Featured Article

Predicting progression to dementia in persons with mild cognitive impairment using cerebrospinal fluid markers

Ron L. H. Handels[a,*], Stephanie J. B. Vos[a], Milica G. Kramberger[b], Vesna Jelic[c], Kaj Blennow[d], Mark van Buchem[e], Wiesje van der Flier[f,g], Yvonne Freund-Levi[c,h], Harald Hampel[i,j], Marcel Olde Rikkert[k], Ania Oleksik[e], Zvezdan Pirtosek[b], Philip Scheltens[f], Hilkka Soininen[l], Charlotte Teunissen[m], Magda Tsolaki[n], Asa K. Wallin[o], Bengt Winblad[c], Frans R. J. Verhey[a], Pieter Jelle Visser[a,f]

[a]Department of Psychiatry and Neuropsychology, Alzheimer Centre Limburg, School for Mental Health and Neurosciences, Maastricht University, Maastricht, The Netherlands
[b]Department of Neurology, Ljubljana University Medical Centre, Ljubljana, Slovenia
[c]Division of Clinical Geriatrics, Department of NVS, Karolinska Institutet, Center for Alzheimer Research, Division of Neurogeriatrics, Huddinge, Sweden
[d]Clinical Neurochemistry Laboratory, Institute of Neuroscience and Physiology, The Sahlgrenska Academy at University of Gothenburg, Sahlgrenska University Hospital, Mölndal, Sweden
[e]Department of Radiology, Leiden University Medical Centre, Leiden, The Netherlands
[f]Alzheimer Centre and Department of Neurology, Neuroscience Campus Amsterdam, VU University Medical Centre, Amsterdam, The Netherlands
[g]Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands
[h]Department of Psychiatry, Tiohundra AB Norrtälje Hospital, Stockholm Sweden
[i]AXA Research Fund & UPMC Chair, Paris, France
[j]Sorbonne Universités Pierre et Marie Curie (UPMC) Paris 06, Inserm, CNRS, Institut du cerveau et de la moelle (ICM), Département de Neurologie, Institut de la Mémoire et de la Maladie d'Alzheimer (IM2A), Hôpital Pitié-Salpêtrière, Boulevard de l'hôpital, Paris, France
[k]Department of Geriatrics, Radboudumc Alzheimer Centre, Donders Institute for Brain Cognition and Behavior, Radboud University Medical Centre, Nijmegen, The Netherlands
[l]Department of Neurology, Institute of Clinical Medicine, University of Eastern Finland, Neurocenter–Neurology, Kuopio University Hospital, Kuopio, Finland
[m]Neurochemistry Laboratory and Biobank, Department of Clinical Chemistry, VU University Medical Centre, Amsterdam, The Netherlands
[n]Memory and Dementia Outpatient Clinic, 3rd Department of Neurology, Medical School, Aristotle University of Thessaloniki, Thessaloniki, Greece
[o]Clinical Memory Research Unit, Department of Clinical Sciences, Lund University, Malmö, Sweden

Abstract

Introduction: We aimed to determine the added value of cerebrospinal fluid (CSF) to clinical and imaging tests to predict progression from mild cognitive impairment (MCI) to any type of dementia.

Methods: The risk of progression to dementia was estimated using two logistic regression models based on 250 MCI participants: the first included standard clinical measures (demographic, clinical, and imaging test information) without CSF biomarkers, and the second included standard clinical measures with CSF biomarkers.

Results: Adding CSF improved predictive accuracy with 0.11 (scale from 0–1). Of all participants, 136 (54%) had a change in risk score of 0.10 or higher (which was considered clinically relevant), of whom in 101, it was in agreement with their dementia status at follow-up.

Discussion: An individual person's risk of progression from MCI to dementia can be improved by relying on CSF biomarkers in addition to recommended clinical and imaging tests for usual care.
© 2017 the Alzheimer's Association. Published by Elsevier Inc. All rights reserved.

Keywords: Dementia; Alzheimer's disease; Mild cognitive impairment; Prognosis; Risk; Progression; Predict; Conversion; Reclassification; Risk prediction model

*Corresponding author. Tel.: +31 43 3881036; Fax: +31 43 3884092.
E-mail address: ron.handels@maastrichtuniversity.nl

1. Introduction

Diagnostic research criteria for Alzheimer's disease (AD) have recommended the use of cerebrospinal fluid (CSF) biomarkers to determine etiology and prognosis in persons with mild cognitive impairment (MCI) [1–3]. Because the recommended clinical diagnostic workup already contains much information to identify the risk of dementia progression, it is important to estimate the added value of CSF biomarkers for AD, relative to clinical assessment and brain imaging.

Previous research [4–10] has indicated the increased accuracy when using CSF measures in addition to neuropsychological tests or magnetic resonance imaging (MRI) atrophy scores to predict progression to AD-type dementia in persons with MCI. However, their generalizability to clinical practice was limited for three reasons. First, most relied on odds ratios, receiver operating characteristic (ROC) curves, and sensitivity and specificity outcome measures. These measures typically reflect the performance of a model to test if improvements are significant and valuable for research purposes [11]. However, it is difficult to determine whether significant improvements in such performance indicators are clinically relevant. Therefore, alternative outcomes measures have been developed such as the reclassification table or the reclassification index. These measures distinguish between false-positive and false-negative outcomes, which likely are differently weighted in clinical practice [11]. Second, the previous research limitedly reflected the test information that is available from the standard usual care diagnostic workup. Some of the previous research [3,7] did not consider the degree of cognitive performance on neuropsychological test results reflecting the degree of cognitive performance and did not focus on the added value or were mainly for methodological purposes, respectively. Other previous research [4,5,9] selected the best model based on statistical significance omitting information available in standard practice (such as demographics and neuropsychological tests). Other previous research [6,8] only analyzed CSF in addition to neuropsychological test or CSF in addition to MRI. Omitting information from the standard diagnostic workup reduces the generalizability to practice and possibly overestimated the added value of CSF because nonsignificant measures could still contribute as covariates to the overall predictive value of a model when applied in practice. Third, all previous research focused on progression to AD-type dementia while progression to other dementia subtypes is also relevant in clinical practice.

