Prediction of Mild Cognitive Impairment that Evolves into Alzheimer's Disease Dementia within Two Years using a Gene Expression Signature in Blood: A Pilot Study

Roed, Line; Grave, Gisle; Lindahl, Torbjorn; Rian, Edith; Horndalsveen, Peter O.; Lannfelt, Lars; Nilsson, Christer; Swenson, Frank; Lonneborg, Anders; Sharma, Praveen; Sjogren, Magnus

Published in:
Journal of Alzheimer's Disease

DOI:
10.3233/JAD-122404

2013

Journal of Alzheimer's Disease 35 (2013) 611–621
DOI 10.3233/JAD-122404
IOS Press

Prediction of Mild Cognitive Impairment that Evolves into Alzheimer's Disease Dementia within Two Years using a Gene Expression Signature in Blood: A Pilot Study

Line Roed[a,*], Gisle Grave[a], Torbjørn Lindahl[a], Edith Rian[a], Peter O. Horndalsveen[b], Lars Lannfelt[c], Christer Nilsson[d], Frank Swenson[e], Anders Lönneborg[a], Praveen Sharma[a] and Magnus Sjögren[a]
[a]*DiaGenic ASA, Oslo, Norway*
[b]*Sykehuset Innlandet, Alderspsykiatrisk avd. Sanderud, Ottestad, Norway*
[c]*Geriatriska Minnesmottagningen, Akademiska sjukhuset, Uppsala, Sweden*
[d]*Department of Clinical Sciences, Lund University, Lund, Sweden*
[e]*Pfizer Inc., New York, NY, USA*

Handling Associate Editor: Brit Mollenhauer

Accepted 8 February 2013

Abstract.
Background: The focus on Alzheimer's disease (AD) is shifting from dementia to the prodromal stage of the disorder, to a large extent due to increasing efforts in trying to develop disease modifying treatment for the disorder. For development of disease-modifying drugs, a reliable and accurate test for identification of mild cognitive impairment (MCI) due to AD is essential.
Objective: In the present study, MCI progressing to AD will be predicted using blood-based gene expression.
Material and Methods: Gene expression analysis using qPCR was performed on blood RNA from a cohort of patients with amnestic MCI (aMCI; $n = 66$). Within the aMCI cohort, patients progressing to AD within 1 to 2 years were grouped as MCI converters ($n = 34$) and the patients remaining at the MCI stage after 2 years were grouped as stable MCI ($n = 32$). AD and control populations were also included in the study.
Results: Multivariate statistical method partial least square regression was used to develop predictive models which later were tested using leave-one-out cross validation. Gene expression signatures that identified aMCI subjects that progressed to AD within 2 years with a prediction accuracy of 74%–77% were identified for the complete dataset and subsets thereof.
Conclusion: The present pilot study demonstrates for the first time that MCI that evolves into AD dementia within 2 years may be predicted by analyzing gene expression in blood. Further studies will be needed to validate this gene signature as a potential test for AD in the predementia stage.

Keywords: Alzheimer's disease, biomarkers, diagnostic tests, gene expression signatures, mild cognitive impairment

INTRODUCTION

The World Alzheimer Report in 2009 estimated that 36 million people worldwide are living with dementia, with numbers nearly doubling every 20 years to 66

*Correspondence to: Dr. Line Roed, DiaGenic ASA, Grenseveien 92, NO-0663 Oslo, Norway. Tel.: +47 23 24 89 50; Fax: +47 23 24 89 59; E-mail: line.roed@diagenic.com.

ISSN 1387-2877/13/$27.50

million by 2030, and 115 million by 2050 [1], that is unless no effective treatment is developed to prevent its progression. The 2010 report calculated the worldwide costs of dementia (US$604 billion in 2010) to amount to more than 1% of global GDP [2]. The rapidly increasing impact and cost for people and society represents a strong driving force for the development of new drugs and diagnostic tools for early treatment and disease management. As of now, no disease modifying treatment for Alzheimer's disease (AD) or any other dementia disorder is available. The development of diagnostic tools may aid in identifying individuals eligible for clinical trials as well as in clinical practice aiding in the decisions of whom to treat.

In clinical practice today, the diagnosis of AD is made on the basis of clinical symptoms and signs, requiring the presence of memory impairment as well as impairment in one additional cognitive domain, and with a concomitant functional decline lasting for more than six months. However, several studies have shown that there are biochemical [3, 4] and neuroanatomical [5] signs of a developing disorder many years before the onset of clinical symptoms in AD. To enable an early diagnosis and incorporate these repeated new findings, new guidelines have been developed and published [6]. These include the use of biomarkers in the diagnosis of mild cognitive impairment (MCI) that evolves into AD [6]. Several biomarkers such as volumetric magnetic resonance imaging (MRI) [7], amyloid imaging using positron emission tomography (PET) [8], and cerebrospinal fluid (CSF) biomarkers such as amyloid-β (Aβ) and total-tau have been proposed in the early diagnosis of AD [9]. However, these are either more invasive, time-consuming, procedure complex, or very expensive compared to that of a simple blood test. A cost benefit analysis by Yu et al. [10] demonstrated that removal of both CSF biomarkers and FDG-PET from the enrichment strategy for clinical trials provided both considerable cost savings and shorter screening times. Methodological and technical variation has also been described [11]. Thus, there is need to develop biomarkers that are practical to use, minimally invasive and reliable, yet still highly accurate [12–15].

Several independent studies have indicated that a blood-based test could be used for diagnostic profiling in neurological diseases [16–24]. Lönneborg [21] summarized in 2008 that of several studies that investigated plasma Aβ levels in AD; one study showed an increase in Aβ levels [25] and the majority of studies found no significant differences between AD and controls [26–30]. Later, a systematic review and meta-analysis in 2012 by Koyama et al. [31] concluded that plasma $A\beta_{42}$:$A\beta_{40}$ ratio may predict development of AD, however, significant heterogeneity in the meta-analysis underlines the need for substantial further investigation of plasma Aβ levels as a preclinical biomarker. Thambisetty et al. have identified plasma clusterin as a potential biomarker for AD [32] and present data that suggests that plasma clusterin is associated with rate of atrophy in MCI [33].

As the first company in the world, DiaGenic developed a CE marked blood-based diagnostic test for AD (ADtect®) which identifies patients with mild to moderate AD with an accuracy of 72%. The development and validation of this 96-assay gene expression signature has been described in Booij et al. and Rye et al. [34, 35]. This test detects the blood-based gene expression signature in blood samples and adopts a combination of multiple gene expression assays to obtain a prediction value for disease classification. Alterations in a gene expression signature in peripheral blood are postulated as a result of a systemic disease response in AD as verified by many previous studies [34–38].

