Journal of Alzheimer's Disease 45 (2015) 1157–1173
DOI 10.3233/JAD-142925
IOS Press

A Blood-Based, 7-Metabolite Signature for the Early Diagnosis of Alzheimer's Disease

Javier Olazarán[a], Luis Gil-de-Gómez[b], Andrés Rodríguez-Martín[b], Meritxell Valentí-Soler[c], Belén Frades-Payo[c], Juan Marín-Muñoz[d], Carmen Antúnez[d], Ana Frank-García[e], Carmen Acedo-Jiménez[e], Lorenzo Morlán-Gracia[f], Roberto Petidier-Torregrossa[f], María Concepción Guisasola[g], Félix Bermejo-Pareja[h], Álvaro Sánchez-Ferro[i], David A. Pérez-Martínez[j], Sagrario Manzano-Palomo[j], Ruth Farquhar[k], Alberto Rábano[l] and Miguel Calero[m,*]

[a]*Servicio de Neurología, HGU Gregorio Marañón, Alzheimer Center Reina Sofía Foundation–CIEN Foundation, Madrid, Spain*
[b]*BIOCROSS, S.L., Boecillo, Valladolid, Spain*
[c]*CIEN Foundation, Carlos III Institute of Health, Alzheimer Center Reina Sofia Foundation, Madrid, Spain*
[d]*Hospital Clínico Universitario Virgen de la Arrixaca de Murcia, Unidad de Demencias, El Palmar, Murcia, Spain*
[e]*IdiPaz Hospital Universitario La Paz. Paseo de la Castellana nº 261, Subdireccíon médica, Edificio Maternidad, Madrid, Spain*
[f]*Servicio de Neurología; Servicio de Geriatría, Hospital Universitario de Getafe, Getafe, Madrid, Spain*
[g]*Unidad de Medicina y Cirugia Experimental, Hospital General Universitario Gregorio Marañon, Madrid, Spain*
[h]*Head Neurology Department U. H. "12 de Octubre", Madrid, Spain; CIBERNED, Carlos III National Research Institute, Madrid, Spain; Medicine Department, Complutense University, Madrid, Spain*
[i]*Research Fellow at the Madrid-MIT M+Vision Consortium, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA*
[j]*Hospital Universitario Infanta Cristina, Sección de Neurología, Parla-Madrid, Spain*
[k]*Exploristics, Titanic Suites Belfast, Belfast, UK*
[l]*Departamento de Neuropatología y Banco de Tejidos, Fundación CIEN, Instituto de Salud Carlos III c/ Valderrebollo, Madrid, Spain*
[m]*Chronic Disease Programme, CIBERNED and Alzheimer Disease Research Unit, CIEN Foundation, Alzheimer Center Reina Sofía Foundation, Carlos III Institute of Health, Madrid, Spain*

Accepted 20 January 2015

Abstract. Accurate blood-based biomarkers of Alzheimer's disease (AD) could constitute simple, inexpensive, and non-invasive tools for the early diagnosis and treatment of this devastating neurodegenerative disease. We sought to develop a robust AD biomarker panel by identifying alterations in plasma metabolites that persist throughout the continuum of AD pathophysiology. Using a multicenter, cross-sectional study design, we based our analysis on metabolites whose levels were altered both in AD patients and in patients with amnestic mild cognitive impairment (aMCI), the earliest identifiable stage of AD. UPLC coupled to mass spectrometry was used to independently compare the levels of 495 plasma metabolites in aMCI ($n = 58$) and AD ($n = 100$) patients with those of normal cognition controls (NC, $n = 93$). Metabolite alterations common to both aMCI and AD patients were used to generate a logistic regression model that accurately distinguished AD from NC patients. The final panel consisted of seven metabolites: three amino acids (glutamic acid, alanine, and aspartic acid), one non-esterified fatty acid (22:6n-3, DHA), one bile

*Correspondence to: Miguel Calero, Chronic Disease Programme, CIBERNED and Alzheimer Disease Research Unit, CIEN Foundation, Alzheimer Center Reina Sofía Foundation, Carlos III Institute of Health, Ctra. Majadahonda-Pozuelo Km 2, 28220, Madrid, Spain. Tel.: +34 918223709; E-mail: mcalero@isciii.es.

ISSN 1387-2877/15/$35.00

acid (deoxycholic acid), one phosphatidylethanolamine [PE(36:4)], and one sphingomyelin [SM(39:1)]. Detailed analysis ruled out the influence of potential confounding variables, including comorbidities and treatments, on each of the seven biomarkers. The final model accurately distinguished AD from NC patients (AUC, 0.918). Importantly, the model also distinguished aMCI from NC patients (AUC, 0.826), indicating its potential diagnostic utility in early disease stages. These findings describe a sensitive biomarker panel that may facilitate the specific detection of early-stage AD through the analysis of plasma samples.

Keywords: Alzheimer's disease, biomarkers, diagnosis, mild cognitive impairment, plasma

INTRODUCTION

Alzheimer's disease (AD) is a neurodegenerative disorder characterized by memory loss, cognitive deterioration, and progressive functional dependence, ultimately leading to death. It is by far the most common cause of dementia, affecting 27 million people worldwide, and is predicted to affect 86 million by 2050 [1, 2]. AD is distinguished from other types of dementia by several pathological features, including the progressive appearance of intracellular neurofibrillary tangles of hyperphosphorylated tau protein and neuritic plaques consisting mainly of amyloid-β (Aβ) peptide [3, 4]. While the underlying pathology appears to precede clinical symptoms by decades [5], the associated molecular mechanisms remain largely unknown.

AD is diagnosed using clinical criteria and by excluding other types of dementia. The 2011 revision of the criteria for the clinical diagnosis of AD, established by National Institute on Aging (NIA) and the Alzheimer's Association (AA), incorporates biomarkers into the criteria for the first time [6]. These biomarkers have potential diagnostic utility in each of the three phases of AD pathophysiology described in the NIA/AA guidelines, by supporting the presence of AD in the preclinical phase, before memory loss occurs; determining the underlying cause of clinical impairment in the mild cognitive impairment (MCI) phase; and confirming AD in patients with dementia. MCI is defined as impairment of one or more cognitive domains that is more advanced than expected for a patient's age, but is not accompanied by significant impairment of functional abilities [7]. This definition encompasses cognitive impairment due to a wide range of causes. However, MCI with memory impairment (including amnestic MCI and multidomain MCI) is considered the earliest clinical manifestation of AD; a significant proportion of patients with these symptoms progress to AD within 5 years [8]. Molecular alterations that persist throughout the continuum of AD pathophysiology, from preclinical stages through to the dementia phase, may serve as biomarkers that could aid the early diagnosis of AD, allowing these patients to be treated much earlier than current diagnostic techniques allow.

Cerebrospinal fluid (CSF) is currently the best source of validated AD biomarkers in routine clinical setting. Alterations in CSF levels of Aβ, phosphorylated tau protein (p-tau), and total tau help distinguish AD patients from elderly, cognitively normal controls [9] and predict conversion from MCI to AD [10]. However, the invasive nature of CSF sample collection limits the clinical utility of these markers; CSF can only be collected by lumbar puncture, which can be particularly difficult to perform in elderly patients and precludes the collection of multiple samples over time. Neuroimaging approaches such as structural and functional magnetic resonance imaging (MRI), amyloid tracer imaging, and positron emission tomography (PET), are also effective prognostic tools, but are expensive and often difficult to implement in routine settings [11, 12]. Faced with a growing elderly population, current CSF and neuroimaging techniques are not ideal first-line approaches for screening large numbers of candidate AD patients. There is thus a pressing need to identify new, specific and sensitive biomarkers that can be used to establish diagnosis in preclinical and early clinical AD stages and can distinguish cognitive impairment due to AD from that which accompanies aging and other degenerative conditions.

Blood biomarkers of AD are particularly attractive: samples can be collected easily and inexpensively, and blood is a valid source for repeated measures. Blood-based markers are also more suited to less controlled settings than those detected in other matrices. The last decade has seen the development of several panels of biomarkers in serum or plasma. Proteomics approaches have identified blood-based profiles or signatures that can distinguish between healthy controls, MCI, and AD patients [13–16] and predict conversion from cognitive impairment to prodromal AD [17, 18]. Lipidomics approaches have also been incorporated into the search for AD biomarkers, and have revealed alterations in lipid metabolism pathways and lipid carrier proteins such as ApoE [19]. Major deficits in brain structural glycerophospholipids and sphingolipids have been described in AD patients [20], as

well as alterations in ceramide to sphingomyelin ratios [21] and phosphatidylcholine levels [22]. A recent study reported accurate detection of preclinical AD via lipid analysis [19]. However, most of these findings remain to be replicated in larger, prospective, population-based cohort studies, and to date no blood-based biomarkers have been established or accepted as an aid to diagnosis [23, 24].

In this cross-sectional study, we used a metabolomics approach to detect molecules or species in the blood that were altered in both amnestic MCI (aMCI) and AD patients as compared with a normal cognition (NC) control group. Alterations in metabolite levels were identified by separately comparing aMCI and AD patients with the control group. Metabolites that were significantly altered in both comparisons were used to generate a classification rule that accurately distinguished controls from AD patients. Finally, we investigated the ability of this classification rule to discriminate aMCI patients from controls. The influence of comorbidities, pharmacological treatments, and risk factors for sporadic AD (gender, vascular risk factors, and *APOE* genotype) on the diagnostic efficacy of each blood metabolite was carefully analyzed to rule out potential confounding effects. Taken together, our findings describe a robust panel of plasma biomarkers that accurately discriminates both aMCI and AD patients from healthy controls, and thus may constitute an important tool in the early identification of AD pathophysiology.

MATERIALS AND METHODS

Participants

This multicenter, cross-sectional study sought to identify potential blood markers of early-stage AD. Participants were aged 60 to 85 years and were recruited from six Spanish university hospitals and one Spanish research institution between 18 October 2011 and 3 December 2012. All participants provided written informed consent, and the protocols were in accordance with the Declaration of Helsinki (1975) and approved by the Ethics Committee of the *Instituto de Salud Carlos III*, Spain.