To enable the translation of findings on CSF biomarkers in the research setting to clinical practice, we approached CSF biomarkers as a risk factor to predict individual risks of progression from MCI to any-type dementia in addition to measures available in usual care diagnostic workup. We aimed to determine the added clinical value of CSF biomarkers relative to clinical and imaging tests that are recommended in usual care, to predict progression to dementia.

2. Methods

2.1. Study participants

We selected participants with MCI who consecutively attended a memory clinic from a range of cohorts in various European countries: the Development of Screening Guidelines and Clinical Criteria for Predementia AD (DESCRIPA) multicenter study [12] (inclusion between 2003 and 2005, with an additional sample of participants seen outside the DESCRIPA inclusion period at one of the sites VU medical center (VUmc), inclusion between 1998 and 2007 [13]), LEARN multicenter study [14] (inclusion between 2009 and 2011), Ljubljana University Medical Centre [15] (inclusion between 2011 and 2014), and Karolinska University Hospital Huddinge memory clinic [16] (inclusion between 2007 and 2011).

Eligibility criteria for each cohort separately are described elsewhere [4,12,14,15,17]. Inclusion criteria for the present study were new referral to a memory clinic because of cognitive complaint; age 50 years or older; baseline diagnosis of MCI; baseline data of CSF markers of amyloid β 1–42 ($A\beta_{1-42}$), total tau (t-tau), and phosphorylated tau (p-tau) levels; at least one follow-up measurement with information on progression to dementia; and no diagnosis of a somatic psychiatric or neurological disorder that might have caused the cognitive impairment at baseline. We excluded two participants with CSF t-tau values more than five times the absolute deviation to the median t-tau in the sample, leaving 250 participants for the analyses.

Local ethical committees approved the studies, and all participants provided informed consent to use their data.

MCI was either diagnosed by a clinician according to the criteria as applied in usual practice (Petersen [18] for LEARN and Ljubljana, and Winblad [19] for Karolinska) or by a researcher using the criteria by Petersen [18] operationalized as a score lower than −1.5 standard deviation on standardized neuropsychological examination results (DESCRIPA and VUmc sample). Not all persons who assessed the diagnosis were blind for the CSF analyses as part of the CSF results were used for clinical purposes.

2.2. Clinical measures

Clinical measures were selected when recommended in clinical guidelines [20] and when available to the authors. Demographic information included age, gender, and years of education. Overall cognition was measured by the Mini–Mental State Examination (MMSE). Memory performance was measured by delayed recall of a Word Learning Test (WLT) (Rey Auditory Verbal Learning Test [RAVLT] [21] for the Karolinska, LEARN, and VUmc samples; the California Verbal Learning Test [22] for the Ljubljana sample; and for the DESCRIPA sample, the RAVLT and CERAD [23] were key tests). Raw scores were transformed to a z-score by adjusting them for age, gender, and/or education using norm scores from healthy control cohorts outside this

study. Using norm scores ensured, the different scales were harmonized to the same outcome. Depressive symptoms were measured by various scales and were categorized into three groups as "mild" if $14 \leq$ Beck Depression Inventory ≤ 19 or $5 \leq$ Geriatric Depression Scale ≤ 8 or $8 \leq$ Hamilton Depression Rating Scale ≤ 14 or $7 \leq$ Cornell Scale for Depression ≤ 9 or $7 \leq$ Montgomery Asberg Depression Rating Scale ≤ 19; "no" if lower than these ranges; and "high" if higher than these ranges.

2.3. Imaging measures

All participants underwent MRI at baseline. A qualitative rating of medial temporal lobe atrophy (MTA) was performed on the MRI [24] with the scores of the right and left side summed (0–8).

2.4. CSF measures

CSF samples were obtained at baseline by lumbar puncture, stored between $-80°C$ and $-70°C$ in polypropylene tubes, and analyzed in the laboratory of the VU Medical Centre (VUmc center cohort and LEARN), the laboratory in Gothenburg (DESCRIPA study), the laboratory in Karolinska University Hospital Huddinge (for the Karolinska sample), and the Laboratory in Ljubljana University Medical Center (for the Ljubljana sample). CSF $A\beta_{1-42}$, t-tau, and p-tau concentrations were measured in all laboratories using commercially available sandwich ELISAs (Innotest $A\beta_{1-42}$; Innotest hTAU-Ag; Innogenetics, Ghent, Belgium) in a single batch within each of the four laboratories. CSF procedures of each cohort are described in detail elsewhere [4,16,25].

2.5. Outcome measures

The main outcome measure was clinical diagnosis of dementia at follow-up (mean follow-up 26 months) according to the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision* (DSM-IV-TR): Diagnostic and statistical manual of mental disorders [26]. Etiology was diagnosed using NINCDS-ADRDA criteria [27] for AD, NINCDS-AIREN criteria [28] for vascular dementia, McKeith criteria [29] for Lewy body dementia, and Neary criteria [30] for frontotemporal lobe dementia. This diagnosis was set by an experienced specialist in dementia during each of the annual clinical follow-up measurement moments in which the participant's physical, neurological, psychiatric, and neuropsychological status was evaluated based on the same clinical measurement scales as at baseline (often except for MRI and CSF tests).