Based on the same approach, DiaGenic has investigated the gene expression signature in blood of patients with amnestic MCI (aMCI) who within 2 years progress to AD dementia and compared this with the gene expression pattern in stable MCI. Diagnostic performance (prediction accuracy, sensitivity, and specificity) of the identified multivariate biomarker is presented.

MATERIAL AND METHODS

Participants

Patients with MCI ($n = 66$) were recruited from 8 centers in Norway and Sweden from January 2007 until January 2011. All patients were clinically examined longitudinally on a yearly basis.

MCI patients progressing to AD within 2 years were grouped as MCI converters ($n = 34$) and the patients maintaining an MCI diagnosis after 2 years (i.e., at the 3rd visit) were grouped as stable MCI ($n = 32$).

The clinical investigations were performed by clinicians with expertise in diagnosing MCI and dementia. Clinical interview and medical examination were performed, and physical diseases and drug use were recorded. Imaging (either MRI or CT) and routine blood tests were performed for all subjects to exclude other causes of MCI or dementia than MCI or AD, respectively.

The cognitive and functional test battery included an overall cognitive screening test (Mini-Mental State

Examination, MMSE) [39] and a set of standardized cognitive tests assessing verbal and non-verbal memory, processing speed and executive function (shifting attention): The Clock Test [40, 41], Kendrick OLT test [42], 10-word test [43, 44], and Trail Making test A and B [45]. In addition, the clinical dementia rating (CDR) scale [46] was scored based on a detailed clinical interview with the patient and a caregiver.

MCI (amnestic MCI) was diagnosed in accordance with the Petersen criteria [47] with symptoms present for at least 6 months, and a diagnosis of AD was in accordance with International Classification of Diseases (ICD) adopted by the WHO member states in 1994 (ICD-10) and the National Institute of Neurological and Communication Disorders and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria [48].

All diagnoses and the staging of cognitive impairment/dementia were evaluated blindly by an expert group of clinicians having access to all relevant clinical, laboratory, and imaging longitudinal patient data in order to independently establish a consensus clinical diagnosis.

The procedures included blood sample collection and were approved by local ethics committees. The blood samples were obtained at the clinical centers at the time of clinical evaluation. Written informed consent was obtained from each included subject.

Sample collection

Venous blood samples (2.5 mL) were drawn into PAXgene™ tubes (Becton & Dickenson, Qiagen Inc., Valencia, CA) according to the manufacturer's instructions. Tubes were incubated at room temperature overnight prior to freezing and stored at −70°C or below. RNA was extracted from all samples ($n = 66$) within 6 months of blood draw and stored at −70°C or below in the DiaGenic biobank.

Assay selection and micro-fluidic card (MFC) layout

A total of 1,239 gene probes have previously been selected [35] and served as a basis for the identification of informative genes. The 1,239 gene probes were selected based on an AD classifier developed following a whole genome screen comprising in total 126 clinically diagnosed AD patients, 98 age-matched cognitively healthy controls, 28 young controls, 28 Parkinson's disease patients, and 10 subjects with MCI. TaqMan assays with adequate efficiency, expression level in blood RNA, and absence of signal from genomic DNA, were available for 970 of the 1,239 initial genes (Applied Biosystems, Foster City, CA). In addition, reference assay candidates and assays for transcripts identified by an extended literature search were also included. A total of 1,152 assays distributed on three different micro-fluidic cards (MFCs) were custom ordered from Applied Biosystem (Foster City, CA).

In addition to the study MFCs, a DiaGenic technical control MFC (4 × 96 assay format) was used to monitor the temporal stability during the gene expression analyses.

RNA extraction and cDNA synthesis

Total RNA was extracted from blood samples using PAXgene™ Blood RNA kit for manual extraction (PreAnalytix, Hombrechtikon, Switzerland) according to the manufacturer's instructions. Total RNA was stored at −70°C or below until analysis. The RNA was assessed for quality using the NanoDrop ND-1000 spectrophotometer (NanoDrop, Wilmington, DE) and the Agilent 2100 BioAnalyzer (Agilent, Santa Clara, CA), with sample acceptance limits RIN ≥7.3; 28S/18S ≥1.0; A260/A230 ≥1.0; A260/A280 ≥1.8; RNA concentration ≥30 ng/μL. The cDNA was prepared according to a predefined randomization scheme using the High-Capacity cDNA Reverse Transcriptase kit (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. From each subject sample, 2,210 ng RNA was used to prepare adequate amounts of cDNA for gene expression analysis on the complete set of MFCs (3 × 384-assay cards). The final concentration of RNA in the cDNA reaction mixture was 8.5 ng/μl. This concentration was tested and shown to be below the limit of inhibition of the cDNA synthesis. The cDNA was stored at −70°C or below until analysis.

Real time qPCR

Real time qPCR (RT-qPCR) was performed using the TaqMan® Universal PCR Master Mix II (2X) with uracil-N-glycosylase according to the manufacturer's instructions in the standard run mode on 2 ViiA7 Dx systems (Life Technologies, Carlsbad, CA) equipped with MFC blocks and robotic feeders (TwisterII, Caliper, Hopkinton, MA). The ViiA7 systems were operated by the ViiA7 Software version 1.0 (Life Technologies). The two instruments used during the study were qualified according to internal procedures prior to use and normalized against each other

according to the manufacturer's procedures. Subsequently, side-by-side studies were performed to verify that the two instruments yielded comparable results (data not shown).

Samples were run on MFCs according to a predefined randomization scheme different from that used for cDNA synthesis. The donors were randomized independently for each of the MFCs. All samples were run on a complete set of MFCs, i.e., 3×384-assay cards. Prepared cDNA was diluted 1/10 in the PCR reaction mixture, which was tested and shown to be below the level of inhibition of the qPCR. Each of the 8 lanes of each card was loaded with 97 µl PCR reaction mixture. Five MFCs were filled simultaneously and loaded into the RT-qPCR using the robotic feeder. To avoid potential fading of the fluorescent dye and to ensure stability of the reaction mixture on the MFCs, the instruments were in a room with dimmed light and temperature control.