Participants were prospectively recruited. Before inclusion in the study, all participants underwent a medical examination, which included a semi-structured interview, a physical examination, a neurological examination, and a mental-state assessment. In all recruiting centers, the basic protocol for mental-state assessment included the Mini-Mental State Examination (MMSE) test [25], the Blessed Dementia Scale [26], the Hachinski Ischemic Score [27], and the Clinical Dementia Rating (CDR) scale [28]. In the case of participants with normal cognition or aMCI, an additional battery of tests of memory, attention, executive functions, and visuospatial ability was administered by a psychologist. All participants underwent standard blood tests, which included a hemogram and analysis of the levels of glucose, creatinine, transaminases, ions, thyroid stimulating hormone, vitamin B12, and folate. Structural brain imaging (either computerized tomography or MRI) was also performed in all participants. Exclusion criteria included severe hepatic disease, severe kidney disease, disseminated cancer, alcohol or drug abuse, Down syndrome, moderate or severe cranial trauma, and any systemic or acute disease that could compromise completion of the study.

Participants were divided in three groups according to previously described clinical guidelines. Those included in the NC group scored within the normal range for their age and education level (i.e., mean $\pm$ 1.5 standard deviations) in the battery of neuropsychological tests. To ensure that no individuals with early-stage AD were included in this group, participants who scored $\geq$0.5 in the memory box of the CDR were excluded. Participants with MCI were defined using criteria described by Petersen et al. [29] and probable AD patients were defined according to NINCDS-ADRDA criteria [30]. MCI patients were classified as amnestic or non-amnestic MCI. For the purpose of this study, only MCI patients with amnestic MCI (single or multidomain MCI) were recruited (henceforth referred to as the aMCI group). Based on these criteria, 93 individuals were assigned to the NC group, 58 to the aMCI group, and 100 to the AD group.

Blood extraction and determinations

Blood was collected by peripheral venipuncture by trained personnel between 8 a.m. and 10 a.m. after overnight fasting. In all cases, 10-mL samples were collected in EDTA BD Vacutainer® blood collection tubes. To obtain plasma, blood samples were first centrifuged at 2,280 g for 10 min. The supernatant (3 mL) was then transferred to three 1.5-mL tubes and centrifuged for 10 min at 13,000 rpm. The resulting platelet-free plasma (PFP) was stored at -80°C until analysis.

APOE genotyping

Total DNA was isolated from peripheral blood following standard procedures. *APOE* polymorphisms

(rs429358 and rs7412) were determined by Real-Time PCR as previously described [31].

Metabolomics analyses

Metabolomics analyses were performed by OWL Metabolomics (Bizkaia, Spain). Endogenous plasma analytes were analyzed by mass spectrometry coupled to ultra-performance liquid chromatography (UPLC-MS). Samples were analyzed in parallel with a test mixture of standard compounds before and after the entire set of randomized sample injections. Moreover, duplicate samples were injected in order to evaluate the retention time stability (generally <6 s variation, injection-to-injection), mass accuracy and sensitivity of the system throughout the course of the run. The overall quality of the analysis procedure was monitored using 5 repeat extracts of the Quality Control (QC) sample.

To successfully profile a broad concentration range of chemically diverse metabolites, metabolites were extracted by fractionating plasma samples into pools of species with similar physicochemical properties, using appropriate combinations of organic solvents [32, 33]. Three separate UPLC-MS platforms were used to ensure optimal metabolite profiling. A total of 495 molecules were detected and quantified.

Platform 1: UPLC/MS analysis of acylcarnitines, bile acids, fatty acids, lysoglycerophospholipids, and steroids

This platform was used to analyze 210 metabolites belonging to the following categories: acylcarnitines (AC), bile acids (BA), non-esterified fatty acids (NEFAs), oxidized fatty acids, steroids, and choline, ethanolamine, and inositol lysoglycerophospholipids (lysoPC, lysoPE, and lysoPI, respectively). Plasma samples (75 μL) were thawed and proteins precipitated by adding 4 volumes of methanol at room temperature. The methanol used for protein extraction was spiked with the appropriate internal standards, which are not detected in unspiked human plasma extracts using the same method: tryptophan-d5(indole-d5), lysoPC(13:0/0:0), NEFA(19:0), and dihydrocholic acid. After vortexing briefly, the samples were incubated overnight at −20°C. Supernatants (300 μL) were collected after centrifugation at 16,000 × g for 15 min, dried and reconstituted in 75 μL methanol, and were then introduced into the ACQUITY UPLC® system (Waters Corp., Milford, USA) with a 1.0 × 100 mm ACQUITY 1.7-μm C18 BEH column (Waters Corp.) maintained at 40°C.

Samples (2 μL) were injected onto the column at a flow rate of 140 μL/min, for a total run time of 18 min. The following linear elution gradien was used: 100% solvent A (0.05% formic acid in water), to which solvent B (acetonitrile containing 0.05% formic acid) was added incrementally to reach a concentration of 50% B after 2 min, increasing to 100% B over the next 11 min, and returning to the initial composition over the final 5 min. Analysis was performed using the aforementioned UPLC system coupled online to a Waters QTOF Premier™ (Waters Corp.) with electrospray ionization. Capillary and cone voltages were set in negative ion mode at 2800 V and 50 V, respectively. The nebulizer gas was set at a flow rate of 600 L/h and a temperature of 350°C and the cone gas at 30 L/h and a source temperature of 120°C.

Platform 2: UPLC/MS analysis of amino acids

This platform was used to analyze amino acids and their derivatives ($n = 29$). Aliquots (10 μL) from the extracts prepared for Platform 1 were transferred to microtubes and derivatized for amino acid analysis [34]. They were then analyzed using an ACQUITY UPLC® system (Waters Corp.) with a 1.0 × 100 mm ACQUITY 1.7-μm C18 BEH column (Waters Corp.) maintained at 40°C. Samples (2 μL) were injected onto the column at a flow rate of 140 μL/min with a total run time of 14 min. The following elution gradient consisting of solvent A (10 mM ammonium bicarbonate [pH 8.8] in water) and solvent B (acetonitrile) was used: 2% solvent B, increasing linearly to 8% solvent B over 6.5 min, to 20% solvent B over 3.5 min, to 30% solvent B over 1 min, and finally to 100% solvent B over 2 min, before returning to initial composition over the last 2 min. The eluents were then introduced into an Acquity-SQD system (Waters Corp.) in positive ion mode with a capillary voltage of 3200 V and a cone voltage of 30 V. The nebulization gas was set to a flow rate of 600 L/h and a temperature of 350°C, and the cone gas at a flow rate of 10 L/h and a source temperature of 120°C.

Platform 3: UPLC/MS analysis of glycerolipids, cholesterol esters, sphingolipids, and glycerophospholipids

Platform 3 was used to analyze 256 metabolites belonging to the following categories: diacylglyc-

erols (DAG), triacylglycerols (TAG), cholesterol esters (ChoE), sphingomyelins (SM), ceramides (Cer), monohexosyl ceramides (CMH), choline glycerophospholipids (PC), ethanolamine glycerophospholipids (PE), and phosphatidylinositol (PI). For PC and PE species containing an ether moiety, the prefix O- denotes the presence of an alkyl ether substituent, while the suffix e indicates the presence of an ether-linked substituent with one or more double bonds. Plasma extracts (10 μL) were mixed with 10 μL sodium chloride (50 mM) and 110 μL chloroform/methanol (2:1) in 1.5-mL microtubes at room temperature. The extraction solvent was spiked with the following compounds, which are not detected in unspiked human plasma extracts in platform 3: SM(d18:1/6:0), PE(17:0/17:0), PC(19:0/19:0), TAG(13:0/13:0/13:0), TAG(17:0/17:0/17:0), Cer(d18:1/17:0), ChoE(12:0). After vortexing briefly, the samples were incubated for 1 h at −20°C. They were then centrifuged at 16,000 × g for 15 min, 70 μL of the lower organic phase was collected, and the solvent removed. The dried extracts were then reconstituted in 100 μL acetronitrile/isopropanol (50:50), centrifuged (16,000 × g for 5 min) and transferred to vials for analysis. Analysis was performed using the ACQUITY UPLC® system (Waters Corp.) coupled online to a Waters QTOF PremierTM (Waters Corp.) with a 2.1 × 100 mm ACQUITY 1.7 μm C18 BEH column (Waters Corp.) maintained at 60°C. Samples (3 μL) were injected onto the column and eluted at a flow rate of 400 μL/min with a total run time of 17 min. The mobile phase consisted of solvent A (water, acetonitrile, and 10 mM ammonium formate) and solvent B (acetonitrile, isopropanol, and 10 mM ammonium formate) and the following elution gradient was used: 40% solvent B, increasing linearly to 100% over 10 min and returning to the initial composition over 5 min, at which it was maintained for a further 2 min. Mass spectrometry was used in positive ion modes with the capillary current set at 3200 V and the cone voltages at 30 V. The nebulizer gas was set at a flow rate of 1000 L/h and a temperature of 500°C and the cone gas at a flow rate of 30 L/h and a source temperature of 120°C.

Data processing

All data were processed using the TargetLynx application manager for MassLynx 4.1 software (Waters Corp.). A set of predefined retention times, mass-to-charge ratio pairs, and Rt-m/z values corresponding to the metabolites included in the analysis were fed into the program. Associated extracted-ion chromatograms (mass tolerance window = 0.05 Da) were then peak-detected and noise-reduced in both the liquid chromatography (LC) and mass spectrometry (MS) domains to ensure that only true metabolite-related features were processed by the software.

Normalization factors were calculated for each metabolite by dividing their intensities in each sample by the recorded intensity of an appropriate internal standard in the same sample [35]. Linear regression (internal standard-corrected response as a function of sample injection order) was used to detect any intra-batch drift in the QC calibration samples that was not corrected by internal standard correction. For all variables, the internal standard-corrected response in each batch was divided by its corresponding intra-batch drift trend, such that the normalized abundance values of the study samples were expressed with respect to the batch-averaged QC calibration plasma samples (arbitrarily set to 1).

After normalization, the concordance between duplicate sample injection response values was assessed. Where coefficients of variation >30% were found, corresponding sample injection data were returned for manual inspection of the automated integration performed by the TargetLynx software, and modifications made where appropriate. For identified metabolites, representative MS detection-response curves were generated using an internal standard for each chemical class included in the analysis. By assuming similar detector response levels for all metabolites belonging to a given chemical class, a linear detection range was defined for each variable.