2.6. Statistical analysis

Missing values occurred on years of education ($n = 3$, 1.2%), MMSE ($n = 2$, 0.8%), WLT ($n = 44$, 17.6%), depression ($n = 51$, 20.4%), and MRI MTA ($n = 27$, 10.8%).

Students' *t*-tests revealed that MMSE, $A\beta$, t-tau, and p-tau significantly differed between cases with and without missing WLT scores, that $A\beta$ significantly differed between cases with and without missing depressive symptom scores, and that WLT significantly differed between cases with and without missing MTA scores. Therefore, missing data were not missing completely at random and were imputed using multiple imputation in the statistical software package STATA12 to generate 10 imputed data sets. This was done using chained equations—for each variable, a linear regression model was built using all other variables as predictors. In an iterative approach, each missing value was replaced with a predicted value. To ensure each predicted value was within the plausible distribution, the predicted value was matched to a closest observed value within the sample. The latter was done including random variation to ensure each imputed value varied across the 10 imputed data sets. This increased the uncertainty of the outcomes due to the uncertainty of the actual value of the missing data. Visual evaluation of the convergence plots and histograms indicated valid imputed values following the same distribution as in the observed data (see Supplemental Material 1).

Two logistic regression prediction models were developed using statistical software package R version 3.2.5 [31] to determine the risk of progression to any type of dementia. In the first model, we included all available standard clinical measures (i.e., age, gender, education, MMSE, WLT, depressive symptoms, and MRI MTA) without CSF biomarkers. In the second model, we included the same available standard clinical measures with the CSF biomarkers concentrations on $A\beta_{1-42}$, t-tau, and p-tau on a continuous scale. For each participant, we calculated a risk score of developing any-type dementia, based on the predicted values by the logistic regression model without CSF as well as based on the model with CSF. A memory clinic center was included as a random intercept to correct for center differences. The difference between the two models (meaning, adding the three CSF markers as a block to the first model) was tested for significance using a Wald test in the statistical software package STATA12.

Model performance was assessed by the concordance index (c-index, which equals the area under the ROC curve; confidence interval was estimated using the "rcorrcens" function from the "Hmisc" package (version 3.17-4) in the statistical software package R). The change in predicted probability of progression to dementia was calculated by subtracting the predicted probability based on the first model (a value between 0 and 1) from the predicted probability based on the second model. A positive difference reflected an increased risk due to CSF test outcomes, whereas a negative value reflected a decreased risk due to CSF test outcomes. Participants who had an increased risk and in whom progression to dementia was observed within the follow-up period were labeled as "correctly reclassified," as well as the participants who had a decreased risk and in whom no progression to dementia was observed. The Net

Reclassification Index (NRI) was assessed separately for participants who did convert to dementia (i.e., events) and who did not convert to dementia (i.e., nonevents). This index represents the sum of the following two proportions: the proportion of individuals with an increased risk minus the proportion with a decreased risk for those with a positive outcome, and the proportion of individuals with a decreased risk minus the proportion with an increased risk for those with a negative outcome [32]. The confidence interval was estimated using the "reclassification" function from the "PredictABEL" package (version PredictABEL_1.2-2) in the statistical software package R.

Both prediction models were validated by visually assessing the relation between the observed and predicted proportion of progression to dementia [33,34] (see Supplemental Material 2). Furthermore, a bootstrap validation was performed. First, 250 participants were sampled with replacement, and the two models without and with CSF were fitted to this sampled data. Second, predicted values were calculated using this bootstrap-based model on the sampled data and on the original data. The performance difference between them was averaged over 6500 bootstrap samples to produce a shrinkage factor, which can be subtracted from the original model's performance in the original data to obtain an optimism-corrected estimate of performance.

In addition to the single measures for model performance (c-index, proportion correct reclassification, and NRI), we generated a reclassification table [35] to provide a more detailed overview of reclassification. Because specific predefined risk categories to guide clinical management upon have not been reported as far as the authors know, the changes in risk scores were cut in quintiles. However, categories based on quintiles are arbitrary because, for example, a change from 0.19 to 0.21 could result into a reclassification while it has most probably no effect on clinical management or impact on the patient. Therefore, a minimum risk change of 0.10 was assumed clinically relevant. As an alternative to the reclassification table, the change in predicted probability was classified and tabulated.

3. Results

Two hundred fifty participants met the inclusion criteria. Table 1 presents the baseline characteristics. The participants were on average 68.6 years old (standard deviation = 7.5) and consisted of 111 (44%) females. Median follow-up time for the 151 (60%) nonconverters was 24 months (mean = 26; range 7–87) and 24 months for the 99 (40%) dementia converters (mean = 27; range 4–130). The proportion of progression to AD-type dementia among the converters was 87%, for vascular dementia (VAD) 4%, for dementia of lewy bodies (DLB) 3%, for frontotemporal dementia (FTD) 5%, and for other types of dementia 1%.