During the study RNA reference samples were run at regular time intervals randomized among and processed as the study samples. The reference samples were used to monitor technical aspects such as instrument variation, inter-card and inter-day variability. Analysis of the gene expression of the reference material demonstrated no significant difference between the two instruments for the 20 assay set with a $p = 0.3$ using multivariate ffmanova (an R package implementing fifty-fifty ANOVA) [49].

Reference samples were both run on study MFCs and on technical control MFCs. Temporal drift was investigated by monitoring the average C_T of the technical replicates of the reference material. No systematic drift of average C_T value was observed following analysis of the technical replicates (Fig. 1), and the average C_T was in the range of 1 C_T.

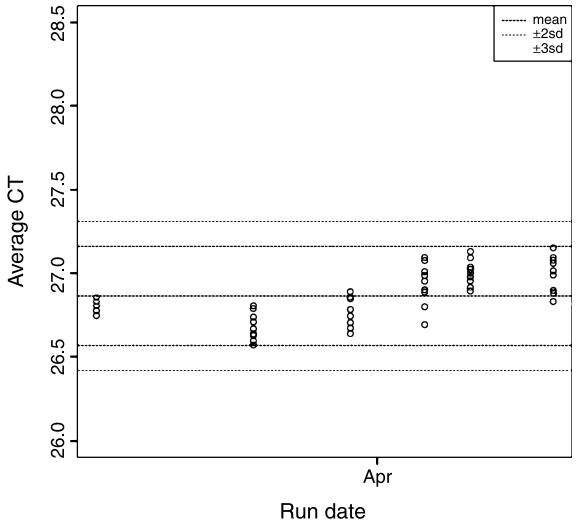

Fig. 1. Variation of average C_T of an RNA reference material during the study.

Data analysis

The present study was performed as a retrospective analysis. Subject data collection and blood draw were both performed prior to inclusion in the present study.

Optimal thresholds were set by the ViiA7 software for determination of C_T for each assay. Missing values were imputed using k-Nearest Neighbors (k-NN) [50]. A panel of reference assays was investigated. Beta-actin demonstrated to be stably expressed and was used for reference gene normalization of the data.

Partial Least Square Regression (PLSR) and Leave One Out Cross-Validation (LOO-CV) were used to build the gene expression algorithm (model) and to estimate the diagnostic performance. Jack-knife feature [51] was used for variable selection. A number of models were developed and evaluated, investigating complete datasets which excluding reference assays constituted 1,123 assays or subsets thereof. For each model the LOO-CV estimates of sensitivity, specificity, prediction accuracy, and area under the receiver operating characteristic curve (AUC) were used as a measure of the models performance as compared to the clinical diagnosis. The statistical computing language R version 2.10.0 was used [52].

Clinical interferences

Clinical interferences were studied in the study population. Subjects with co-morbidities such as cancer, hypertension, diabetes, coronary disease, depression, and allergies were compared with subjects without a history of chronic illness using PCA score analysis. This was done to evaluate if co-morbidities were likely to affect the test result.

Biological significance

To investigate the biological significance of proteins encoded by the selected gene assays pathway analyses were performed. The pathway analyses were performed using Ingenuity Pathway Analysis software (IPA, Ingenuity Systems, Redwood City, CA). Benjamini-Hochberg multiple testing correction [53] at $p < 0.05$ was used to correct for multiple testing when identifying biological functions and canonical pathways significantly associated with assay sets.

Table 1
Demographic distribution of subjects ($n = 66$) used in the development of the disease model

		Stable MCI $n = 32$	MCI converters $n = 34$
Educational level	Primary/secondary school (up to 9 y)	1 (3%)	2 (6%)
	High school (up to 12 y)	9 (28%)	7 (21%)
	College and university	11 (34%)	15 (44%)
	Unknown	11 (34%)	10 (29%)
Gender	Male	14 (44%)	15 (44%)
	Female	18 (56%)	19 (56%)
Baseline MMSE	Mean	28	28
	Range (min-max)	(23–30)	(23–30)
MMSE (1 y visit)	Mean	28	27
	Range (min-max)	(24–30)	(22–30)
MMSE (2 y visit)	Mean	28	26
	Range (min-max)	(22–30)	(19–29)
Baseline CDR	Mean	0.5	0.5
	Range (min-max)	(0.0–1.0)	(0.5−0.5)
CDR (1 y visit)	Mean	0.5	0.5
	Range (min-max)	(0.0–1.0)	(0.5–1.0)
CDR (2 y visit)	Mean	0.5	0.5
	Range (min-max)	(0.5−0.5)	(0.5–1.0)
Baseline CDR-SOB	Mean	1	1
	Range (min-max)	(0–5)	(1–4)
CDR-SOB (1 y visit)	Mean	1	2
	Range (min-max)	(0–5)	(1–6)
CDR-SOB (2 y visit)	Mean	1	3
	Range (min-max)	(1–4)	(1–7)
Age (y)	Mean	67.4	71.5
	Range (min-max)	(52–81)	(52–84)

MCI, mild cognitive impairment; MMSE, Mini-Mental Status Examination; CDR, Clinical Dementia Rating, SOB, sum of boxes.

The functional categories used in IPA were developed by Ingenuity staff of Ph.D. scientists who are experts across various domains of biology. These categories strive to encompass all types of biological functions that are most important to researchers using IPA.

The dynamic Canonical Pathways are well-characterized metabolic and cell signaling pathways that have been curated and hand-drawn by Ph.D. level scientists based on previous literature findings. The information contained in Canonical Pathways comes from specific journal articles, review articles, text books, and HumanCyc.

RESULTS

Subject characteristics

The included individuals in this study were recruited through DiaGenic blood collection studies at 8 clinical centers in Norway and Sweden using a standardized protocol. Clinical and demographic features of the study population are shown in Table 1. A total of 66 subjects were included in the present study and blood collected at baseline was used to investigate the gene expression signature.

Prediction of MCI that evolves into AD dementia within two years in subjects with aMCI

For development of the gene expression algorithm, PLSR and LOO-CV were used both to build the model and to estimate the diagnostic performance of the generated models. A number of models were developed and evaluated, and, similarly, a number of different datasets were investigated. The selection of these datasets is described below.

Our first model was developed using the complete dataset compiled of gene expression data obtained from 1,123 assays. Using the gene expression data from the stable MCI and the MCI converters, a PLSR model was developed discriminating the two classes. The model was evaluated by LOO-CV from which a prediction accuracy of 74% was demonstrated. The model performance characteristics are summarized in Table 2.

