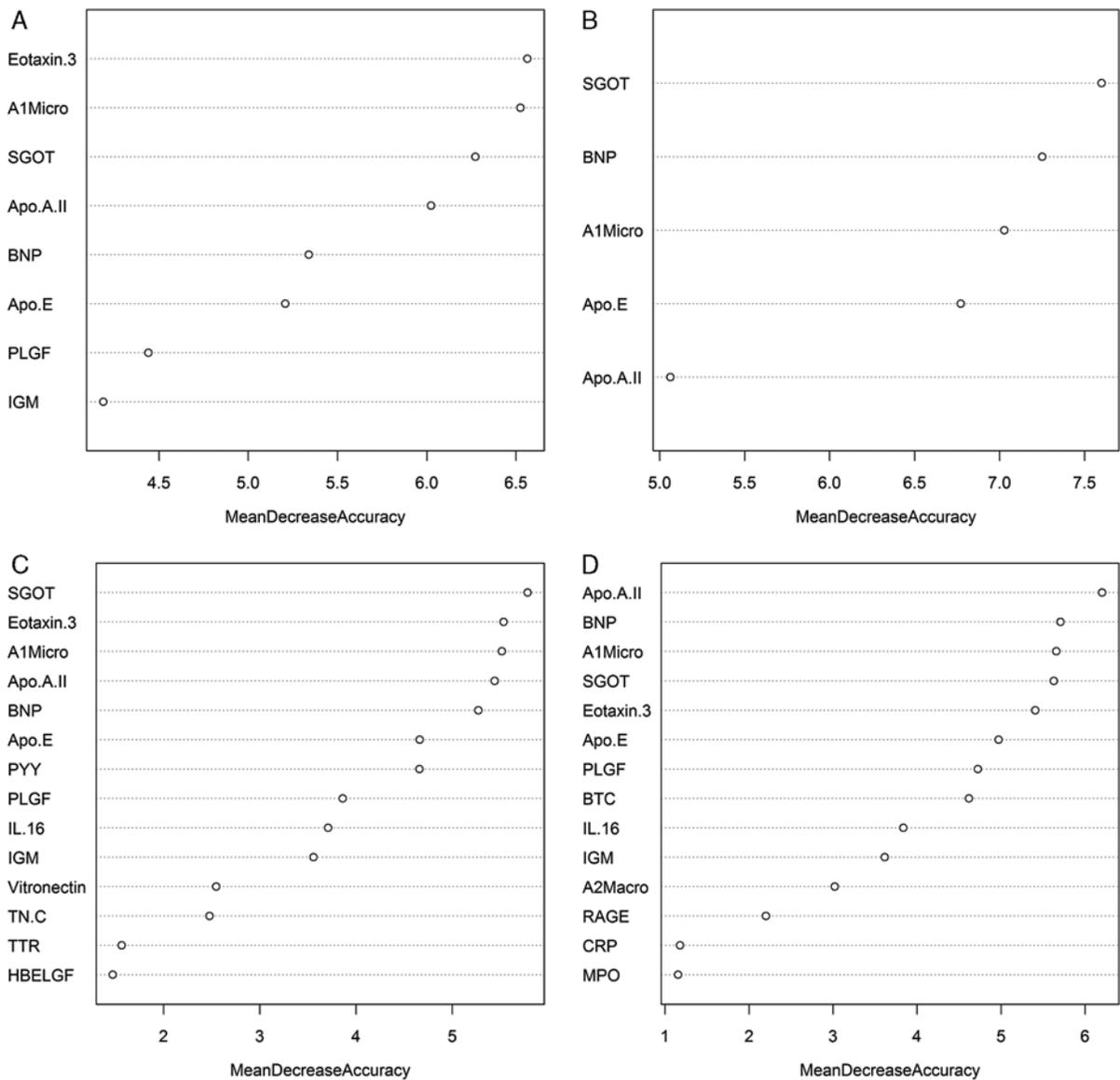
FIGURE 4. A–D, The relative importance of the markers in the first 4 signatures (A–D) with respect to the random forests model is summarized. The relative importance of each marker is determined by the drop in mean predictive accuracy in out-of-bag (hold-out) samples after replacing the concerned marker in the random forests model with a random noise. A1Micro indicates α-1-microglobulin; ApoA-II, apolipoprotein A-II; ApoE, apolipoprotein E; BNP, brain natriuretic peptide; IgM, immunoglobulin M; PLGF, placenta growth factor; PYY, peptide YY; SGOT, serum glutamic oxaloacetic transaminase.