Table 2 summarizes the univariate logistic regression results and the results of the multivariate prediction model without and with CSF. In the model without CSF, female gender, lower

Table 1
Baseline characteristics and number of observations from 250 participants before imputation of the sample

Characteristic	Mean (SD) or n (%)	Range	% Missing values
Age	68.6 (7.5)	52.0 to 89.0	0
Female gender	111 (44%)	n/a	0
Education years	11.2 (3.1)	4 to 18	1
MMSE	26.9 (2.4)	18 to 30	1
WLT delayed recall z-score	−1.6 (1.2)	−5.6 to 2.3	18
Depression	0.3 (0.6)	0 to 2	20
MRI MTA left and right	2.3 (1.8)	0 to 6	11
CSF Aβ$_{1-42}$ levels, pg/mL	666 (296)	157 to 1538	0
CSF t-tau levels, pg/mL	446 (237)	52 to 1179	0
CSF p-tau levels, pg/mL	71 (30)	16 to 172	0
Conversion to dementia	99 (40%)	n/a	0
AD	86 (87%)	n/a	0
VAD	4 (4%)	n/a	0
DLB	3 (3%)	n/a	0
FTD	5 (5%)	n/a	0
Other	1 (1%)	n/a	0
Time to conversion (months)	26.5 (14.3)	4 to 130.4	0

Abbreviations: SD, standard deviation; n, number of participants; MMSE, Mini–Mental State Examination; WLT, Word Learning Test; MRI, magnetic resonance imaging; MTA, medial temporal lobe atrophy; CSF, cerebrospinal fluid; n/a, not applicable; AD, Alzheimer's disease; p-tau, phosphorylated tau; t-tau, total tau; VAD, vascular dementia; DLB, dementia of lewy bodies; FTD, frontotemporal dementia.

MMSE, and higher MTA scores significantly predicted progression to any-type dementia. The model including CSF Aβ$_{1-42}$, t-tau, and p-tau markers was significantly better than the model without the three CSF markers ($P < .001$; F = 12.93).

The concordance index of the model without CSF was 0.741 (95% confidence interval [CI]: 0.679–0.802) and with CSF 0.851 (95% CI: 0.805–0.896) reflecting an improved accuracy of 0.11 on a scale from 0 to 1.

Each participant's risk of progression to any-type dementia using the model without CSF and the model with CSF is shown in Fig. 1. The proportion correct reclassifications in participants who converted to dementia was 0.70 (95% CI: 0.64–0.76) and in participants who did not convert to dementia was 0.72 (95% CI: 0.66–0.78). This indicates that of the 99 participants who progressed to dementia (represented by both filled green circles and filled red triangles in Fig. 1), 69 (70%) participants received a higher risk after CSF testing compared with those received risk before CSF testing (represented by filled green circles in Fig. 1). For the 151 participants who did not progress to dementia (open green circles and open red triangles in Fig. 1), 109 (72%) received a lower risk after CSF testing (open green circles in Fig. 1). Furthermore, 54% of the participants' risk scores changed more than 0.10, which is reflected by the fact that most reclassifications fell outside the gray-shaded area around the diagonal in Fig. 1.

The NRI was 0.40 (95% CI: 0.26–0.55) within the participants who converted to dementia and 0.09 (95% CI: −0.01 to 0.19) within the participants who did not convert to dementia, when classifying risks into categories of quintiles

Table 2
Logistic regression results for any-type dementia prediction using the model without and with CSF biomarkers

Predictor	Univariate		Multivariate model without CSF		Multivariate model with CSF	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Intercept	n/a	n/a	0.33 (0.00–74.04)	.686	10.61 (0.01–7617.88)	.481
Age (years)	1.05 (1.01–1.09)	.015	1.03 (0.98–1.08)	.249	0.99 (0.94–1.05)	.764
Female gender	1.94 (1.12–3.35)	.018	2.15 (1.12–4.13)	.021	1.48 (0.70–3.12)	.299
Education (years)	1.03 (0.94–1.12)	.531	1.10 (0.99–1.22)	.073	1.06 (0.94–1.19)	.367
MMSE	0.83 (0.73–0.94)	.003	0.84 (0.73–0.98)	.025	0.84 (0.70–1.00)	.046
Word learning test (z-score)	0.69 (0.53–0.91)	.007	0.76 (0.56–1.04)	.084	0.82 (0.57–1.20)	.307
Depression	0.84 (0.49–1.44)	.535	0.95 (0.51–1.74)	.857	0.99 (0.52–1.89)	.970
MRI MTA left and right	1.43 (1.20–1.71)	.000	1.39 (1.12–1.71)	.003	1.49 (1.16–1.91)	.002
CSF Aβ$_{1-42}$	1.00 (0.99–1.00)	.000	n/a	n/a	1.00 (1.00–1.00)	.000
CSF t-tau	1.00 (1.00–1.01)	.000	n/a	n/a	1.00 (1.00–1.00)	.453
CSF p-tau	1.03 (1.02–1.04)	.000	n/a	n/a	1.02 (0.99–1.04)	.219

Abbreviations: CSF, cerebrospinal fluid; OR, odds ratio; CI, confidence interval; MMSE, Mini–Mental State Examination; MRI, magnetic resonance imaging; MTA, medial temporal lobe atrophy; n/a, not applicable; t-tau, total tau; p-tau, phosphorylated tau.