A large assay set of 1,123 assays is unpractical and costly for clinical use. A reduction of the number of assays used by investigating a subset of the data was

explored aiming at maintaining the most informative assays. Using a jack-knife procedure, a total of 225 of the most informative or significant assays ($p < 0.2$) were identified. The gene expression data obtained from this data subset was used for statistical modeling using PLSR for the discrimination of stable MCI from MCI converters. A model demonstrating a prediction accuracy of 77% following LOO-CV was obtained. The performance characteristics of this model based on the 225 assay set are summarized in Table 2.

No further jack-knife reduction was performed for the present dataset as subsequent product development did not aim at final assay selection at that time. For product development, further adaptation of the qPCR protocol for regulatory approval purposes is planned prior to locking of the algorithm; moreover a larger number of assays into the optimization phase is preferred. However, in this study we wanted to investigate a subset of 20 informative assays that we have previously identified in independent in-house studies (results not shown) during the development of ADtect® [34, 35]. The identification of these assays was based on the discrimination of AD and cognitively healthy controls. Based on this, these 20 assays were included as part of the complete assay set in the present study. Hence, the 20 assays were not a jack-knife product of the complete assay set as the 225 assays. The applicability of these 20 informative assays for the prediction of MCI that evolves into AD dementia within 2 years in an aMCI population was investigated. Using this exploratory dataset, a PLSR model was developed using the gene expression data for the discrimination of stable MCI from MCI converters. The developed algorithm using the 20 informative assays selected based on AD pathology demonstrated a prediction accuracy of 77%. The model performance when discriminating MCI converters from stable MCI is summarized in Table 2. This demonstrates that it is possible to retain the same level of prediction accuracy even when reducing the assay set to only a fraction of the initial dataset, in this case 20 assays. A summary of the origin of the investigated assay sets is shown in Fig. 2.

Clinical interferences

In the present study, the possibility of co-morbidity interferences was investigated, including a history of, for example, cancer, hypertension, diabetes, coronary disease, depression, and allergies. In the cohort of stable MCI patients, 11 of the 34 subjects reported a history of chronic disease. Similarly, 13 of 32 of the MCI converters had a history of chronic disease.

Table 2
Summary of model performances

Performance parameter	Complete dataset (1123 assays) %	Reduced dataset (225 assays) %	Exploratory dataset (20 assays) %
Prediction accuracy	74	77	77
Sensitivity	70	73	70
Specificity	77	81	84

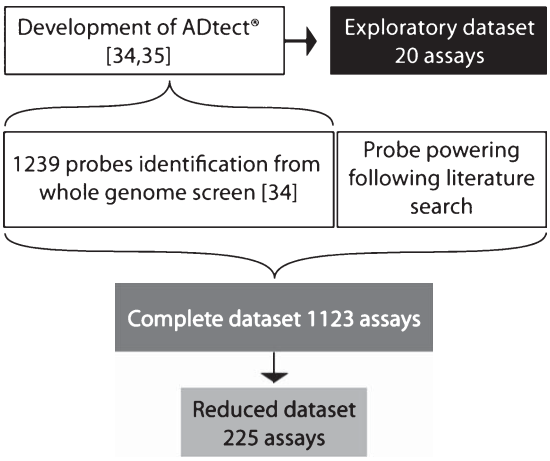

Fig. 2. Summary of origin of investigated assay sets.

PCA scores revealed no clear effect for individuals suffering from at least one co-morbidity. It was therefore concluded that co-morbidities did not affect the results in the present study.

The use of medication could not be investigated in the present study as only 3 of 32 stable MCI and 3 of 32 MCI converters were on acetylcholinesterase inhibitor.

Biological significance

For all assays selected, i.e., both the complete dataset and data subsets, the assays were chosen based on the predictive ability in an algorithm and not selected based on the individual informative nature of the genetic origin of the targeted transcripts. Still, we find that a majority of the genes included in the whole assay set as well as in the 225 assay set and the 20 assay set encode proteins with biological functions associated with AD and AD-related biological processes.

Of the 225 assays, 207 encode proteins with known identity. Adjusting for multiple testing using Benjamini-Hochberg correction [53], Aβ processing is the canonical pathway that separates from other canonical pathways with the lowest p-value ($-\log 3.599$). Seven of the 54 proteins (ratio = 0.13) included in this canonical pathway are among the 207 proteins with a

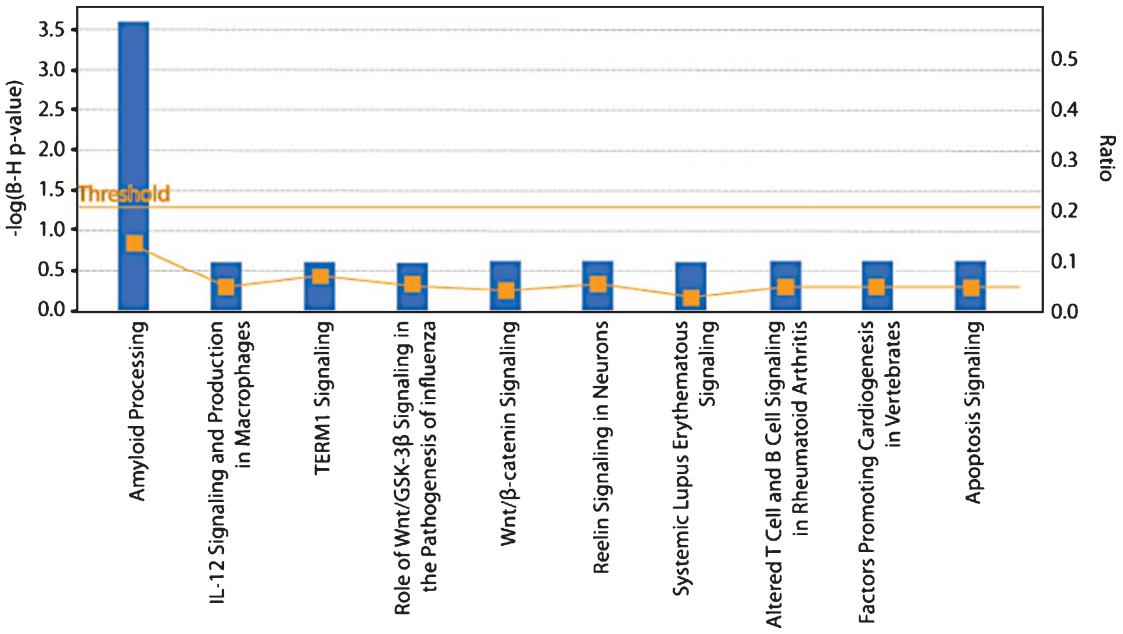

Fig. 3. Canonical pathways associated with the proteins encoded by the 225 assay set. Blue bars: −log (Benjamini-Hochberg *p*-value); Orange squares: Ratio of the number of proteins in the pathway represented by the assay set/total number of proteins in the pathway. Orange horizontal line denotes the cutoff for significance (*p*-value of 0.05).