The particular markers identified by the current study are of interest, because several of them (BNP, SGOT, and A1micro) are not classically associated with AD. For example, A1micro, which demonstrated significance in both the univariate and multivariate analysis, is a liver-derived protein whose expression is upregulated during inflammatory states.[46–48] The progressive elevation of A1micro in AD versus MCI versus controls observed in the current study is consistent with the findings that AD is associated with central nervous system inflammation. Although there are several studies demonstrating increases in central nervous system inflammation in AD,[49,50] we are not aware of any previous studies that have assessed the relationship between plasma A1micro and AD. Another marker not

classically associated with AD, BNP, was identified in this study. BNP, although initially identified in brain tissue, is typically associated with myocardial stretch and congestive heart failure. Even though it is not known whether plasma BNP levels in the current study reflect brain or cardiac origin, CSF levels of BNP are significantly lower compared with plasma levels, suggesting that much or most plasma BNP is not brain derived.[51] Another marker identified in this study, SGOT (also known as aspartate transaminase), is typically associated with liver or skeletal muscle injury. SGOT levels were found to be lower in AD patients relative to control or MCI patients and showed significance (uncorrected for multiple comparisons) in the univariate and the multivariate analysis. Similar decreases in the
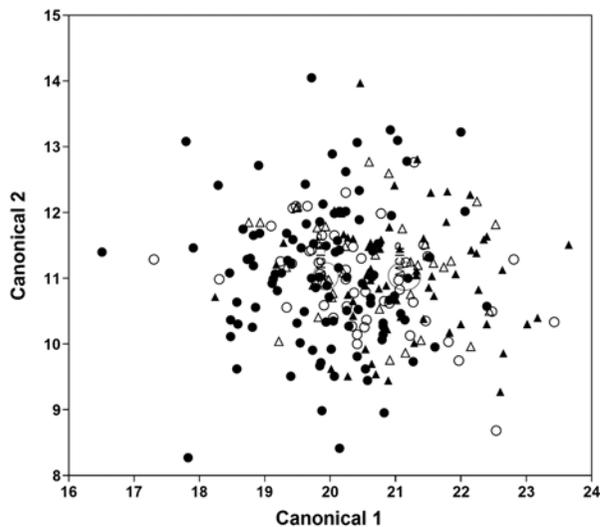
FIGURE 5. A, Canonical plot of a multivariate signature illustrates poor performance in training and test sets. The closed circles and closed triangles represent MCI-AD non-progressors and progressors respectively in the Training set, and that the open circles and open triangles represent non-progressors and progressors respectively in the Test set. A1Micro indicates α-1-microglobulin; ApoA-II, apolipoprotein A-II; ApoE, apolipoprotein E; BNP, brain natriuretic peptide; SGOT, serum glutamic oxaloacetic transaminase.

SGOT levels in AD patients have been seen variably in other studies.[52,53] One potential explanation is that because SGOT levels (as well as other liver function tests) are associated with body mass index,[54] and because AD patients may have a lower body mass index compared with controls,[55] then the lowered SGOT levels in the current study may be related to altered body mass in AD patients. However, we found that SGOT levels in this study did not vary with body weight. Eotaxin-3 is a chemokine usually associated with pulmonary inflammation, although recent findings have shown an association between CSF, eotaxin-3 (in a multivariate analysis[56]), and AD and between peripheral eotaxin-3 and Huntington disease.[57] Interestingly, a related molecule, eotaxin-1, was recently found to increase with age in an animal model, and high levels were found to impair cognitive function when artificially elevated in young animals.[58] The findings that BNP, SGOT, eotaxin-3, and A1micro show novel associations with AD, illustrate the hypothesis-generating potential of a broad-based multivariate approach to biomarker discovery.