(see Table 3). This indicates that of the 99 participants who progressed to dementia, 55 were correctly classified in a higher risk category after CSF testing and 15 were incorrectly classified in a lower risk category after CSF testing, resulting in 40 (40% of all participants who progressed) net correctly reclassified participants. For the 151 participants who did not progress to dementia, 37 were correctly and 23 incorrectly reclassified, resulting in 14 (9% of all
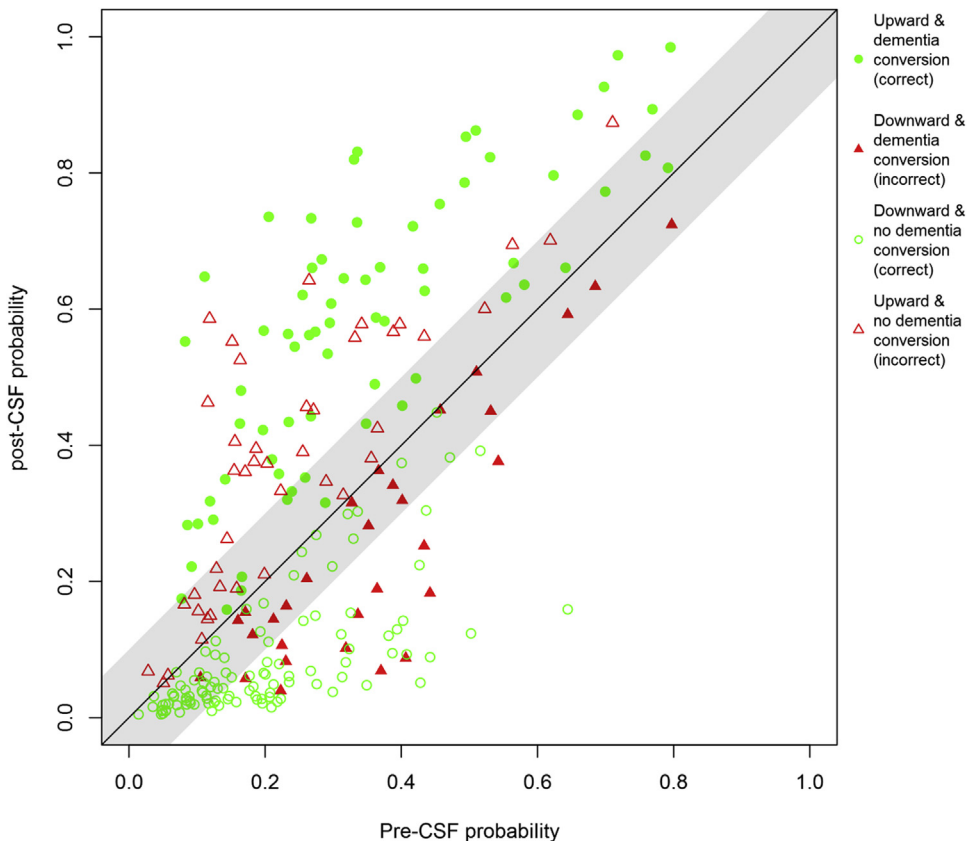

Fig. 1. Dementia risk score of each participant based on the logistic regression model without CSF (x-axis) and with CSF (y-axis) in addition to standard clinical measures. Participants in whom the predicted risk increased with CSF, compared to the risk without CSF, are presented in the top left of the diagonal line, and in whom the risk decreased in the bottom right. Filled dots represent participants who converted to dementia at follow-up and open dots participants who did not convert. Increased risks by adding CSF were considered correct if the participant converted (filled green circle) and incorrect if the participant did not convert (open red triangle), and vice versa for decreased risks. The gray area indicates a 10% or less change in pre- and post-CSF dementia risk. Abbreviations: CSF, cerebrospinal fluid.

Table 3
Reclassification table representing the number of participants in each risk category based on the risk prediction model (rows) without CSF and the risk prediction model with CSF (columns) in addition to standard clinical measures for the total group (A) and split up into nonconverters (B) and converters (C)

| | Dementia risk classification with CSF | | | | | | |
Dementia risk classification without CSF	QU1	QU2	QU3	QU4	QU5	Reclassified n (%)	Correctly reclassified n (%)
(A) Reclassification in the total sample							
QU1	83	14	10	1	0	25 (23)	13 (52)
QU2	35	24	19	11	2	67 (74)	50 (75)
QU3	7	8	7	10	3	28 (80)	21 (75)
QU4	1	0	1	6	8	10 (63)	8 (80)
QU5	0	0	0	0	0	n/a	n/a
(B) Reclassification in the sample of subjects not converted to dementia							
QU1	75	7	5	0	0	12 (14)	0 (0)
QU2	26	13	7	1	0	34 (72)	26 (76)
QU3	5	5	2	2	0	12 (86)	10 (83)
QU4	1	0	0	1	1	2 (67)	1 (50)
QU5	0	0	0	0	0	n/a	n/a
(C) Reclassification in the sample of subjects converted to dementia							
QU1	8	7	5	1	0	13 (62)	13 (100)
QU2	9	11	12	10	2	33 (75)	24 (73)
QU3	2	3	5	8	3	16 (76)	11 (69)
QU4	0	0	1	5	7	8 (62)	7 (88)
QU5	0	0	0	0	0	n/a	n/a

Abbreviations: CSF, cerebrospinal fluid; n, number of participants; QU1 to QU5 represent the five quintiles with probabilities [0%, 20%], [20%, 40%], [40%, 60%], [60%, 80%), and [80%, 100%], respectively; n/a, not applicable.

NOTE. The last two columns present the total number and percentage of reclassifications and the proportion of correct reclassifications (i.e., reclassified into higher risk category and observed dementia progression, or reclassified into a lower risk category and no observed dementia progression). For example, 91 participants received a risk score between 0.20 and 0.40. Of them, 67 (74%) were reclassified of whom 35 to a lower risk category (26 eventually converted and 9 did not) and 32 to a higher risk category (24 eventually converted and 8 did not).

participants who did not progress) net correct reclassified participants.

Table 3 summarizes the number of reclassifications of risk categories (i.e., quintiles) after using the prediction model with CSF compared with the prediction model without CSF for the total sample and for the subgroup of participants who did and did not convert to dementia. Participants with a risk score between 0.40 and 0.60 (i.e., third quintile) were most often reclassified (80%). Reclassifications were most often correct (80%) in participants with a risk score between 0.60 and 0.80.