Data analysis

Statistical analyses of nominal or categorical variables (gender, *APOE* genotype, comorbidities, and treatments) were performed using Pearson's chi-squared test. Quantitative variables (e.g., age) that deviated from normality were identified using the Kolmogorov-Smirnov statistic with Lilliefors' significance, and subsequently analyzed using the Mann-Whitney U-test.

Univariate analyses were used to compare the levels of the 495 starting metabolites between (i) the AD and NC groups and (ii) the aMCI and NC groups. For each of the two comparisons, a logistic regression model was built for each variable of interest and corresponding receiver operating characteristic (ROC) curves were generated for each of these metabolites. Metabolites that were significantly altered in both comparisons ($p < 0.05$; *t*-test for continuous vari-

ables and Pearson's chi-squared test for categorical variables) were retained for further analysis. These metabolites were then used as independent variables in a logistic regression model to develop a classification rule to distinguish AD patients from healthy controls.

The performance of this classification rule was evaluated using the conservative method of cross-validation. The cross-validations consisted of 100 random samples using a 70% sample for the training dataset, and the remaining 30% for validation. Logistic regression assuming stepwise selection, applying entry criteria of $p < 0.05$ and stay criteria of $p < 0.1$, was used to generate a final model for each iteration of the cross-validation. The number of times a variable was included within the final models across all validations was recorded. Variables found within at least 25% of the derived models were then included within a logistic regression model without selection to determine the final model for classification.

Based upon the derived classification rule, a second round of cross-validation was performed to determine the accuracy of the model in distinguishing the AD from the NC group using ROC curves. This allowed the calculation of the mean, standard deviation, and 95% confidence intervals of the accuracy rates for the AD versus NC comparison. The ability of the final model to distinguish between the aMCI and NC groups was also evaluated.

The potential influence of well-established risk factors for late onset AD such as *APOE* genotype, gender, and vascular risk factors, as well as clinical treatments and comorbidities, was analyzed using Pearson's chi-squared test and logistic regression analysis, respectively.

RESULTS

Demographic and clinical analysis of the study populations

A total of 251 participants fulfilled the inclusion criteria. Based on the criteria described in the Materials and Methods, individuals were assigned to the following groups: NC ($n = 93$), aMCI ($n = 58$), and AD ($n = 100$). The demographic and clinical characteristics of the three study populations are show in Table 1.

For all three groups, age at study inclusion followed a non-normal distribution. Analysis of gender distribution revealed no significant differences in the aMCI ($p = 0.10$) or AD ($p = 0.38$) groups with respect to the NC group. Analysis of age distribution revealed significant differences in AD ($p < 0.001$), but not aMCI ($p = 0.54$) patients, as compared with the NC group. The distribution of *APOE* genotypes in the NC, aMCI, and AD groups is shown in Table 1. *APOE* polymorphic variants at codon 112 and codon 158, as well as the allelic haplotypes $\varepsilon 2/\varepsilon 3/\varepsilon 4$, were in Hardy-Weinberg equilibrium in all three groups. As expected, the *APOE* $\varepsilon 4$ haplotype was over-represented in both aMCI ($p < 0.001$) and AD ($p < 0.001$) patients as compared with controls. No significant differences were observed in the distribution of most comorbidities and risk factors between controls and the aMCI or AD populations, with the exceptions of alcohol consumption and smoking history (Table 1). Drug treatments were also comparably distributed across groups in each of the two comparisons, with the exceptions of the following medications: antihypertensives, neuroleptics, and AD-specific treatments (acetylcholinesterase inhibitors, NMDA receptor antagonists, and neuroleptics) in the AD group; anticoagulants and bronchodilators in the aMCI group; and antidepressants in both the aMCI and AD groups (Table 1).

Metabolite analyses

From one plasma sample taken from each of the 251 study participants, 517 metabolites were either fully or partially identified and subsequently quantified. Of these, 22 were classified as non-identifiable, and were excluded from the analysis. The remaining 495 metabolites comprised the following groups: glycerophospholipids ($n = 238$), glycerolipids ($n = 87$), NEFAs ($n = 48$), sphingolipids ($n = 45$), amino acids ($n = 21$), sterols ($n = 17$), BA ($n = 13$), AC ($n = 11$), amino acid derivatives ($n = 8$), oxidized fatty acids ($n = 7$). A further 35 metabolites were only partially identified, and therefore excluded from the analysis. Before proceeding with group analysis of the remaining 460 metabolites, principal components analysis was performed to rule out potential bias associated with sample provenance or the presence of xenobiotics; no such effect was observed, as evidenced by arbitrary distribution shown in Fig. 1.

Metabolite alterations in aMCI and AD patients

We performed separate univariate analyses to compare the levels of the 460 metabolites in each patient group (aMCI and AD) with those of the NC group. Of these metabolites, 78 were significantly altered in the aMCI group and 100 in the AD group ($p < 0.05$; Fig. 2), as compared with the NC group. We identified

Table 1
Distribution of demographic variables, treatments and comorbidities across the normal cognition (NC), amnestic mild cognitive impairment (aMCI), and Alzheimer's disease (AD) groups. Variables that were non-parametrically distributed (age and Mini-Mental State Examination, MMSE) were identified using the Kolmogorov-Smirnov test and subsequently analyzed using a Mann Whitney U-test. Categorical variables including gender, various comorbidities, and *APOE* genotype were analyzed using Pearson's chi-squared test. Differences were considered significant at $p < 0.05$

Variable/Comorbidity/Treatment		NC ($n = 93$) Value	aMCI ($n = 58$) Value	*p*-value	AD ($n = 100$) Value	*p*-value
Age (years)	Median	74	74	0.538	78	<0.001
	Range [min-max]	[64–86]	[62–86]		[60–90]	
MMSE	Median	29	25	<0.001	17	<0.001
	Range [min-max]	[24–30]	[16–30]		[1–27]	
Gender	% Female	53.8	67.2	0.102	60	0.382
	% Male	46.2	32.8		40	
APOE (allele ε4)	% positive	16.9	41.1	0.001	60	<0.001
	% negative	83.1	58.9		40	
Alcohol consumption history	Yes (%)	18.3	3.4	0.008	2	<0.001
	No (%)	81.7	96.6		98	
Anxiety disorders	Yes (%)	16.1	15.5	0.920	4	0.005
	No (%)	83.9	84.5		96	
Arrhythmia	Yes (%)	7.5	13.8	0.211	6	0.672
	No (%)	92.5	86.2		94	
Depression	Yes (%)	19.4	29.3	0.159	17	0.671
	No (%)	80.6	70.7		83	
Diabetes mellitus	Yes (%)	10.8	10.3	0.937	13	0.630
	No (%)	89.2	89.7		87	
Dyslipidemia	Yes (%)	45.2	51.7	0.432	46	0.907
	No (%)	54.8	48.3		54	
Hypertension	Yes (%)	49.5	51.7	0.787	59	0.184
	No (%)	50.5	48.3		41	
Ictus	Yes (%)	4.3	5.2	0.804	7	0.419
	No (%)	95.7	94.8		93	
Ischemic heart disease	Yes (%)	7.5	10.3	0.548	9	0.711
	No (%)	92.5	89.7		91	
Liver disease	Yes (%)	3.2	5.2	0.551	1	0.278
	No (%)	96.8	94.8		99	
Smoking history	Yes (%)	35.5	8.6	<0.001	7	<0.001
	No (%)	64.5	91.4		93	
Thyroid disease	Yes (%)	18.3	10.3	0.187	9	0.059
	No (%)	81.7	89.7		91	
NSAIDs	Yes (%)	5.4	6.9	0.701	2	0.210
	No (%)	94.6	93.1		98	
Anxiolytics	Yes (%)	19.4	32.8	0.063	16	0.541
	No (%)	80.6	67.2		84	
Antiplatelet agents	Yes (%)	19.4	27.6	0.239	27	0.209
	No (%)	80.6	72.4		73	
Anti-arrhythmic agents	Yes (%)	1.1	5.2	0.127	4	0.201
	No (%)	98.9	94.8		96	
Anticoagulants	Yes (%)	3.2	13.8	0.015	4	0.774
	No (%)	96.8	86.2		96	
Anticholinesterases	Yes (%)	0	1.7	0.204	72	<0.001
	No (%)	100	98.3		28	
Anticonvulsants	Yes (%)	1.1	3.4	0.309	3	0.348
	No (%)	98.9	96.6		97	
Antidepressants	Yes (%)	10.8	29.3	0.004	30	0.001
	No (%)	89.2	70.7		70	
Bronchodilators	Yes (%)	2.2	10.3	0.029	1	0.519
	No (%)	97.8	89.7		99	
Antihypertensives	Yes (%)	47.3	56.9	0.252	62	0.040
	No (%)	52.7	43.1		38	
Memantine	Yes (%)	0	0	NA	28	<0.001
	No (%)	100	100		72	
Neuroleptics	Yes (%)	0	1.7	0.204	8	0.005
	No (%)	100	98.3		92	

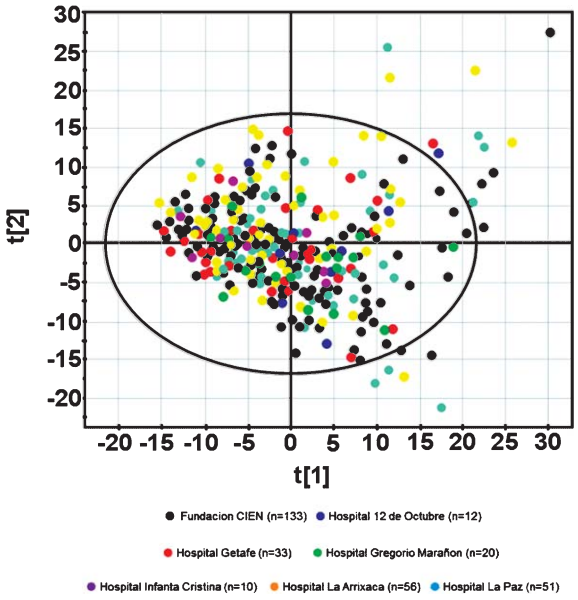

Fig. 1. Principal components analysis of plasma samples from the 251 patients recruited reveals no effect of sample provenance. Samples are color-coded according to the site at which they were collected.