Neither a disease state differentiating signature nor an optimized signature for MCI-AD conversion showed high predictive accuracy for MCI-AD conversion over 12 and 24 months. These differ from previous work using a similar approach, but with cognitive, imaging, and CSF data. We previously showed that signatures comprised of only cognitive endpoints, or combinations of biomarkers that were optimized for disease state differentiation, showed good predictive accuracy for MCI-AD conversion at 12 months (AUC values of 65% to 70%[59]). The reason for the differential performance of plasma markers for disease state differentiation and MCI-AD prediction is not known but may have to do with the complex relationship between values of these analytes and disease state. For example, for some markers, there seemed to be a transition state, with values showing their largest deviation from normal in

the MCI state and seemed to "normalize" in AD (eg, HBELGF, MIP-1a, PAPPA, ANG-2, ApoA-IV, ILGFBP, and TTR). The scenario of maximal deviation of a biomarker in early disease with pseudonormalization has been seen with other AD markers, such as hippocampal activation on fMRI,[60] and may represent a compensatory response. Whether the changes seen in these markers also represent compensatory responses will have to be determined in future studies.

The current study illustrates the potential power of a multivariate analytical approach to develop relatively noninvasive and cost-effective diagnostic procedures to assist in the diagnosis of AD, as well as to open new avenues for research into the underlying biology of AD. It does not yet open an avenue for prognostic use but continued work is essential. Future work will hopefully address potential limitations in the present study. For example, it will be critical to test the 5 to 14 analyte plasma signatures against an independent data set to assess their predictive validity. In addition, the current work does not assess these biomarker signatures in other related disease states, such as vascular dementia, Lewy body disease, frontotemporal dementia, Creutzfeldt-Jacob disease, or Parkinson disease. As new disease-modifying therapies with significant adverse effect profiles begin to emerge, it will become imperative for clinicians to differentiate AD from these other dementias and from normal aging. Further evaluation of these multivariate blood-derived signatures identified in the current study in prospective studies will help determine whether these signatures can serve as such a diagnostic tool.

REFERENCES

1. McKhann G, Drachman D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34:939–944.
2. Dubois B, Feldman HH, Jacova C, et al. Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurol*. 2010;9:1118–1127.
3. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7:263–269.
4. Petersen R, Smith G, Waring S, et al. Aging, memory, and mild cognitive impairment. *Int Psychogeriatr*. 1997;9(suppl 1):65–69.
5. Seppälä TT, Herukka SK, Hänninen T, et al. Plasma Aβ42 and Aβ40 as markers of cognitive change in follow-up: a prospective, longitudinal, population-based cohort study. *J Neurol Neurosurg Psychiatry*. 2010;81:1123–1127.
6. Le Bastard N, Leurs J, Blomme W, et al. Plasma amyloid-β forms in Alzheimer's disease and non-Alzheimer's disease patients. *J Alzheimers Dis*. 2010;21:291–301.
7. Lui JK, Laws SM, Li Q-X, et al. Plasma amyloid-β as a biomarker in Alzheimer's disease: the AIBL study of aging. *J Alzheimers Dis*. 2010;20:1233–1242.
8. Lewczuk P, Kornhuber J, Vanmechelen E, et al. Amyloid [beta] peptides in plasma in early diagnosis of Alzheimer's disease: a multicenter study with multiplexing. *Exp Neurol*. 2010;223:366–370.
9. Blennow K, De Meyer G, Hansson O, et al. Evolution of Abeta42 and Abeta40 levels and Abeta42/Abeta40 ratio in plasma during progression of Alzheimer's disease: a multi-center assessment. *J Nutr Health Aging*. 2009;13:205–208.

10. Roher AE, Esh CL, Kokjohn TA, et al. Amyloid beta peptides in human plasma and tissues and their significance for Alzheimer's disease. *Alzheimers Dement.* 2009;5:18–29.

11. Kester MI, Verwey NA, van Elk EJ, et al. Evaluation of plasma A[beta]40 and A[beta]42 as predictors of conversion to Alzheimer's disease in patients with mild cognitive impairment. *Neurobiol Aging.* 2010;31:539–540.

12. Schupf N, Tang MX, Fukuyama H, et al. Peripheral Aβ subspecies as risk biomarkers of Alzheimer's disease. *Proc Natl Acad Sci.* 2008;105:14052–14057.

13. Gupta VB, Laws SM, Villemagne VL, et al. Plasma apolipoprotein E and Alzheimer disease risk. *Neurology.* 2011;76: 1091–1098.