Table 4 describes the proportions of participants in each category of risk change from the prediction model without CSF to the prediction model with CSF. Of all participants 136 (54%) had a change in risk score of 0.10 or higher, which was considered clinically relevant. Of these 136, 52 correctly received an increased risk after CSF (because they were converted to dementia at follow-up), and 49 correctly received a decreased risk. Respectively, 23 incorrectly received an increased risk after CSF because they did not convert to dementia at follow-up, and 12 incorrectly received a decreased risk.

4. Discussion

We evaluated the added value of CSF biomarkers to a prognostic prediction model consisting of demographic,

clinical, and imaging information for progression to any-type dementia in a memory clinic sample of 250 participants with MCI. Our results indicated that adding CSF biomarkers to the usual care diagnostic workup improved the ability to differentiate between those who do and do not progress to dementia with 0.11 on a scale from 0 to 1 (concordance

Table 4
Participants categorized by their change in risk score between the prediction model without CSF and the prediction model with CSF and the proportion of progression to any-type dementia

Risk change	n (% of total sample)	n, conversion	n, no conversion
Increased risk of dementia conversion			
0% ≤ change < 10%	36 (14%)	17 (47%)	19 (53%)
10% ≤ change < 25%	41 (16%)	23 (56%)	18 (44%)
25% ≤ change < 50%	32 (13%)	27 (84%)	5 (16%)
50% ≤ change ≤ 100%	2 (1%)	2 (100%)	0 (0%)
Total clinically relevant (change ≥ 10%)	75 (30%)	52 (69%)	23 (31%)
Decreased risk of dementia conversion			
0% ≤ change < 10%	78 (31%)	18 (23%)	60 (77%)
10% ≤ change < 25%	46 (18%)	9 (20%)	37 (80%)
25% ≤ change < 50%	15 (6%)	3 (20%)	12 (80%)
50% ≤ change ≤ 100%	0 (0%)	n/a	n/a
Total clinically relevant (change ≥ 10%)	61 (24%)	12 (20%)	49 (80%)

Abbreviations: CSF, cerebrospinal fluid; n, number of participants; n/a, not applicable.

index: 0.74–0.85; 0.50 indicates no predictive ability, 1.00 indicates full predictive ability). In 40% of all participants, the dementia risk score changed and showed a correct clinically relevant increased or decreased risk score of 0.10 (on a risk scale of 0–1) or more after CSF biomarkers were added to the prediction model. However, 26% of the participants were incorrectly reclassified with a risk change of 0.10 or more, which might result into incorrect prognostic disease management.

The addition of CSF markers to the model resulted mostly in reclassifications of individuals who had a midrange risk score for dementia (between 0.20 and 0.60). If one adopts the assumption that persons with an improved individual risk score experience less feelings of uncertainty because they have better knowledge on what to expect in the future, then CSF can to, a limited extent, improve a person's well-being related to certainty. For example, addition of CSF markers was, in particular, useful to identify persons with a low risk for short-term decline because in a large group of participants, the post-CSF risk reduced and fell between 0.00 and 0.10.

After using CSF biomarkers, 54% of all participants received a dementia risk change that was considered clinically relevant. Still, 46% of the risk changes caused by CSF were lower than 0.10, which was considered clinically irrelevant. This urges a careful selection in performing a lumbar puncture for clinical purposes to ensure relevant risk changes and prevent unnecessary exposure to lumbar puncture adverse events. From the results of this study, a clinically relevant and correct risk change occurred most often in participants with a pre-CSF risk between 20% and 60%. Furthermore, post hoc analyses identified various subgroups of clinically relevant and correct risk changes, using Students' t-test and chi-square tests ($P < .05$). Participants with a risk change in risk score of ≥ 0.10 had a significantly higher age (70 vs. 68), lower MMSE (26 vs. 28), and higher MTA score (2.6 vs. 1.9) compared to participants with a change in risk score lower than 0.10. Participants with a clinically relevant correct reclassification had a significantly lower age (69 vs. 72) and were more likely female (56% vs. 29%) compared to participants with a clinically relevant incorrect reclassification. These identified subgroups, however, only show subtle differences. Therefore, future research should more precisely identify subgroups of participants with correct reclassifications based on sets of multiple characteristics, for example, using cluster analysis such as by van Giessen et al. [36].

Various studies have evaluated the predictive value of CSF in addition to clinical markers and MRI in participants with MCI using a prediction model [4–8]. All reported a significant improvement in predictive accuracy when relying on CSF biomarkers. However, as explained in Section 1, the models used in these studies omit test information from the standard diagnostic workup making their results not generalizable to medical practice. In addition,

our study showed 54% relevant reclassifications of which 75% correctly. This detailed information enables clinicians to judge whether the benefits outweigh the harms (46% irrelevant changes and 26% of all relevant changes incorrect, as well as possible lumbar puncture side effects). Our study also differed by using dementia risk predictions on a continuous scale allowing more subtle reclassifications instead of cutoff-based dichotomized risk classifications because we considered a risk change of 10% or more as clinically relevant. Because the results indicated an added value of CSF biomarkers on this risk, it suggests that the common search for optimal cutoff scores in biomarker research to dichotomize persons as either positive or negative undervalues the complexity and uncertainty in actual practice.

4.1. Limitations

The sample size was moderately large and included 99 dementia progression events for 10 predictor variables, which could have led to an overfitted model. Therefore, our results must be validated in an external cohort. Nevertheless, participants were recruited from different centers, over a large time period (1998–2014) and using different inclusion criteria, making it generalizable to this heterogeneous population. However, generalizability to old age might be limited as the average age was 69 years, and diagnostic performance has been shown to decrease with age [37], and a 40% a priori risk is higher than other observed proportions [38].