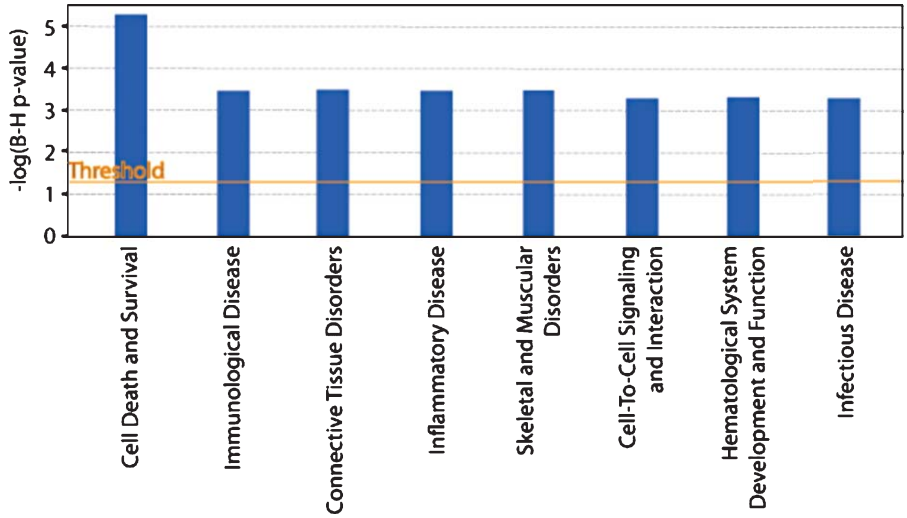

Fig. 4. Biological functions associated with the proteins encoded by the 225 assay set. Blue bars: −log (Benjamini-Hochberg *p*-value); Orange horizontal line denotes the cutoff for significance (*p*-value of 0.05).

known identity (Fig. 3) suggesting that this pathway has a pronounced influence on the model. Cell death and survival is the most significant biological function represented by these proteins (Fig. 4). Of the proteins, 73 are associated with cell death and survival and 64 of these proteins are associated with apoptosis.

Due to the limited number of proteins encoded by the 20 assays, no reliable analysis of significant canonical pathways or biological functions represented by these proteins could be done. Therefore the pathway analysis was done on the larger 225 assay set. Of the proteins encoded by the 20 assay set, 10 are directly associated with and cover all biological processes that have been listed above for the 225 assay set.

Among the 207 proteins with known identity, 157 are directly associated with AD-related proteins

according to KEGG or with biological processes that have been associated with the disease (data not shown). These processes include neurodegeneration, amyloid-β, and tau processing, mitochondria transport and function, inflammation, apoptosis, and calcium signaling. Our findings of transcript profiles in blood suggest that these disease-related processes are not affected only in CNS tissue but also in other parts of the body.

DISCUSSION

We and others have previously shown that disease biomarkers to classify AD with dementia from cognitively healthy individuals are present in blood [16, 23, 35] and an algorithm based on the gene expression of a selected set of genes can be used to identify those with an AD dementia [34–36]. We have now further extended these findings to also include AD in the pre-dementia stage of the disease. The results of the present study demonstrate that information to classify MCI due to AD in an aMCI population is present in blood. The results show the potential of using a prediction algorithm based on the expression of a limited set of expressed genes in blood for identification of MCI that evolves into AD dementia within 2 years among individuals with aMCI. When balancing the sensitivity and specificity, a prediction accuracy ranging from 74% to 77% was obtained for PLSR models built using gene expression data from stable MCI and MCI that evolves into AD dementia within 2 years.

A majority of the gene assays included in both the 225 assay set and in the 20 assay set encode proteins that are directly associated with biological processes that has been associated with AD. This suggests that these processes not only are affected in brain tissue but also are reflected in a systemic response that can be detected using gene expression in blood. This is also well in agreement with similar findings in previous AD studies [34–38]. The results also support the concept of AD as a multifactorial sporadic disorder [54] with multiple genes and alterations in gene expression involved [55, 56]. Aβ processing appears to be a significant canonical pathway and apoptosis a significant biological function represented among the proteins encoded by the 225 assay set. However, it is worth noting that there are several proteins encoded by the assay set that do not appear to be involved in any of the processes that has been associated with AD and it could be of interest to explore these proteins further to see if novel biological functions may be found that are essential to extend our understanding of the development of this devastating disease.

The clinical applicability of a multivariate index assay test based on gene expression in blood as described here is as a selection tool in clinical trials by enriching the MCI population with higher risk to progress to AD within the next few years as well as a clinician tool as an aid in the diagnosis for early treatment, future planning, and disease management. A blood-based biomarker is a patient-friendly, less invasive, and more affordable alternative to other investigated biomarkers such as CSF biomarkers or amyloid imaging using PET that are either invasive or very expensive [12, 14].

In the present study the possibility of co-morbidity interferences was investigated, but not found to affect the results. However, it cannot be ruled out that some of the co-morbidities may result in an alteration in gene expression. Future investigation may reveal clearer which co-morbidity has the highest impact of gene expression and if the magnitude has an effect on results from downstream analysis.

Parnetti et al. recently developed CSF biomarker models based on logistic regression analysis for the prediction of MCI that evolves into AD dementia within 2 years among individuals with MCI [57]. Using the CSF biomarkers $A\beta_{1\text{-}42}$ and p-tau for the model development, they obtained a sensitivity of 75% and specificity of 96%. In a multimodal classification approach combining PET-FDG, MR features following volumetric MR imaging, and the level of selected CSF biomarkers, Zhang et al. achieved a sensitivity of 92% and specificity of 73% which was superior to the performance obtained when investigating the individual modality of biomarkers [58]. The sensitivity and specificity using either CSF or CSF in combination with PET-FDG and volumetric MRI are higher than what is achieved with either of our blood-based gene expression models. However, our prodromal AD biomarker is the first blood-based gene expression biomarker described, and further refinement of the biomarker both by assay selection and increased sample size may demonstrate the full potential of such a biomarker. Still, the use of a blood-based gene biomarker is both a non-invasive low-risk procedure as well as cost effective compared to that of lumbar puncture, making blood-based gene expression biomarker tests an attractive alternative to CSF biomarkers.