14. Lee KS, Chung JH, Choi TK, et al. Peripheral cytokines and chemokines in Alzheimer's disease. *Dement Geriatr Cogn Disord.* 2009;28:281–287.

15. Kamer AR, Craig RG, Pirraglia E, et al. TNF-[alpha] and antibodies to periodontal bacteria discriminate between Alzheimer's disease patients and normal subjects. *J Neuroimmunol.* 2009;216:92–97.

16. Bermejo P, Martín-Aragón S, Benedí J, et al. Differences of peripheral inflammatory markers between mild cognitive impairment and Alzheimer's disease. *Immunol Lett.* 2008;117: 198–202.

17. Thambisetty M. Do extracellular chaperone proteins in plasma have potential as Alzheimer's disease biomarkers? *Biomarkers Med.* 2010;4:831–834.

18. Akatsu H, Ogawa N, Kanesaka T, et al. Higher activity of peripheral blood angiotensin-converting enzyme is associated with later-onset of Alzheimer's disease. *J Neurol Sci.* 2011; 300:67–73.

19. Güntert A, Campbell J, Saleem M, et al. Plasma gelsolin is decreased and correlates with rate of decline in Alzheimer's disease. *J Alzheimers Dis.* 2010;21:585–596.

20. Mayeux R, Tang M, Jacobs D, et al. Plasma amyloid beta-peptide 1-42 and incipient Alzheimer's disease. *Ann Neurol.* 1999;46:412–416.

21. Mayeux R, Tang HL, Manly MX, et al. Plasma A[beta]40 and A[beta]42 and Alzheimer's disease: relation to age, mortality, and risk. *Neurology.* 2003;61:1185–1190.

22. Mehta PD, Pirttila T, Mehta SP, et al. Plasma and cerebrospinal fluid levels of Amyloid {beta} proteins 1-40 and 1-42 in Alzheimer disease. *Arch Neurol.* 2000;57:100–105.

23. Fukumoto H, Tennis M, Locascio JJ, et al. Age but not diagnosis is the main predictor of plasma Amyloid {beta}-protein levels. *Arch Neurol.* 2003;60:958–964.

24. Kosaka T, Imagawa M, Seki K, et al. The beta APP717 Alzheimer mutation increases the percentage of plasma amyloid-beta protein ending at A beta42(43). *Neurology.* 1997;48:741–745.

25. Scheuner D, Eckman C, Jensen M, et al. Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. *Nat Med.* 1996;2:864–870.

26. Tamaoka A, Fukushima T, Sawamura N, et al. Amyloid beta protein in plasma from patients with sporadic Alzheimer's disease. *J Neurol Sci.* 1996;141:65–68.

27. Vanderstichele H, Van Kerschaver E, Hesse C, et al. Standardization of measurement of beta-amyloid(1-42) in cerebrospinal fluid and plasma. *Amyloid.* 2000;7:245–258.

28. Thambisetty M, Tripaldi R, Riddoch-Contreras J, et al. Proteome-based plasma markers of brain amyloid-β deposition in non-demented older individuals. *J Alzheimers Dis.* 2010;22: 1099–1109.

29. Li N-J, Liu W-T, Li W, et al. Plasma metabolic profiling of Alzheimer's disease by liquid chromatography/mass spectrometry. *Clin Biochem.* 2010;43:992–997.

30. Soares HD, Chen Y, Sabbagh M, et al. Identifying early markers of Alzheimer's disease using quantitative multiplex proteomic immunoassay panels. *Ann NY Acad Sci.* 2009; 1180:56–67.

31. Marksteiner J, Kemmler G, Weiss EM, et al. Five out of 16 plasma signaling proteins are enhanced in plasma of patients with mild cognitive impairment and Alzheimer's disease. *Neurobiol Aging.* 2011;32:539–540.

32. Hye A, Lynham S, Thambisetty M, et al. Proteome-based plasma biomarkers for Alzheimer's disease. *Brain.* 2006;129: 3042–3050.

33. Ray S, Britschgi M, Herbert C, et al. Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. *Nat Med.* 2007;13:1359–1362.