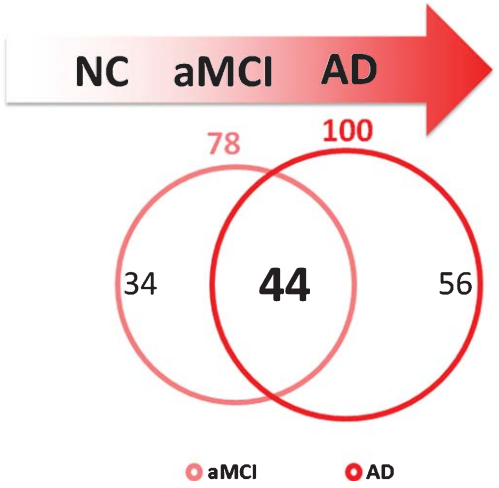

Fig. 2. Venn diagram showing the 44 significantly altered metabolites common to both comparisons (NC versus aMCI and NC versus AD; $p < 0.05$). These metabolites were retained for further analysis and used to develop the final model.

44 significantly altered metabolites that were common to both groups (Fig. 2, Table 2). In all cases the direction of change was the same in both patient groups.

We observed alterations in the levels of a wide range of metabolites, which are implicated in diverse metabolic pathways. The majority were amino acids and lipid metabolites, although changes were also observed in the levels of glucocorticoids, different BA such as deoxycholic acid, glycodeoxycholic acid, and lithocholic acid, and two AC species, AC(10:0) and AC(10:1).

The altered amino acid metabolites constituted a heterogeneous group of compounds with diverse molecular structures and metabolic functions. Levels of acidic amino acids (glutamic acid and aspartic acid) were reduced in AD and aMCI patients with respect to the NC group, while those of glycine, alanine, asparagine, methionine, and arginine were increased (Table 2).

Levels of the vast majority of lipid metabolites identified were dramatically reduced in AD and aMCI patients as compared with healthy controls, suggesting widespread dysregulation of lipid metabolism in AD (Fig. 3). These included one PI species [PI(40:6)], several diacyl PC species [PC(36:5), PC(37:6), PC(38:5), PC(38:6), PC(40:5) and PC(40:6)], one ether-linked PC species [PC(O-36:4)], and ether-linked PE species [PE(38:7e) and PE(40:6e)]. Exceptions to this pattern included diacyl PE species [PE(36:4) and PE(38:5)] and the monoacyl PE species [PE(18:0/0:0) and PE(18:1/0:0)], all of which were detected at higher levels in aMCI and AD patients as compared with healthy controls. However, despite these increases in diacyl PE species, overall levels of PE species, the majority of which were ether-linked, were decreased (Fig. 3).

In line with the observed depletion of most lipid metabolites both patient groups, the levels of several NEFAs were significantly diminished in aMCI and AD patients versus healthy controls. These included one saturated fatty acid (NEFA 16:0, palmitic acid), unsaturated fatty acids (including NEFA 18:1n-9 [oleic acid]) and numerous omega-3 fatty acids, including 18:3n-3 (α-linolenic acid), 20:5n-3 (eicosapentanoic acid; EPA), 22:5n-3 (docosapentanoic acid; DPA), and 22:6n-3 (docosahexanoic acid; DHA) (Fig. 2). aMCI and AD patients also showed marked decreases in the levels of many sphingolipids, including the sphingomyelins SM(39:1), SM(41:1), and SM(42:1) and the ceramides Cer(39:1), Cer(40:1), Cer(41:1), Cer(42:1), and Cer(43:1). The levels of diverse TAG species were significantly reduced in patient groups as compared with controls. Alterations in two specific TAG species, TAG(56:7) and TAG(56:8), were common to both comparisons. In both patient groups, we observed significant increases in two AC species, three BA species, and in the levels of the stress hormone cortisol.

Table 2
List of metabolites significantly altered ($p < 0.05$) in independent comparisons of AD and aMCI patients with controls (NC). The mean and standard deviation (SD) of the levels of each metabolite are shown. Values are expressed relative to batch-averaged quality-control plasma samples (arbitrarily set to 1)

	NC ($n = 93$)		aMCI ($n = 58$)			AD ($n = 100$)		
Metabolite	Mean	SD	Mean	SD	p-value	Mean	SD	p-value
Glutamic acid	0.114	0.046	0.089	0.047	<0.01	0.073	0.029	<0.01
Glycine	0.477	0.142	0.563	0.181	<0.01	0.592	0.233	<0.01
Alanine	0.502	0.133	0.589	0.219	<0.01	0.579	0.173	<0.01
Asparagine	0.748	0.136	0.815	0.171	0.01	0.807	0.147	<0.01
Aspartic acid	0.044	0.016	0.036	0.019	<0.01	0.033	0.012	<0.01
Methionine	0.776	0.147	0.928	0.186	<0.01	0.876	0.190	<0.01
Arginine	0.472	0.094	0.514	0.134	0.04	0.531	0.152	<0.01
AC (10:0)	2.939	1.728	3.934	3.138	0.03	4.252	4.008	<0.01
AC(10:1)	1.473	0.644	1.832	0.852	<0.01	1.825	0.910	<0.01
Deoxycholic acid	0.666	0.738	1.006	1.136	0.05	1.164	1.265	<0.01
Lithocholic acid	1.236	0.807	1.833	1.987	0.03	1.785	1.927	0.01
Glycodeoxycholic acid	0.260	0.373	0.498	0.761	0.03	0.476	0.662	<0.01
Cortisol	1.047	0.378	1.477	0.596	<0.01	1.591	0.534	<0.01
NEFA 16:0	0.737	0.249	0.635	0.220	0.01	0.659	0.288	0.05
NEFA 18:1n-9	1.212	0.370	1.021	0.346	<0.01	1.007	0.415	<0.01
NEFA 20:2n-6	0.783	0.329	0.611	0.261	<0.01	0.664	0.280	<0.01
NEFA 20:3n-3	0.349	0.133	0.289	0.099	<0.01	0.307	0.142	0.04
NEFA 20:5n-3	0.453	0.345	0.328	0.231	<0.01	0.300	0.209	<0.01
NEFA 22:5n-3	0.812	0.491	0.602	0.305	<0.01	0.623	0.392	<0.01
NEFA 22:6n-3	1.741	0.805	1.367	0.730	<0.01	1.349	0.683	<0.01
PC(36:5)	5.082	3.877	3.678	2.886	0.01	3.662	3.426	<0.01
PC(37:6)	4.769	2.429	4.053	1.846	0.04	4.073	1.895	0.03
PC(38:5)	3.510	2.178	2.712	1.487	<0.01	2.780	1.952	0.02
PC(38:6)	3.597	0.749	3.215	0.800	<0.01	3.287	0.849	<0.01
PC(40:5)	1.342	0.495	1.108	0.330	<0.01	1.197	0.409	0.03
PC(40:6)	5.045	1.658	4.398	1.496	0.02	4.338	1.596	<0.01
PC(O-36:4)	1.727	0.590	1.558	0.438	0.05	1.557	0.421	0.02
PE(36:4)	1.521	0.780	1.898	1.149	0.03	1.896	1.077	<0.01
PE(38:5)	1.557	0.981	1.924	1.011	0.03	1.825	0.863	0.05
PE(38:7e)	5.332	1.906	4.522	1.809	0.01	4.216	1.576	<0.01
PE(40:6e)	1.727	0.590	1.558	0.438	<0.01	1.557	0.421	<0.01
LysoPE(18:0/0:0)	0.418	0.124	0.463	0.122	0.03	0.476	0.154	<0.01
LysoPE(18:1/0:0)	0.692	0.301	0.870	0.298	<0.01	0.789	0.293	0.03
PI(40:6)	3.079	2.205	2.104	1.729	<0.01	2.376	1.859	0.02
Cer(39:1)	1.959	0.958	1.512	0.599	<0.01	1.601	0.681	<0.01
Cer(40:1)	1.431	0.563	1.232	0.418	0.01	1.280	0.458	0.04
Cer(41:1)	1.105	0.414	0.971	0.331	0.03	0.958	0.336	<0.01
Cer(42:1)	1.496	0.631	1.256	0.480	<0.01	1.326	0.539	0.05
Cer(43:1)	1.335	0.709	1.070	0.549	0.01	1.008	0.524	<0.01
SM(39:1)	1.446	0.582	1.269	0.421	0.03	1.172	0.410	<0.01
SM(41:1)	1.670	0.533	1.494	0.471	0.04	1.376	0.451	<0.01
SM(42:1)	1.397	0.448	1.231	0.379	0.02	1.224	0.413	<0.01
TAG(56:7)	46.228	34.547	31.007	27.162	<0.01	30.049	23.144	<0.01
TAG(56:8)	36.141	29.406	24.819	29.346	0.02	24.265	19.116	<0.01

Multivariate analysis and development of the diagnostic algorithm

The 44 metabolites that were altered in both aMCI and AD patients as compared with controls (Table 2) were used as independent variables to create a multivariate diagnostic algorithm based on a logistic regression model using cross-validation of data. The model-building process was applied within a series of cross-validations to provide a robust method for the derivation of a model from the available data. First, we developed a classification rule to discriminate AD patients from NC controls. Using logistic regression models we identified the following set of seven metabolites, whose levels most consistently showed the greatest differences between the NC and AD groups: glutamic acid, alanine, aspartic acid, deoxycholic acid, PE(36:4), NEFA 22:6n-3, and SM(39:1) (Table 3). The normalized levels of these metabolites in each of the populations studied are shown in the