In the present study a few donors had been followed beyond the 3rd visit. Of these there were 5 donors with stable MCI for 2 years that were diagnosed with AD at the 4th visit (3 years after baseline) during the longitudinal trial. The present PLSR model predicted

all of these donors as stable MCI. However, all these subjects may be considered slow converters that possibly is somewhere in between the stable MCI and MCI converters. The subgroup of stable MCI is likely a heterogeneous group and as time passes a fraction of these are expected to progress to AD. The prediction accuracy is therefore dependent on where the follow-up cutoff is set. For the present study we have chosen 2 year follow-up as a cut-off to identify MCI in an aMCI population that will progress to AD within 2 years. Although a substantial portion of MCI progress to AD within 1–2 years, the progression from MCI to AD has been demonstrated to take as much as 10 years [59]. Compared to CSF biomarkers, our prodromal AD biomarker has a dynamic characteristics compared to the more static feature of CSF biomarkers. Our test predicts changes to occur within 2 years, which is a relatively short period of time. CSF biomarkers may have a longer time perspective. This static feature was recently reported by Buchhave et al. [59] who demonstrated that baseline CSF $A\beta_{42}$ levels were equally reduced in patients with MCI who converted to AD within 0 to 5 years (early converters) compared with those who converted between 5 and 10 years (late converters).

When building a PLSR model including only the available clinical parameters (scores from the neuropsychological test battery), a prediction accuracy of 67% was obtained, which is inferior to that obtained for the gene expression model. Including the clinical parameters in the gene expression model only marginally improved the diagnostic performance by increasing the prediction accuracy from 74% to 75%. This is in contrast to that observed by Gomar et al. [60]. They obtained a prediction accuracy of 71% using cognitive markers only and concluded these to be generally stronger predictors than investigated biomarkers including brain volume, CSF biomarkers, and APOE genotype.

The present study was a pilot study with a limited number of samples. The aim of the study was to demonstrate the potential of using the prodromal AD gene expression signature for the development of a multivariate index assay test for the prediction of MCI that evolves into AD dementia within 2 years in an aMCI population. The regulatory requirement to further optimize, technically verify and clinically validate the prodromal AD gene expression signature is strongly emphasized.

To conclude, we have shown for the first time that it is possible to detect MCI that evolves into AD dementia within 2 years before the onset of dementia by analysis of the gene expression in blood and that gene expression in peripheral blood is sensitive to a pathological process manifesting in the brain.

ACKNOWLEDGMENTS

The following hospitals or clinics contributed with clinical samples: Stavanger University Hospital (Norway), Haraldsplass Deaconess Hospital (Bergen, Norway), Karolinska University Hospital (Huddinge, Sweden), Skånes universitetssjukhus (Lund, Sweden), Innlandet Hospital (Sanderud, Norway), Stockholmsgeriatriken (Stockholm, Sweden), Ullevål University Hospital (Oslo), and Akademiska sjukhuset (Uppsala, Sweden). At the Akademiska sjukhuset Malin Degerman-Gunnarsson and Lena Kilander are acknowledged for their expertise in the recruitment of MCI subjects.

Funding was partly provided by The Norwegian Research Council Functional Genomics program and by Pfizer Inc.

Authors' disclosures available online (http://www.j-alz.com/disclosures/view.php?id=1676).

REFERENCES

[1] (2009) World Alzheimer Report 2009, Prince M, Jackson J, eds. Alzheimer's Disease International.

[2] (2010) World Alzheimer Report 2010, Wimo A, Prince M, eds. Alzheimer's Disease International.

[3] Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L (2006) Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: A follow-up study. *Lancet Neurol* **5**, 228-234.

[4] Hansson O, Zetterberg H, Vanmechelen E, Vanderstichele H, Andreasson U, Londos E, Wallin A, Minthon L, Blennow K (2010) Evaluation of plasma Abeta(40) and Abeta(42) as predictors of conversion to Alzheimer's disease in patients with mild cognitive impairment. *Neurobiol Aging* **31**, 357-367.

[5] Spulber G, Niskanen E, Macdonald S, Kivipelto M, Padilla D, Julkunen V, Hallikainen M, Vanninen R, Wahlund L, Soininen H (2012) Evolution of global and local grey matter atrophy on serial MRI scans during the progression from MCI to AD. *Curr Alzheimer Res* **9**, 516-524.

[6] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carillo MC, Thies B, Phelps CH (2011) The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the Nathional Institute on Aging and Alzheimer's Association workgroup. *Alzheimers Dement* **7**, 270-279.

[7] Chan D, Fox N, Jenkins R, Scahill R, Crum W, Rossor M (2001) Rates of global and regional cerebral atrophy in AD and frontotemporal dementia. *Neurology* **57**, 1756-1763.

[8] Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, Delacourte A, Galasko D, Gauthier S, Jicha G, Meguro K, O'Brien J, Pasquier F, Robert P, Rossor M,

Salloway S, Stern Y, Visser PJ, Scheltens P (2007) Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDS-ADRDA criteria. *Lancet Neurol* **6**, 734-746.

[9] Mattsson N, Zetterberg H, Hansson O, Andreasen N, Parnetti L, Jonsson M, Herukka S-K, Flier WMvd, Blankenstein MA, Ewers M, Rich K, Kaiser E, Verbeek M, Tsolaki M, Mulugeta E, Rosén E, Aarsland D, Visser PJ, Schröder J, Marcusseon J, Leon Md, Hampel H, Scheltens P, Pirttilä T, Wallin A, Jönhangen ME, Minthon L, Winblad B, Blennow K (2009) CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. *JAMA* **302**, 385-393.

[10] Yu P, Dean RA, Hall SD, Qi Y, Sethuraman G, Willis BA, Siemers ER, Martenyi F, Tauscher JT, Schwarz AJ (2012) Enriching amnestic mild cognitive impairment populations for clinical trials: Optimal combination of biomarkers to predict conversion to dementia. *J Alzheimers Dis* **32**, 373-385.

[11] Mattsson N, Blennow K, Zetterberg H (2010) Inter-laboratory variation in cerebrospinal fluid biomarkers for Alzheimer's disease: United we stand, divided we fall. *Clin Chem Lab Med* **48**, 603-607.