34. Petersen RC, Smith GE, Waring SC, et al. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol.* 1999;56:303–308.

35. Strimmer K. A unified approach to false discovery rate estimation. *BMC Bioinform.* 2008;9:303.

36. Strimmer K. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bionformatics.* 2008;24:1461–1462.

37. Breiman L. Random forests. *Machine Learning.* 2001;45: 5–32.

38. Breiman L. Bagging predictors. *Machine Learning.* 1996;24: 123–140.

39. Cadima J, Cerdeira JO, Minhoto M. Computational aspects of algorithms for variable selection in the context of principal components. *Comput Stat Data Anal.* 2004;47:225–236.

40. Shi L, Campbell G, Jones WD, et al. The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. *Nat Biotech.* 2010;28:827–838.

41. Ambroise C, McLachlan GJ. Selection bias in gene extraction on the basis of microarray gene-expression data. *Proc Natl Acad Sci.* 2002;99:6562–6566.

42. Dupuy A, Simon R. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. *J Natl Cancer Inst.* 2007;99:147–157.

43. R Foundation for Statistical Computing; 2011. Available at: http://www.r-project.org/. Accessed on September 13, 2012.

44. Kuhn M, Wing J, Weston S, et al. *caret: Classification and Regression Training. R package version 4.83*; 2011. Available at: http://CRAN.R-project.org/package = caret/.

45. Liaw A, Wiener M. Classification and regression by Random Forest. *R News.* 2002;2/3:18–22.

46. Daveau M, Jean L, Soury E, et al. Hepatic and extra-hepatic transcription of inter-α-inhibitor family genes under normal or acute inflammatory conditions in rat. *Arch Biochem Biophys.* 1998;350:315–323.

47. Falkenberg C, Blom A, Fries E, et al. Alpha 1-microglobulin and bikunin in rats with collagen II-induced arthritis: plasma levels and livier mRNA content. *Scand J Immunol.* 1997;46: 122–128.

48. Penders J, Delanghe JR. Alpha 1-microglobulin: clinical laboratory aspects and applications. *Clin Chim Acta.* 2004;346: 107–118.

49. Llano D, Li J, Waring J, et al. Cerebrospinal fluid cytokine dynamics differ between Alzheimer's disease patients and elderly controls. *Alzheimer Dis Assoc Disord.* 2011. [Epub ahead of print].

50. Gorelick PB. Role of inflammation in cognitive impairment: results of observational epidemiological studies and clinical trials. *Ann NY Acad Sci.* 2010;1207:155–162.

51. Kondziella D, Göthlin M, Fu M, et al. B-type natriuretic peptide plasma levels are elevated in subcortical vascular dementia. *Neuroreport.* 2009;20:825–827.

52. Yavuz BB, Yavuz B, Halil M, et al. Serum elevated gamma glutamyltransferase levels may be a marker for oxidative stress in Alzheimer's disease. *Int Psychogeriatr.* 2008;20: 815–823.

53. Squitti R, Ventriglia M, Barbati G, et al. 'Free' copper in serum of Alzheimer's disease patients correlates with markers of liver function. *J Neural Transm.* 2007;114:1589–1594.

54. Salvaggio A, Periti M, Miano L, et al. Body mass index and liver enzyme activity in serum. *Clin Chem.* 1991;37:720–723.

55. Doruk H, Naharci MI, Bozoglu E, et al. The relationship between body mass index and incidental mild cognitive impairment, Alzheimer's disease and vascular dementia in elderly. *J Nutr Health Aging*. 2010;14:834–838.

56. Craig-Schapiro R, Kuhn M, Xiong C, et al. Multiplexed immunoassay panel identifies novel CSF biomarkers for Alzheimer's disease diagnosis and prognosis. *PLoS ONE*. 2011;6:e18850.

57. Wild E, Magnusson A, Lahiri N, et al. Abnormal peripheral chemokine profile in Huntington's disease. *PLoS Curr*. 2011;3:RRN1231.

58. Villeda SA, Luo J, Mosher KI, et al. The ageing systemic milieu negatively regulates neurogenesis and cognitive function. *Nature*. 2011;477:90–94.

59. Llano DA, Laforet G, Devanarayan V. Derivation of a new ADAS-cog composite using tree-based multivariate analysis: prediction of conversion from mild cognitive impairment to Alzheimer disease. *Alzheimer Dis Assoc Disord*. 2011;25:73–84.

60. Dickerson B, Salat D, Greve D, et al. Increased hippocampal activation in mild cognitive impairment compared to normal aging and AD. *Neurology*. 2005;65:404–411.