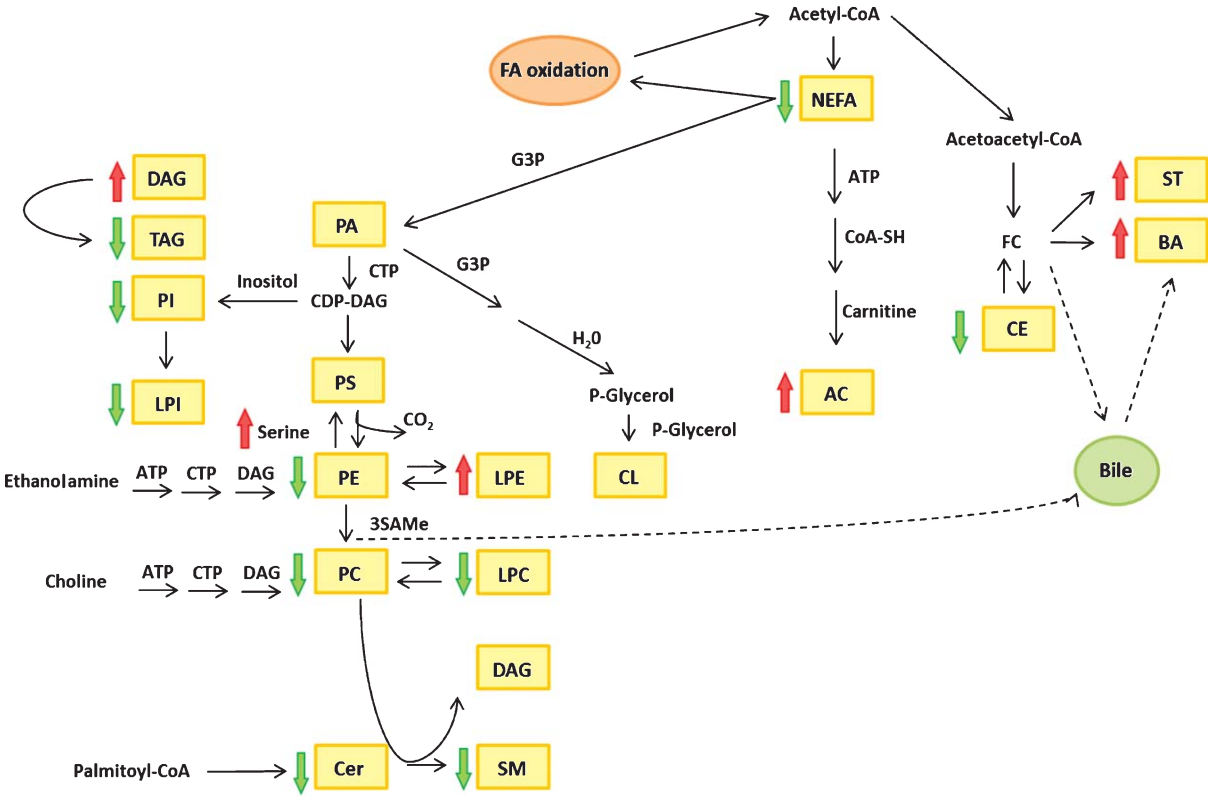

Fig. 3. Lipid biosynthesis pathways: lipid metabolite groups whose levels were altered in AD patients with respect to normal cognition controls are indicated with red (increased) and green (decreased) arrows. AC, acylcarnitines; BA, bile acids; CE, cholesterol esters; CL, cardiolipins; Cer, ceramides; DAG, diacylglycerides; FC, unesterified or free cholesterols; LPC, lysophosphatidylcholine; LPE, lysophosphatidylethanolamine; LPI, lysophosphatidylinositol; NEFA, non-esterified fatty acid, PA, phosphatidic acid; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PI, phosphatidylinositol; PS, phosphatidylserine; SAMe, S-adenosylmethionine; SM, sphingomyelins; ST, steroids; TAG, triacylglycerides.

boxplots in Fig. 4. Based on these findings, a final model consisting of these seven metabolites was generated and cross-validated. To determine the accuracy of the model in distinguishing the AD from the NC group, we generated ROC curves from the second round of cross-validation analysis (Fig. 5). Determining the area under the ROC curve (AUC) provides a measure of the discriminatory power of a diagnostic model. An AUC of 1.0 indicates a perfect test (100% specificity and 100% sensitivity), while lower values indicate less than optimal performance. The model yielded a consistent mean AUC of 0.905 and a mean accuracy (1-Misclassification Rate) of 0.83 across the 100 cross-validations. Moreover, application of the derived model to the full set of AD and NC patients yielded an AUC of 0.918 (Fig. 5). Importantly, the application of this algorithm to the aMCI versus NC comparison yielded a mean AUC of 0.836 and a mean accuracy of 0.80 across the 100 cross-validations. An AUC of 0.826 was achieved when the model was applied to the full set of aMCI and NC patients (Fig. 5).

Potential confounding variables

The potential influence of *APOE* genotype, gender, and age on the diagnostic capacity of each metabolite included in the final model was explored by logistic regression analysis. Comparison of the controlled versus the uncontrolled model revealed no significant changes in the diagnostic capacity of any of the seven metabolites (data not shown). Interestingly, analysis of stratified ROC curves for the AD and NC groups revealed that aspartic acid was more informative in males than females (AUC: 0.183 and 0.366, respectively), while the opposite effect was observed for alanine (AUC: 0.545 and 0.729, respectively) (data not shown).

Table 3
Metabolites included in the final diagnostic algorithm. AUC, Area under curve; SE, Standard error; CL, 95% confidence limits. *AUC, Area under ROC curve. This is a measure of the discriminatory power of a diagnostic test. An AUC of 1.0 indicates a perfect test (100% specificity and 100% sensitivity), while an AUC of 0.5 is considered non-informative

	NC versus aMCI				NC versus AD			
Metabolite	AUC*	SE	Lower CL	Upper CL	AUC	SE	Lower CL	Upper CL
Glutamic acid	0.677	0.048	0.583	0.770	0.785	0.032	0.722	0.848
Alanine	0.614	0.051	0.513	0.716	0.643	0.040	0.565	0.722
Aspartic acid	0.676	0.046	0.585	0.766	0.709	0.037	0.638	0.781
Deoxycholic acid	0.582	0.049	0.486	0.677	0.640	0.040	0.561	0.718
NEFA 22:6n-3	0.656	0.047	0.565	0.748	0.665	0.039	0.589	0.742
PE(36:4)	0.610	0.048	0.517	0.703	0.630	0.040	0.551	0.709
SM(39:1)	0.573	0.047	0.481	0.664	0.638	0.040	0.559	0.717

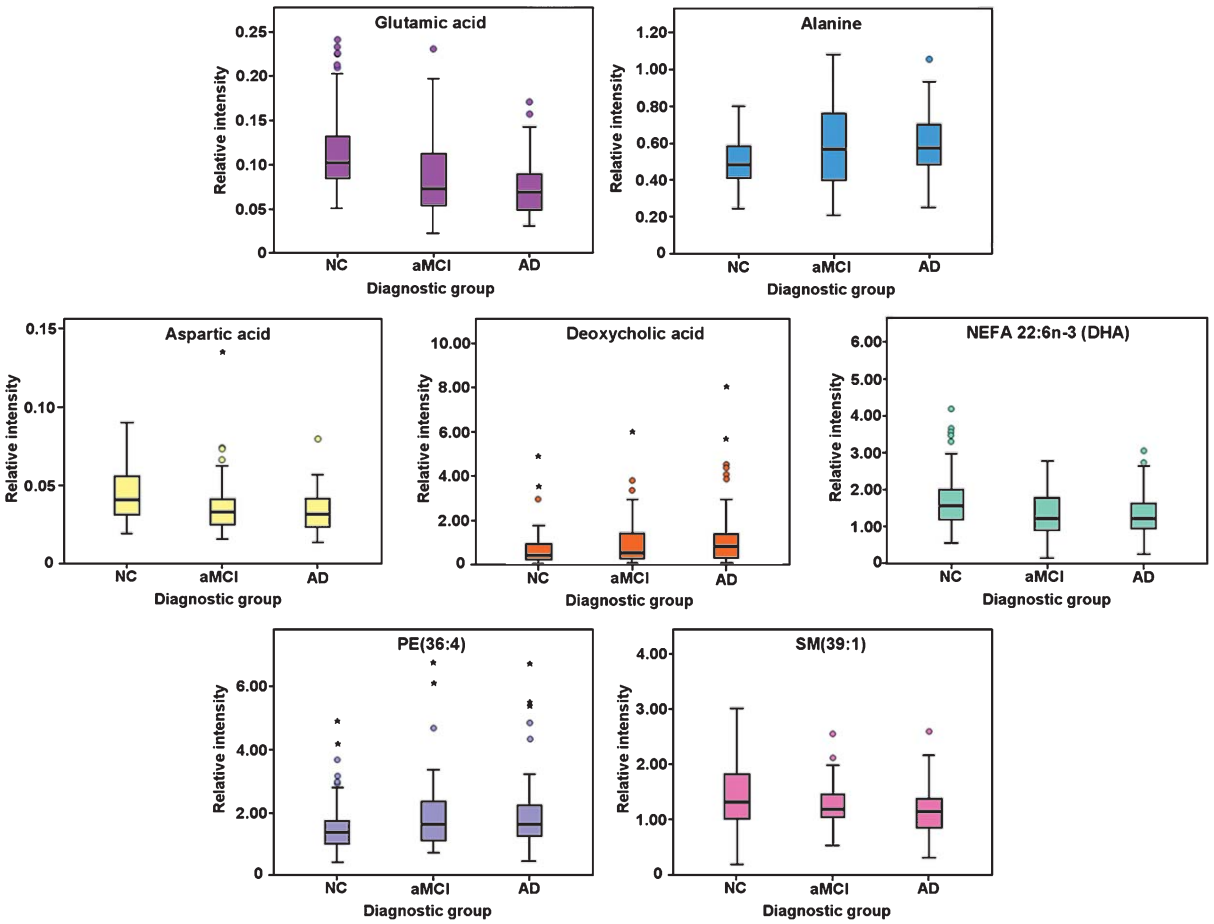

Fig. 4. Box plots show comparative levels of each of the seven metabolites included in the final model. Values are expressed relative to batch-averaged quality-control plasma samples (arbitrarily set at 1). Horizontal lines within each box represent the median of the sample, while the bottom and top of each box represent the first and fourth quartiles. Error bars represent the standard deviation. Outliers are represented as small circles and stars.

We also investigated the potential confounding effects on the final model's seven metabolites of well-established risk factors for late onset AD (e.g., vascular risk factors), as well as clinical treatments and comorbidities (Table 4). Correlations were observed between treatments and certain metabolites included in the final model, mainly in the AD population. SM(39:1) was the only metabolite for which a correlation with a specific treatment (antihypertensives) was observed in both the aMCI and AD groups.