[12] Li Y, Rinne JO, Mosconi L, Pirraglia E, Rusinek H, DeSanti S, Kemppainen N, Nagren K, Kim BC, Tsui W, de Leon MJ (2008) Regional analysis of FDG and PIB-PET images in normal aging, mild cognitive impairment, and Alzheimer's disease. *Eur J Nucl Med Mol Imaging* **35**, 2169-2181.

[13] Brys M, Glodzik L, Mosconi L, Switalski R, De Santi S, Pirraglia E, Rich K, Kim BC, Mehta P, Zinkowski R, Pratico D, Wallin A, Zetterberg H, Tsui WH, Rusinek H, Blennow K, de Leon MJ (2009) Magnetic resonance imaging improves cerebrospinal fluid biomarkers in the early detection of Alzheimer's disease. *J Alzheimers Dis* **16**, 351-362.

[14] Buchhave P, Blennow K, Zetterberg H, Stomrud E, Londos E, Andreasen N, Minthon L, Hansson O (2009) Longitudinal study of CSF biomarkers in patients with Alzheimer's disease. *PLoS One* **4**, e6294.

[15] Aksu Y, Miller DJ, Kesidis G, Bigler DC, Yang QX (2011) An MRI-derived definition of MCI-to-AD conversion for long-term, automatic prognosis of MCI patients. *PLoS One* **6**, e25074.

[16] Maes OC, Xu S, Yu B, Chertkow HM, Wang E, Schipper HM (2007) Transcriptional profiling of Alzheimer blood mononuclear cells by microarray. *Neurobiol Aging* **28**, 1795-1809.

[17] Scherzer CR, Eklund AC, Morse LJ, Liao Z, Locascio JJ, Fefer D, Schwarzschild MA, Schlossmacher MG, Hauser MA, Vance JM, Sudarsky LR, Standaert DG, Growdon JH, Jensen RV, Gullans SR (2007) Molecular markers of early Parkinson's disease based on gene expression in blood. *Proc Natl Acad Sci U S A* **104**, 955-960.

[18] Sharp FR, Xu H, Lit L, Walker W, Apperson M, Gilbert DL, Glauser TA, Wong B, Hershey A, Liu DZ, Pinter J, Zhan X, Liu X, Ran R (2006) The future of genomic profiling of neurological diseases using blood. *Arch Neurol* **63**, 1529-1536.

[19] Tang Y, Nee AC, Lu A, Ran R, Sharp FR (2003) Blood genomic expression profile for neuronal injury. *J Cereb Blood Flow Metab* **23**, 310-319.

[20] Tang Y, Lu A, Ran R, Aronow BJ, Schorry EK, Hopkin RJ, Gilbert DL, Glauser TA, Hershey AD, Richtand NW, Privitera M, Dalvi A, Sahay A, Szaflarski JP, Ficker DM, Ratner N, Sharp FR (2004) Human blood genomics: Distinct profiles for gender, age and neurofibromatosis type 1. *Brain Res Mol Brain Res* **132**, 155-167.

[21] Lonneborg A (2008) Biomarkers for Alzheimer disease in cerebrospinal fluid, urine, and blood. *Mol Diagn Ther* **12**, 307-320.

[22] Winblad B, Lönneborg A (2008) Blood biomarkers for Alzheimer's disease. *Eur Neurol Rev* **3**, 28-30.

[23] Fehlbaum-Beurdeley P, Jarrige A-C, Pallares D, Carriere J, Guihal C, Soucaille C, Rouet F, Drouin D, Sol O, Jordan H, Wu D, Lei L, Einstein R, Schweighoffer F, Bracco L(2010) Toward an Alzheimer's disease diagnosis via high-resolution blood gene expression. *Alzheimers Dement* **6**, 25-38.

[24] Grunblatt E, Bartl J, Zehetmayer S, Ringel TM, Bauer P, Riederer P, Jacob CP (2009) Gene expression as peripheral biomarkers for sporadic Alzheimer's disease. *J Alzheimers Dis* **16**, 627-634.

[25] Mehta PD, Pirttila T, Mehta SP, Sersen EA, Aisen PS, Wisniewski HM (2000) Plasma and cerebrospinal fluid levels of amyloid beta proteins 1-40 and 1-42 in Alzheimer disease. *Arch Neurol* **57**, 100-105.

[26] Scheuner D, Eckman C, Jensen M, Song X, Citron M, Suzuki N, Bird TD, Hardy J, Hutton M, Kukull W, Larson E, Levy-Lahad L, Viitanen M, Peskind E, Poorkaj P, Schellenberg G, Tanzi R, Wasco W, Lannfelt L, Selkoe D, Younkin S (1996) Secreted amyloid β-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* **2**, 864-870.

[27] Kosaka T, Imagawa M, Seki K, Arai H, Sasaki H, Tsuji S, Asami-Odaka A, Fukushima T, Imai K, Iwatsubo T (1997) The beta APP717 Alzheimer mutation increases the percentage of plasma amyloid-beta protein ending at A beta42(43). *Neurology* **48**, 741-745.

[28] Tamaoka A, Fukushima T, Sawamura N, Ishikawa K, Oguni E, Komatsuzaki Y, Shoji S (1996) Amyloid beta protein in plasma from patients with sporadic Alzheimer's disease. *J Neurol Sci* **141**, 65-68.

[29] Vanderstichele H, Van Kerschaver E, Hesse C, Davidsson P, Buyse MA, Andreasen N, Minthon L, Wallin A, Blennow K, Vanmechelen E (2000) Standardization of measurement of beta-amyloid(1-42) in cerebrospinal fluid and plasma. *Amyloid* **7**, 245-258.

[30] Fukomoto H, Tennis M, Locascio JJ, Hyman BT, Growdon JH, Irizarry MC (2003) Age but not diagnosis is the main predictor of plasma amyloid beta-protein levels. *Arch Neurol* **60**, 958-964.

[31] Koyama A, Okereke OI, Yang T, Blacker D, Selkoe DJ, Grodstein F (2012) Plasma amyloid-β as a predictor of dementia and cognitive decline: A systematic review and meta-analysis. *Arch Neurol* **69**, 824-831.