The average follow-up period of 26 months may have been too short to observe all progressions to AD-type dementia (lowering the probability of progression in CSF positive) and also too short to observe other-type dementia progression (also lowering the probability of progression in CSF negative). Furthermore, the follow-up duration might impact reclassifications due to CSF. Therefore, these risk predictions are only generalizable to the study's average follow-up period of 26 months. To check the stability of the results, the model performance was analyzed in a subsample of participants with at least 12-month follow-up (95%). All model performance outcomes were similar to the original analyses and did not change the conclusions of this study (the largest change was the NRI in participants who did not convert to dementia and was 0.23 instead of 0.25 in the original analyses). Post hoc analysis on participants with at least 24-month follow-up (60%) also resulted in a significant added value of CSF biomarkers, although there was a smaller improvement in c-index.

Data on clinical history, neurological and physical examination, activities of daily living, comorbidities, and blood tests were not available for some of the data samples. This limited the extent to which this study reflects clinical practice and could have led to an overestimation of the added value of CSF biomarkers as the missing clinical

and diagnostic variables could increase the predictive value of the prediction model without CSF. A study by Handels et al. [39] found no added value of CSF biomarkers when added to patient vignettes presenting all test information from the standard diagnostic workup in a memory clinic. They simulated prognostic decision making by a panel of clinical experts instead of using a risk prediction model and concluded that the standard workup probably appeared already sufficient for them to predict short-term decline fairly accurately.

The results of this study relied on the assumption that a risk change of 0.10 was clinically relevant, over the full range of prior probabilities (between 0 and 1). Possibly, the relevance of a change is dependent on its prior probability (e.g., a risk change from 0.15 to 0.05 might not change the clinical conclusion and management, whereas a risk change from 0.40 to 0.50 does change the clinical conclusion and management).

Both the timing of the dementia progression and the possible positive and negative effects of prognostic information on patient-important health outcomes were not taken into account. Future studies, especially on patient-important outcomes, are required to decide on the position of CSF in clinical practice, as well as improvements due to standardization and harmonization [40]. Studies on patient-important outcomes are ideally performed in a randomized controlled setting measuring the impact of communicating the diagnostic and prognostic conclusion on patient well-being [41]. To enable cost-effectiveness studies, we added Supplemental Material 3 for decision analytic modeling.

This study followed the argument from the Alzheimer's Disease International report [42] that regardless of the cause, it is the dementia syndrome which is most relevant in the domain of public health and dementia prevention. Focusing on AD-type dementia would leave a patient in uncertainty in case of negative biomarkers as the risk of progression to other-type dementia is still present. This study takes into account this risk as, for example, in case of AD-negative CSF results, progression to any-type dementia can still be indicated by the other markers. The study is, however, limited as imaging or biomarkers specific to other neurodegenerative diseases such as Lewy body and vascular causes were not taken into account.

4.2. Conclusion

An individual patient's risk estimation of any-type dementia progression can be improved by adding CSF biomarker information to clinical and imaging tests as recommended in usual care.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jalz.2016.12.015.

RESEARCH IN CONTEXT

- Systematic review: Publications on diagnostic or prognostic value of CSF in addition to clinical measures in MCI were searched in PubMed and in references and citations of retrieved relevant publications. Previous research has indicated increased odds on progression to AD-type dementia when using CSF biomarkers in addition to a selected set of clinical measures.

- Interpretation: In our study, we reflected clinical practice by reporting clinically relevant reclassification tables, including an extensive set of clinical measures used in standard practice and reporting on progression to any-type dementia. Our findings indicated improved classification, although this was limited to 40% of the participants.

- Future directions: The critical question remains whose prognosis improves and whether improved prognostic (and diagnostic) accuracy causes improved patient well-being. This needs to be established empirically to introduce and reimburse advanced diagnostics for AD in clinical practice.

References

[1] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. 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 2011;7:270–9.

[2] Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, et al. Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. Lancet Neurol 2014;13:614–29.

[3] Vos SJB, Verhey F, Frölich L, Kornhuber J, Wiltfang J, Maier W, et al. Prevalence and prognosis of Alzheimer's disease at the mild cognitive impairment stage. Brain 2015;138:1327–38.

[4] Vos S, van Rossum I, Burns L, Knol D, Scheltens P, Soininen H, et al. Test sequence of CSF and MRI biomarkers for prediction of AD in subjects with MCI. Neurobiol Aging 2012;33:2272–81.

[5] Ewers M, Walsh C, Trojanowski JQ, Shaw LM, Petersen RC, Jack CR, et al. Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. Neurobiol Aging 2012;33:1203–1214.e2.

[6] Heister D, Brewer JB, Magda S, Blennow K, McEvoy LK. Predicting MCI outcome with clinically available MRI and CSF biomarkers. Neurology 2011;77:1619–28.

[7] Verbeek AJM, Verbeek JFM, van Dijck JAAM, Verbeek ALM. [Predicting chance of disease: Calculation using prediction rules]. Ned Tijdschr Geneeskd 2014;158:A7041.

[8] Richard E, Schmand BA, Eikelenboom P, Van Gool WA. MRI and cerebrospinal fluid biomarkers for predicting progression to Alzheimer's disease in patients with mild cognitive impairment: A diagnostic accuracy study. BMJ Open 2013;3.

[9] Michaud TL, Kane RL, McCarten JR, Gaugler JE, Nyman JA, Kuntz KM. Risk stratification using cerebrospinal fluid biomarkers in patients with mild cognitive impairment: An exploratory analysis. J Alzheimers Dis 2015;47:729–40.