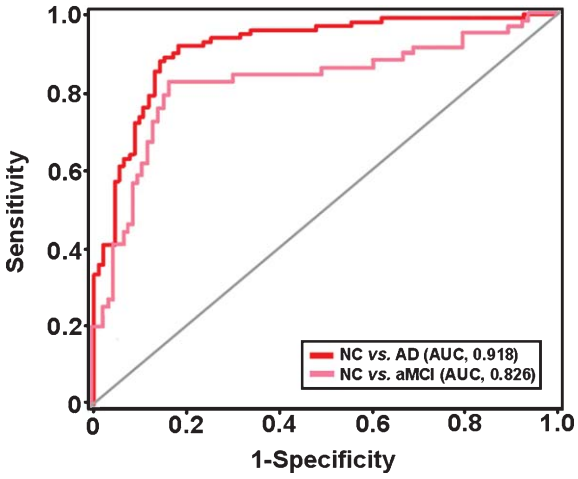

Fig. 5. Performance of the final model when applied to the NC versus AD (AUC, 0.9183) and the NC versus aMCI (AUC, 0.8259) comparisons, based on the full population for each group.

DISCUSSION

In this study, we analyzed the levels of 495 plasma metabolites in NC ($n = 93$), aMCI ($n = 58$), and AD ($n = 100$) subjects using a multicenter, cross-sectional design. By selecting from a group of 44 metabolites whose levels were significantly altered both in aMCI and AD patients, we generated a seven metabolite biomarker panel that accurately distinguished AD from NC subjects (AUC, 0.918). This panel consisted of three amino acids (glutamic acid, alanine, and aspartic acid), one non-esterified fatty acid (22:6n-3, DHA), one bile acid (deoxycholic acid), one phosphatidylethanolamine [PE(36:4)], and one sphingomyelin [SM(39:1)]. Importantly, the panel also accurately distinguished aMCI from NC patients (AUC, 0.826), indicating its potential diagnostic utility in early disease stages. Thus, this signature constitutes a non-invasive tool that may facilitate the rapid, early diagnosis of AD, circumventing the many problems associated with the analysis of CSF, the current diagnostic sample of choice. The study design and the methodology used were carefully considered to ensure the identification of robust and reproducible biomarkers, and procedural and technical details were strictly controlled. A multicenter, cross-sectional design was used. The recruitment procedure included a complete neuropsychological evaluation of both patients and controls, and strict sample procurement procedures were implemented to ensure sample homogeneity. To identify biomarkers of early disease stages, we recruited two independent populations at different disease stages; patients with aMCI (the earliest clinical manifestation of AD) and AD patients with dementia. Each of these groups was individually compared with the NC population. The use of three separate metabolomics platforms allowed the analysis of a wide range of different metabolites, and identified and quantified 495 metabolites in each participant. These metabolites comprised several molecular groups including amino acids, bile acids, acylcarnitines, and several lipid classes such as phospholipids, fatty acids, and sphingolipids. Importantly, variability or clustering associated with sample origin was ruled out by principal component analysis (Fig. 1).

Univariate analysis identified a large number of metabolites ($n = 44$) that were significantly altered in

Table 4
Correlations between treatments and each of the metabolites included the final model ($p < 0.05$ indicates a significant correlation)

	NC			aMCI			AD		
Treatments	Associated biomarker	Correlation coef.	*p*-value	Associated biomarker	Correlation coef.	*p*-value	Associated biomarker	Correlation coef.	*p*-value
NSAIDs	–	–	–	–	–	–	–	–	–
Anxiolytics	–	–	–	–	–	–	–	–	–
Antiplatelet agents	SM(39:1)	–0.31	0.01	Alanine	0.26	0.03	–	–	–
Anti-arrhythmic agents				NEFA 22:6n-3	–0.86	0.01	–	–	–
	NEFA 22:6n-3	1.22	<0.01	Alanine	–0.52	0.03			
				Aspartic acid	–0.55	0.03			
Anticoagulants	–	–	–	–	–	–	PE(36:4)	0.58	0.02
Anticholinesterases	–	–	–	–	–	–	PE(36:4)	–0.28	0.01
Anticonvulsants	Glutamic acid	0.76	0.05	–	–	–	PE(36:4)	0.67	0.02
							SM(39:1)	–0.43	0.05
Antidepressants	–	–	–	–	–	–	SM(39:1)	0.18	0.02
Bronchodilators	–	–	–	–	–	–	–	–	–
Antihypertensives	–	–	–	SM(39:1)	–0.22	0.02	SM(39:1)	–0.19	0.01
				Glutamic acid	0.30	0.03	PE(36:4)	0.22	0.03
Memantine	–	–	–	–	–	–	Alanine	0.22	<0.01
Neuroleptics	–	–	–	Deoxycholic acid	2.42	0.04	–	–	–

both AD and aMCI patients as compared with controls, as expected for a complex, polygenic pathology like AD. Interestingly, levels of all 44 metabolites were altered in the same direction in both patient groups, suggesting that these alterations reflect long-term changes that persist throughout the course of the disease process (Table 2). The selected 44 metabolites were then used as independent variables to develop a logistic regression model based on the AD versus NC comparison. The use of the cross-validation strategy described in the Materials and Methods allowed the generation of a seven-metabolite model that discriminated AD patients from controls, while avoiding overfitting of the data. This biomarker panel consisted of three amino acids, one bile acid, one phospholipid species, one free fatty acid, and one sphingolipid species (Table 3), and distinguished AD patients from controls with an accuracy of over 80% (mean AUC = 0.905, mean accuracy = 0.83). Importantly, this model was also capable of discriminating aMCI patients from healthy controls (mean AUC = 0.836, mean accuracy = 0.80).

Among the 44 metabolites altered in both aMCI and AD patients were seven amino acids; levels of aspartic and glutamic acid were decreased, while those of glycine, alanine, asparagine, methionine, and arginine were increased (Table 2). Three of these amino acids also featured in the final seven-metabolite panel: aspartic and glutamic acids and alanine. The observed decreases in glutamic acid are in line with the well-documented alterations in glutamatergic function that accompany AD [36], and alterations in both blood and CSF levels of alanine have been described in several neurodegenerative processes, including AD [37, 38]. Decreases in CSF levels of aspartic acid have also been reported in AD patients [39], in agreement with the lower levels of this amino acid observed in both our patient groups. Moreover, aspartic acid is an agonist of glutamatergic N-methyl-D-aspartate receptors, which are strongly implicated in the neuropathogenesis of AD [40]. Memantine, a partial NMDA antagonist, stabilizes cognitive decline in patients with moderate-to-severe AD [41]. Importantly, circadian variations of glutamic acid and alanine are unaffected by diet, likely due to their rapid metabolism [42], further supporting their use as biomarkers. A recent study by Mapstone and coworkers also described significant decreases in the plasma levels of several amino acids (proline, lysine, taurine, and phenylalanine) [19]. While none of these amino acids coincided with those of our biomarker panel, this discrepancy may be due to differences in methodology and study design (e.g., the age-range of the patient groups).

aMCI and AD patients showed significant alterations in the levels of several lipid molecules, including AC, NEFAs, phospholipid molecules, Cer, SM, and TAG (Table 2). Three of these lipid metabolites were included in the final 7-metabolite panel: NEFA 22:6n-3, PE(36:4), and SM(39:1) (Table 3). The proper balance of sphingolipids, which include SM and Cer, is essential for normal neuronal function, and several authors have also linked changes in sphingolipid metabolism to the pathophysiology of AD [43, 44]. In accordance with our findings, decreases in plasma SM levels have been reported in AD patients [21]. Apparently conflicting results have been reported regarding blood ceramide levels, with some authors reporting increases in AD versus NC patients [21, 45, 46], and others showing no changes in AD patients but decreases in MCI patients with respect to controls [47]. Diverse studies have proposed that SM and ceramide levels are altered to varying degrees depending on the proximity to the onset of memory impairment, suggesting that these lipids may be useful preclinical markers of the memory impairment that frequently precedes AD [44, 45].

The alterations in plasma phospholipid and NEFA levels observed here are supported by previous findings in both MCI and AD patients [48]. Depleted diacyl PC species were included in a 10-lipid metabolite panel that predicted conversion from cognitive normality to AD [19]. Similar decreases in plasma PC levels [22, 49] and increased CSF levels of PC metabolites [50] have also been previously described in AD patients. In agreement with previous reports [51, 52], we also observed alterations in the levels of PI and PE species in patient groups as compared with healthy controls: levels of diacyl PE species were increased, while those of ether-linked PE species were decreased in aMCI and AD patients. DHA, which featured in the final seven-metabolite panel, accounts for 30–40% of the long chain polyunsaturated fatty acid (LCPUFA) content of the cerebral cortex, and has been implicated in multiple brain functions (e.g., cell membrane fluidity, receptor affinity, modulation of signal transduction molecules). The findings of several epidemiological studies and clinical trials support a role of LCPUFAs, particularly DHA, in AD prevention [53, 54].

aMCI and AD patients also showed significant alterations in the levels of cortisol and bile acids. Given the well-described diurnal fluctuations and inter-individual variability in levels of the stress hormone cortisol, we considered this metabolite unsuitable

as a robust biomarker of AD [55]. The bile acid deoxycholic acid was also included in the final seven-metabolite panel. Bile acids are regulatory molecules with bioactive properties and play an important role in lipid digestion [56]. Moreover, bile acids have been implicated in the etiology of AD [57], and significant increases in the levels of a secondary bile acid have been described in AD patients [19].

The heterogeneity of our final biomarker panel is in keeping with the complexity of AD, in which diverse metabolic pathways are involved, and members of each of these metabolite classes have been previously implicated in the pathophysiology of AD. Thus, the alterations observed at different stages of AD are probably related to long-term metabolic processes, comparable to the mechanisms that underlie other pathologies such as hyperlipidemia and obesity. These results add to a growing body of research indicating that dysregulation of lipid metabolism contributes to the pathophysiology of AD, suggesting that this group of metabolites may include potential biomarkers of AD [20]. Lipids play important roles in myelin and neuronal membrane structure, signal transduction, cholinergic function, AβPP processing, Aβ aggregation, synaptic activity, and neuronal death, indicating that dysregulation of lipid metabolism may contribute to the pathogenesis of AD [20, 43]. Although a general disturbance of lipid metabolism could be caused by the pathophysiological processes underlying AD, the fact that levels of both "unhealthy" and "healthy" lipids were diminished suggests that these specific changes are markers, rather than drivers, of the disease process. While systemic (mainly lipid) metabolic dysregulation is proposed as an important contributor to the induction of AD, peripheral metabolic alterations could also be the consequence of neuronal dysfunction and death due to other causes (e.g., amyloid deposition or neurofibrillary pathologies), even occurring as compensatory responses.