[32] Thambisetty M, Simmons A, Velayudhan L, Hye A, Campbell J, Zhang Y, Wahlund LO, Westman E, Kinsey A, Guntert A, Proitsi P, Powell J, Causevic M, Killick R, Lunnon K, Lynham S, Broadstock M, Choudhry F, Howlett DR, Williams RJ, Sharp SI, Mitchelmore C, Tunnard C, Leung R, Foy C, O'Brien D, Breen G, Furney SJ, Ward M, Kloszewska I, Mecocci P, Soininen H, Tsolaki M, Vellas B, Hodges A, Murphy DG, Parkins S, Richardson JC, Resnick SM, Ferrucci L, Wong DF, Zhou Y, Muehlboeck S, Evans A, Francis PT, Spenger C, Lovestone S (2010) Association of plasma clusterin concentration with severity, pathology, and progression in Alzheimer disease. *Arch Gen Psychiatry* **67**, 739-748.

[33] Thambisetty M, An Y, Kinsey A, Koka D, Saleem M, Güntert A, Kraut M, Ferrucci L, Davatzikos C, Lovestone S, Resnick SM (2012) Plasma clusterin concentration is associated with longitudinal brain atrophy in mild cognitive impairment. *Neuroimage* **59**, 212-217.

[34] Rye PD, Booij BB, Grave G, Lindahl T, Kristiansen L, Andersen HM, Horndalsveen PO, Nygaard HA, Naik M, Hoprekstad D, Wetterberg P, Nilsson C, Aarsland D, Sharma P,

Lonneborg A (2011) A novel blood test for the early detection of Alzheimer's disease. *J Alzheimers Dis* **23**, 121-129.

[35] Booij BB, Lindahl T, Wetterberg P, Skaane NV, Saebo S, Feten G, Rye PD, Kristiansen LI, Hagen N, Jensen M, Bardsen K, Winblad B, Sharma P, Lonneborg A (2011) A gene expression pattern in blood for the early detection of Alzheimer's disease. *J Alzheimers Dis* **23**, 109-119.

[36] Fehlbaum-Beurdeley P, Sol O, Désiré L, Touchon J, Dantoine T, Vercelletto M, Gabelle A, Jarrige AC, Haddad R, Lemarié JC, Zhou W, Hampel H, Einstein R, Vellas B; EHTAD/002 study group(2012) Validation of AclarusDx™, a blood-based transcriptomic signature for the diagnosis of Alzheimer's disease. *J Alzheimers Dis* **32**, 169-181.

[37] Lunnon K, Ibrahim Z, Proitsi P, Lourdusamy A, Newhouse S, Sattlecker M, Furney S, Saleem M, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Coppola G, Geschwind D, Simmons A, Lovestone S, Dobson R, Hodges A, Consortium A (2012) Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood. *J Alzheimers Dis* **30**, 685-710.

[38] Lunnon K, Sattlecker M, Furney SJ, Coppola G, Simmons A, Proitsi P, Lupton MK, Lourdusamy A, Johnston C, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Geschwind D, Lovestone S, Dobson R, Hodges A (2013) A blood gene expression marker of early Alzheimer's disease. *J Alzheimers Dis* **33**, 737-753.

[39] 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.

[40] Freedman M (1994) *Clock drawing: A neuropsychological analysis*, Oxford University Press, New York.

[41] Scanlan JM, Brush M, Quijano C, Borson S (2002) Comparing clock tests for dementia screening: Naive judgments vs formal systems - what is optimal? *Int J Geriatr Psychiatry* **17**, 14-21.

[42] Kendrick DC (1985) *Kendrick cognitive test for the elderly*, The NFER-NELSON Publishing Company Ltd, Windsor, United Kingdom.

[43] Welsh K, Butters N, Hughes J, Mohs R, Heyman A (1991) Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. *Arch Neurol* **48**, 278-281.

[44] Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Mellits ED, Clark C (1989) The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* **39**, 1159-1165.

[45] Brown EC, Casey A, Fisch RI, Neuringer C (1958) Trail making test as a screening device for the detection of brain damage. *J Consult Psychol* **22**, 469-474.

[46] Morris JC (1993) The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology* **43**, 2412-2414.

[47] 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.

[48] 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.

[49] Langsrud Ø (2002) 50-50 Multivariate analysis of variance for collinear responses. *Statistician* **51**, 305-317.

[50] Dasarathy BV (1991) *Nearest neighbor (NN) norms: NN pattern classification techniques*, IEEE Computer Society Press, Los Alamitos, California.

[51] Gidskehaug L, Anderssen E, Flatberg A, Alsberg BK (2007) A framework for significance analysis of gene expression data using dimension reduction methods. *BMC Bioinformatic* **8**, 346.

[52] Team RDC (2008) *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria.

[53] Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* **57**, 289-300.

[54] Iqbal K, Flory M, Khatoon S, Soininen H, Pirttila T, Lehtovirta M, Alafuzoff I, Blennow K, Andreasen N, Vanmechelen E, Grundke-Iqbal I (2005) Subgroups of Alzheimer's disease based on cerebrospinal fluid molecular markers. *Ann Neurol* **58**, 748-757.

[55] Burns LC, Minster RL, Demirci FY, Barmada MM, Ganguli M, Lopez OL, DeKosky ST, Kamboh MI (2011) Replication study of genome-wide associated SNPs with late-onset Alzheimer's disease. *Am J Med Genet B Neuropsychiatr Genet* **156**, 507-512.

[56] Reitz C, Mayeux R (2009) Use of genetic variation as biomarkers for Alzheimer's disease. *Ann N Y Acad Sci* **1180**, 75-96.

[57] Parnetti L, Chiasserini D, Eusebi P, Giannandrea D, Bellomo G, De Carlo C, Padiglioni C, Mastrocola S, Lisetti V, Calabresi P (2012) Performance of Aβ1-40, Aβ1-42, total tau and phosphorylated tau as predictors of dementia in a cohort of patients with mild cognitive impairment. *J Alzheimers Dis* **29**, 229-238.

[58] Zhang D, Wang Y, Zhou L, Yuan H, Shen D, ADNI (2011) Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage* **55**, 856-867.

[59] Buchhave P, Minthon L, Zetterberg H, Wallin ÅK, Blennow K, Hansson O (2012) Cerebrospinal fluid levels of β-amyloid 1-42, but not of tau, are fully changed already 5 to 10 years before the onset of Alzheimer dementia. *Arch Gen Psychiatry* **69**, 98-106.

[60] Gomar JJ, Bobes-Bascaran MT, Conejero-Goldberg C, Davies P, Goldberg TE (2011) Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to Alzheimer disease in patients in the Alzheimer's Disease Neuroimaging Initiative. *Arch Gen Psychiatry* **68**, 961-969.