[10] Gomar JJ, Conejero-Goldberg C, Davies P, Goldberg TE. Extension and refinement of the predictive value of different classes of markers in ADNI: Four-year follow-up data. Alzheimers Dement 2014; 10:704–12.

[11] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: A framework for traditional and novel measures. Epidemiology 2010; 21:128–38.

[12] Visser PJ, Verhey FRJ, Boada M, Bullock R, De Deyn PP, Frisoni GB, et al. Development of screening guidelines and clinical criteria for predementia Alzheimer's disease. The DESCRIPA Study. Neuroepidemiology 2008;30:254–65.

[13] van Rossum IA, Vos SJB, Burns L, Knol DL, Scheltens P, Soininen H, et al. Injury markers predict time to dementia in subjects with MCI and amyloid pathology. Neurology 2012;79:1809–16.

[14] Handels RLH, Aalten P, Wolfs CAG, OldeRikkert M, Scheltens P, Visser PJ, et al. Diagnostic and economic evaluation of new biomarkers for Alzheimer's disease: The research protocol of a prospective cohort study. BMC Neurol 2012;12:72.

[15] Kramberger MG, Brezovar S, Emersic A, Rot U, Kavcic V, Slana A, et al. The role of neuropsychological assessment in predicting CSF dementia measures in mild cognitive impairment. Alzheimers Dement 2014;10:P505.

[16] Kramberger MG, Kåreholt I, Andersson T, Winblad B, Eriksdotter M, Jelic V. Association between EEG abnormalities and CSF biomarkers in a memory clinic cohort. Dement Geriatr Cogn Disord 2013; 36:319–28.

[17] Enache D, Cavallin L, Lindberg O, Farahmand B, Kramberger MG, Westman E, et al. Medial temporal lobe atrophy and depressive symptoms in elderly patients with and without Alzheimer disease. J Geriatr Psychiatry Neurol 2015;28:40–8.

[18] Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med 2004;256:183–94.

[19] Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund L, et al. Mild cognitive impairment–beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment. J Intern Med 2004;256:240–6.

[20] Waldemar G, Dubois B, Emre M, Georges J, McKeith IG, Rossor M, et al. Recommendations for the diagnosis and management of Alzheimer's disease and other disorders associated with dementia: EFNS guideline. Eur J Neurol 2007;14:e1–26.

[21] Rey A. L' examen clinique en psychologie. Oxford, England: Presses universitaires de France; 1958.

[22] Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test. Second Edition. Adult Version. ManualSan Antonio: TX: Psychological Corporation; 2000.

[23] Welsh KA, Butters N, Mohs RC, Beekly D, Edland S, Fillenbaum G, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. Neurology 1994;44:609–14.

[24] Scheltens P, Leys D, Barkhof F, Huglo D, Weinstein HC, Vermersch P, et al. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: Diagnostic value and neuropsychological correlates. J Neurol Neurosurg Psychiatry 1992;55:967–72.

[25] Vos SJB, Visser PJ, Verhey F, Aalten P, Knol D, Ramakers I, et al. Variability of CSF Alzheimer's disease biomarkers: Implications for clinical practice. PLoS One 2014;9:e100784.

[26] American Psychiatric Association. DSM-IV-TR: Diagnostic and statistical manual of mental disorders. American Psychiatric Association; 1994.

[27] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34:939–44.

[28] Román GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. Vascular dementia: Diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. Neurology 1993;43:250–60.

[29] McKeith IG, Galasko D, Kosaka K, Perry EK, Dickson DW, Hansen LA, et al. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): Report of the consortium on DLB international workshop. Neurology 1996; 47:1113–24.

[30] Neary D, Snowden JS, Gustafson L, Passant U, Stuss D, Black S, et al. Frontotemporal lobar degeneration: A consensus on clinical diagnostic criteria. Neurology 1998;51:1546–54.

[31] R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016. Available at: https://www.R-project.org. Accessed February 27, 2017.

[32] Pencina MJ, D'Agostino Sr RB, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157–72.

[33] Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis

or diagnosis (TRIPOD): The TRIPOD statement. J Clin Epidemiol 2015;68:134–43.

[34] Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. Ann Intern Med 2015;162:W1–73.

[35] Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: The role of reclassification measures. Ann Intern Med 2009;150:795–802.

[36] van Giessen A, Moons KGM, de Wit GA, Verschuren WMM, Boer JMA, Koffijberg H. Tailoring the implementation of new biomarkers based on their added predictive value in subgroups of individuals. PLoS One 2015;10:e0114020.

[37] Mattsson N, Rosén E, Hansson O, Andreasen N, Parnetti L, Jonsson M, et al. Age and diagnostic performance of Alzheimer disease CSF biomarkers. Neurology 2012;78:468–76.

[38] Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia–meta-analysis of 41 robust inception cohort studies. Acta Psychiatr Scand 2009;119:252–65.

[39] Handels RLH, Joore MA, Vos SJB, Aalten P, Ramakers IHGB, Rikkert MO, et al. Added prognostic value of cerebrospinal fluid biomarkers in predicting decline in memory clinic patients in a prospective cohort. J Alzheimers Dis 2016;52:875–85.

[40] Blennow K, Dubois B, Fagan AM, Lewczuk P, de Leon MJ, Hampel H. Clinical utility of cerebrospinal fluid biomarkers in the diagnosis of early Alzheimer's disease. Alzheimers Dement 2015;11:58–69.

[41] Ferrante di Ruffano L, Hyde CJ, McCaffery KJ, Bossuyt PMM, Deeks JJ. Assessing the value of diagnostic tests: A framework for designing and evaluating trials. BMJ 2012;344:e686.

[42] Prince M, Albanese E, Guerchet M, Prina M. World Alzheimer report 2014. London; 2014.