Our seven-biomarker panel discriminated AD patients from the NC group with a comparable accuracy (AUC = 0.918) to that reported for other blood and CSF biomarkers of AD [13, 19, 58–60]. Most importantly, the model was also capable of discriminating aMCI patients from healthy controls, albeit with slightly lower accuracy (AUC = 0.826). This decreased accuracy observed for aMCI patients is unsurprising: aMCI is a heterogeneous syndrome [8], and thus aMCI patients by definition are less precisely classified. While the model's ability to classify aMCI patients could undoubtedly be improved by adjusting coefficients, this would result in unwanted overfitting of the model to our data. Nonetheless, as aMCI and AD can be considered different stages of the same disease, the discriminatory efficacy of the model in the two separate patient groups strongly supports the robustness of the selected biomarker panel, and highlights its potential for the detection of early changes in preclinical disease stages. Interestingly, the observed changes in the levels of most of the seven markers included in the algorithm followed a marked progression from NC to aMCI to AD (see Fig. 4). Detailed analysis of potential modifiers and confounding factors was performed to ensure the robustness of the final biomarker panel. Logistic regression analysis, controlling for *APOE* genotype, gender, and age, was used to study the influence of these variables on the individual metabolites included in our final model. Comparison of the controlled with the uncontrolled model revealed no significant influence of any of these variables. Furthermore, in each of the three study populations, correlation analysis was used to examine the influence of potential confounding variables, including comorbidities and treatments, on the seven biomarkers of the final panel. For certain metabolites in each of the three populations, correlations with certain treatment types were observed. However, as the final model was generated using only metabolites that were significantly altered in both aMCI and AD patients as compared with NC participants, only correlations detected in both patient populations are likely to reflect genuine confounding variables. SM(39:1) was the only metabolite in the final model to show a correlation with any treatment (antihypertensive medications) in both aMCI and AD patients (Table 3). It thus remains unclear whether SM(39:1) is a genuine marker of AD or whether the observed alteration is related to more frequent use of antihypertensives in both patient groups. The latter seems unlikely however, given that differences in the distribution of antihypertensive medications were observed only for AD patients, and not aMCI patients ($p = 0.252$), with respect to the NC group.

This study has several important strengths. Using rigorous recruitment criteria, 251 patients were included from seven different centers, providing the study with an adequate sample size. Moreover, plasma samples were acquired, handled, and stored following strict methodologies to minimize variability. In order to avoid overfitting, all statistical analyses were performed using cross-validation. As aMCI is widely accepted as the earliest clinical manifestation of AD, the inclusion of the aMCI group ensured that we identified metabolites that are consistently altered throughout the continuum of AD pathophysiology. Of

the two patient groups, the AD group was less likely to include other non-AD-related pathologies: accordingly, the development of the final model was based on the NC versus AD comparison. However, internal cross-validation with a separate patient population (aMCI) further supported the diagnostic validity of the model in early stages of AD. Finally, rigorous analysis of the influence of potential confounding variables ensured the development of a robust and accurate biomarker panel.

Our study has some limitations. The first concern is the diagnosis of the patient groups: patient classification based only on clinical criteria resulted in a degree of diagnostic uncertainty, particularly in the case of the aMCI patients, some of whom may not represent early AD. Moreover, the use of clinical criteria meant that we could not rule out potential concomitant pathologies in the patient groups (e.g., vascular and Lewy body pathologies). However, because the aim to this study was to detect early peripheral biomarkers of AD, we did not evaluate the performance of our 7-metabolite signature in distinguishing AD from other dementias. Because we did not compare our diagnostic algorithm with other, "gold standard" biomarkers, the study was self-limited by the precision of the clinical diagnosis. Future studies would thus benefit from a prospective approach, and by analyzing a subset of neuropathologically confirmed cases and controls, or by including CSF biomarkers to increase the accuracy of clinical classification. A second limitation was that the efficacy of our biomarker panel was only tested in plasma, and not in any other matrix (e.g., serum). Finally, given the significant difference in age between the AD and NC populations, it is possible that the observed differences in the levels of some metabolites may reflect normal aging rather than the disease state *per se*. This seems unlikely however, as corresponding alterations were observed in aMCI patients, whose age did not differ significantly from that of the NC group.

The incorporation of biomarkers into AD diagnostic techniques is a key research goal in the field of AD [6]. In particular, blood-based biomarkers show significant promise as simple, inexpensive, non-invasive tools for the diagnosis of AD in early disease stages. Our results describe the accurate and robust discrimination of both aMCI and AD patients from a normal cognition control group using a panel of seven plasma biomarkers. Crucially, these findings will need to be replicated using other analytical platforms and validated in prospective cohort studies in order to establish their clinical utility [61, 62]. Our findings suggest that this biomarker panel constitutes a robust and accurate means of detecting early stage AD through the analysis of plasma samples. Ultimately, the clinical utility of this 7-metabolite signature will need to be established by replicating this study in different populations, and by analyzing its performance in the differential diagnosis of AD from other types of dementias, such as vascular and frontotemporal dementias.

ACKNOWLEDGMENTS

The authors thank all the patients, hospitals, and institutions that made this study possible. We also thank María Jose Gil Moreno for writing the study protocol, Owen Howard for editing assistance, Aiden Flynn for statistical support, and Carlo Zanotti for his decisive role throughout the course of this project. This work was supported by Biocross, SL., which participated in the design, acquisition of data, statistical analysis, study supervision, approval of data, and drafting of the manuscript.

Authors' disclosures available online (http://j-alz.com/manuscript-disclosures/14-2925r1).

REFERENCES

[1] Alzheimer's Association (2012) 2012 Alzheimer's disease facts and figures. *Alzheimers Dement* **8**, 131-168.

[2] Prince MJ, Acosta D, Castro-Costa E, Jackson J, Shaji KS (2009) Packages of care for dementia in low- and middle-income countries. *PLoS Med* **6**, e1000176.

[3] Bossy-Wetzel E, Schwarzenbacher R, Lipton SA (2004) Molecular pathways to neurodegeneration. *Nat Med* **10**(Suppl), S2-S9.

[4] Selkoe DJ (2003) Folding proteins in fatal ways. *Nature* **426**, 900-904.

[5] Holtzman DM (2011) CSF biomarkers for Alzheimer's disease: Current utility and potential future use. *Neurobiol Aging* **32**(Suppl 1), S4-S9.

[6] Jack CR Jr, Albert MS, Knopman DS, McKhann GM, Sperling RA, Carrillo MC, Thies B, Phelps CH (2011) Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 257-262.

[7] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH (2011) The diagnosis of mild cognitive impairment 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* **7**, 270-279.

[8] Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L, Winblad B (2001) Current concepts in mild cognitive impairment. *Arch Neurol* **58**, 1985-1992.

[9] Fagan AM, Mintun MA, Shah AR, Aldea P, Roe CM, Mach RH, Marcus D, Morris JC, Holtzman DM (2009) Cerebrospinal fluid tau and ptau(181) increase with cortical amyloid deposition in cognitively normal individuals: Implications for future clinical trials of Alzheimer's disease. *EMBO Mol Med* **1**, 371-380.

[10] Shaw LM, Vanderstichele H, Knapik-Czajka M, Clark CM, Aisen PS, Petersen RC, Blennow K, Soares H, Simon A, Lewczuk P, Dean R, Siemers E, Potter W, Lee VM, Trojanowski JQ (2009) Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann Neurol* **65**, 403-413.

[11] Frank RA, Galasko D, Hampel H, Hardy J, de Leon MJ, Mehta PD, Rogers J, Siemers E, Trojanowski JQ (2003) Biological markers for therapeutic trials in Alzheimer's disease. Proceedings of the biological markers working group; NIA initiative on neuroimaging in Alzheimer's disease. *Neurobiol Aging* **24**, 521-536.

[12] Schneider P, Hampel H, Buerger K (2009) Biological marker candidates of Alzheimer's disease in blood, plasma, and serum. *CNS Neurosci Ther* **15**, 358-374.

[13] Doecke JD, Laws SM, Faux NG, Wilson W, Burnham SC, Lam CP, Mondal A, Bedo J, Bush AI, Brown B, De Ruyck K, Ellis KA, Fowler C, Gupta VB, Head R, Macaulay SL, Pertile K, Rowe CC, Rembach A, Rodrigues M, Rumble R, Szoeke C, Taddei K, Taddei T, Trounson B, Ames D, Masters CL, Martins RN, Alzheimer's Disease Neuroimaging I, Australian Imaging B, Lifestyle Research G (2012) Blood-based protein biomarkers for diagnosis of Alzheimer disease. *Arch Neurol* **69**, 1318-1325.

[14] O'Bryant SE, Xiao G, Barber R, Huebinger R, Wilhelmsen K, Edwards M, Graff-Radford N, Doody R, Diaz-Arrastia R, Texas Alzheimer's R, Care C, Alzheimer's Disease Neuroimaging I (2011) A blood-based screening tool for Alzheimer's disease that spans serum and plasma: Findings from TARC and ADNI. *PLoS One* **6**, e28092.

[15] Hu WT, Holtzman DM, Fagan AM, Shaw LM, Perrin R, Arnold SE, Grossman M, Xiong C, Craig-Schapiro R, Clark CM, Pickering E, Kuhn M, Chen Y, Van Deerlin VM, McCluskey L, Elman L, Karlawish J, Chen-Plotkin A, Hurtig HI, Siderowf A, Swenson F, Lee VM, Morris JC, Trojanowski JQ, Soares H (2012) Plasma multianalyte profiling in mild cognitive impairment and Alzheimer disease. *Neurology* **79**, 897-905.

[16] Llano DA, Laforet G, Devanarayan V (2011) 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* **25**, 73-84.

[17] Ray S, Britschgi M, Herbert C, Takeda-Uchimura Y, Boxer A, Blennow K, Friedman LF, Galasko DR, Jutel M, Karydas A, Kaye JA, Leszek J, Miller BL, Minthon L, Quinn JF, Rabinovici GD, Robinson WH, Sabbagh MN, So YT, Sparks DL, Tabaton M, Tinklenberg J, Yesavage JA, Tibshirani R, Wyss-Coray T (2007) Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. *Nat Med* **13**, 1359-1362.

[18] Hye A, Riddoch-Contreras J, Baird AL, Ashton NJ, Bazenet C, Leung R, Westman E, Simmons A, Dobson R, Sattlecker M, Lupton M, Lunnon K, Keohane A, Ward M, Pike I, Zucht HD, Pepin D, Zheng W, Tunnicliffe A, Richardson J, Gauthier S, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S (2014) Plasma proteins predict conversion to dementia from prodromal disease. *Alzheimers Dement* **10**, 799-807.e2.

[19] Mapstone M, Cheema AK, Fiandaca MS, Zhong X, Mhyre TR, MacArthur LH, Hall WJ, Fisher SG, Peterson DR, Haley JM, Nazar MD, Rich SA, Berlau DJ, Peltz CB, Tan MT, Kawas CH, Federoff HJ (2014) Plasma phospholipids identify antecedent memory impairment in older adults. *Nat Med* **20**, 415-418.

[20] Wood PL (2012) Lipidomics of Alzheimer's disease: Current status. *Alzheimers Res Ther* **4**, 5.

[21] Han X, Rozen S, Boyle SH, Hellegers C, Cheng H, Burke JR, Welsh-Bohmer KA, Doraiswamy PM, Kaddurah-Daouk R (2011) Metabolomics in early Alzheimer's disease: Identification of altered plasma sphingolipidome using shotgun lipidomics. *PLos One* **6**, :e21643.

[22] Whiley L, Sen A, Heaton J, Proitsi P, Garcia-Gomez D, Leung R, Smith N, Thambisetty M, Kloszewska I, Mecocci P, Soininen H, Tsolaki M, Vellas B, Lovestone S, Legido-Quigley C (2014) Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. *Neurobiol Aging* **35**, 271-278.

[23] Thambisetty M, Lovestone S (2010) Blood-based biomarkers of Alzheimer's disease: Challenging but feasible. *Biomark Med* **4**, 65-79.

[24] Galasko D, Golde TE (2013) Biomarkers for Alzheimer's disease in plasma, serum and blood - conceptual and practical problems. *Alzheimers Res Ther* **5**, 10.

[25] Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**, 189-198.

[26] Blessed G, Tomlinson BE, Roth M (1968) The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br J Psychiatry* **114**, 797-811.

[27] Hachinski VC, Iliff LD, Zilhka E, Du Boulay GH, McAllister VL, Marshall J, Russell RW, Symon L (1975) Cerebral blood flow in dementia. *Arch Neurol* **32**, 632-637.

[28] Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL (1982) A new clinical scale for the staging of dementia. *Br J Psychiatry* **140**, 566-572.

[29] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E (1999) Mild cognitive impairment: Clinical characterization and outcome. *Arch Neurol* **56**, 303-308.

[30] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (1984) 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* **34**, 939-944.

[31] Calero O, Hortiguela R, Bullido MJ, Calero M (2009) Apolipoprotein E genotyping method by real time PCR, a fast and cost-effective alternative to the TaqMan and FRET assays. *J Neurosci Methods* **183**, 238-240.

[32] Baker M (2011) Metabolomics: From small molecules to big ideas. *Nat Meth* **8**, 117-121.

[33] Duportet X, Aggio R, Carneiro S, Villas-Bôas S (2012) The biological interpretation of metabolomic data can be misled by the extraction method used. *Metabolomics* **8**, 410-421.

[34] Grumbach E, Wheat T, Mazzeo J (2010) A novel method for the analysis of amino acids. Waters Application notes. 2007 Lib. Number WA43226.

[35] van der Kloet FM, Bobeldijk I, Verheij ER, Jellema RH (2009) Analytical error reduction using single point calibration for accurate and precise metabolomic phenotyping. *J Proteome Res* **8**, 5132-5141.

[36] Butterfield DA, Pocernich CB (2003) The glutamatergic system and Alzheimer's disease: Therapeutic implications. *CNS Drugs* **17**, 641-652.

[37] Ilzecka J, Stelmasiak Z, Solski J, Wawrzycki S, Szpetnar M (2003) Plasma amino acids percentages in amyotrophic lateral sclerosis patients. *Neurol Sci* **24**, 293-295.
[38] Basun H, Forssell LG, Almkvist O, Cowburn RF, Eklof R, Winblad B, Wetterberg L (1990) Amino acid concentrations in cerebrospinal fluid and plasma in Alzheimer's disease and healthy control subjects. *J Neural Transm Park Dis Dement Sect* **2**, 295-304.
[39] D'Aniello A, Fisher G, Migliaccio N, Cammisa G, D'Aniello E, Spinelli P (2005) Amino acids and transaminases activity in ventricular CSF and in brain of normal and Alzheimer patients. *Neurosci Lett* **388**, 49-53.
[40] Malinow R (2012) New developments on the role of NMDA receptors in Alzheimer's disease. *Curr Opin Neurobiol* **22**, 559-563.
[41] Reisberg B, Doody R, Stoffler A, Schmitt F, Ferris S, Mobius HJ (2003) Memantine in moderate-to-severe Alzheimer's disease. *N Engl J Med* **348**, 1333-1341.
[42] Tsai PJ, Huang PC (1999) Circadian variations in plasma and erythrocyte concentrations of glutamate, glutamine, and alanine in men on a diet without and with added monosodium glutamate. *Metabolism* **48**, 1455-1460.
[43] Haughey NJ, Bandaru VV, Bae M, Mattson MP (2010) Roles for dysfunctional sphingolipid metabolism in Alzheimer's disease neuropathogenesis. *Biochim Biophys Acta* **1801**, 878-886.
[44] Mielke MM, Lyketsos CG (2010) Alterations of the sphingolipid pathway in Alzheimer's disease: New biomarkers and treatment targets? *Neuromolecular Med* **12**, 331-340.
[45] Mielke MM, Bandaru VV, Haughey NJ, Rabins PV, Lyketsos CG, Carlson MC (2010) Serum sphingomyelins and ceramides are early predictors of memory impairment. *Neurobiol Aging* **31**, 17-24.
[46] Mielke MM, Ratnam Bandaru VV, Haughey NJ, Xia J, P, Yasar FL, Varma S, Harris V, Schneider G, Rabins EB, Bandeen-Roche PV, Lyketsos K, Carlson CG, MC (2012) Serum ceramides increase the risk of Alzheimer disease. *Neurology* **79**, 633-641.
[47] Mielke MM, Haughey NJ, Ratnam Bandaru VV, Schech S, Carrick R, Carlson MC, Mori S, Miller MI, Ceritoglu C, Brown T, Albert M, Lyketsos CG (2010) Plasma ceramides are altered in mild cognitive impairment and predict cognitive decline and hippocampal volume loss. *Alzheimers Dement* **6**, 378-385.
[48] Cunnane SC, Schneider JA, Tangney C, Tremblay-Mercier J, Fortier M, Bennett DA, Morris MC (2012) Plasma and brain fatty acid profiles in mild cognitive impairment and Alzheimer's disease. *J Alzheimers Dis* **29**, 691-697.
[49] Schaefer EJ, Bongard V, Beiser AS, Lamon-Fava S, Robins SJ, Au R, Tucker KL, Kyle DJ, Wilson PW, Wolf PA (2006) Plasma phosphatidylcholine docosahexaenoic acid content and risk of dementia and Alzheimer disease: The Framingham Heart Study. *Arch Neurol* **63**, 1545-1550.
[50] Walter A, Korth U, Hilgert M, Hartmann J, Weichel O, Hilgert M, Fassbender K, Schmitt A, Klein J (2004) Glycerophosphocholine is elevated in cerebrospinal fluid of Alzheimer patients. *Neurobiol Aging* **25**, 1299-1303.
[51] Bennett SA, Valenzuela N, Xu H, Franko B, Fai S, Figeys D (2013) Using neurolipidomics to identify phospholipid mediators of synaptic (dys)function in Alzheimer's Disease. *Front Physiol* **4**, 168.
[52] Han X, Holtzman DM, McKeel DW Jr (2001) Plasmalogen deficiency in early Alzheimer's disease subjects and in animal models: Molecular characterization using electrospray ionization mass spectrometry. *J Neurochem* **77**, 1168-1180.
[53] Yurko-Mauro K (2010) Cognitive and cardiovascular benefits of docosahexaenoic acid in aging and cognitive decline. *Curr Alzheimer Res* **7**, 190-196.
[54] Astarita G, Jung KM, Berchtold NC, Nguyen VQ, Gillen DL, Head E, Cotman CW, Piomelli D (2010) Deficient liver biosynthesis of docosahexaenoic acid correlates with cognitive impairment in Alzheimer's disease. *PLoS One* **5**, e12538.
[55] Young EA, Abelson J, Lightman SL (2004) Cortisol pulsatility and its role in stress regulation and health. *Front Neuroendocrinol* **25**, 69-76.
[56] Hylemon PB, Zhou H, Pandak WM, Ren S, Gil G, Dent P (2009) Bile acids as regulatory molecules. *J Lipid Res* **50**, 1509-1520.
[57] Ogundare M, Theofilopoulos S, Lockhart A, Hall LJ, Arenas E, Sjovall J, Brenton AG, Wang Y, Griffiths WJ (2010) Cerebrospinal fluid steroidomics: Are bioactive bile acids present in brain? *J Biol Chem* **285**, 4666-4679.
[58] Sattlecker M, Kiddle SJ, Newhouse S, Proitsi P, Nelson S, Williams S, Johnston C, Killick R, Simmons A, Westman E, Hodges A, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, the AddNeuroMed C, Dobson RJ (2014) Alzheimer's disease biomarker discovery using SOMAscan multiplexed protein technology. *Alzheimers Dement* **10**, 724-734.
[59] Roe CM, Fagan AM, Williams MM, Ghoshal N, Aeschleman M, Grant EA, Marcus DS, Mintun MA, Holtzman DM, Morris JC (2011) Improving CSF biomarker accuracy in predicting prevalent and incident Alzheimer disease. *Neurology* **76**, 501-510.
[60] Fagan AM, Roe CM, Xiong C, Mintun MA, Morris JC, Holtzman DM (2007) Cerebrospinal fluid tau/beta-amyloid(42) ratio as a prediction of cognitive decline in nondemented older adults. *Arch Neurol* **64**, 343-349.
[61] Bjorkqvist M, Ohlsson M, Minthon L, Hansson O (2012) Evaluation of a previously suggested plasma biomarker panel to identify Alzheimer's disease. *PLoS One* **7**, e29868.
[62] Ioannidis JP, Panagiotou OA (2011) Comparison of effect sizes associated with biomarkers reported in highly cited individual articles and in subsequent meta-analyses. *JAMA* **305**, 2200-2210.