Diagnosis and prognosis of Alzheimer's disease in subjects with mild cognitive impairment
van Rossum, I.A.

2013

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl
2.2

Test sequence of CSF and MRI biomarkers for prediction of AD in subjects with MCI


NEUROBIOLOGY OF AGING 2012 OCT; 33 (10): 2272-81
ABSTRACT

Our aim was to identify the best diagnostic test sequence for predicting Alzheimer’s disease (AD)-type dementia in subjects with mild cognitive impairment (MCI) using cerebrospinal fluid (CSF) and MRI biomarkers.

We selected 153 subjects with MCI from a multicenter memory-clinic-based cohort. We tested the CSF Aβ1-42/tau ratio using ELISA and hippocampal volumes (HCVs) using the atlas-based LEAP method. Outcome measure was progression to AD-type dementia in 2 years.

At follow-up, 48 (31%) subjects converted to AD-type dementia. In multivariable analyses, CSF Aβ1-42/tau and HCV predicted AD-type dementia regardless of APOE genotype and cognitive scores. Test sequence analyses showed that CSF Aβ1-42/tau increased predictive accuracy in subjects with normal HCV (p<0.001) and abnormal HCV (p=0.025). HCV increased predictive accuracy only in subjects with normal CSF Aβ1-42/tau (p=0.014). Slope analyses for annual cognitive decline yielded similar results. For selection of subjects for a prodromal AD trial, the best balance between sample size and number of subjects needed to screen was obtained with CSF markers.

These results provide further support for the use of CSF and MRI biomarkers to identify prodromal AD.
INTRODUCTION

Diagnosis of Alzheimer’s disease (AD) in the stage of mild cognitive impairment (MCI) is important in order to give patients a prognosis. Moreover, disease-modifying drugs for AD might be most effective in this early course of the disease. Several biomarkers of AD are now available that can help to identify AD pathology including abnormal levels of beta amyloid (Aβ)1-42 and tau in cerebrospinal fluid (CSF) and medial temporal lobe atrophy on magnetic resonance imaging (MRI).

For clinical practice and for selection of subjects for prodromal AD trials, it is of major importance to investigate which combination of CSF and MRI biomarkers could best predict AD-type dementia in subjects with MCI. Some studies found that CSF biomarkers could best predict AD-type dementia, but others studies found that MRI biomarkers were the best predictor.

Moreover, several studies showed that a combination of these biomarkers could predict AD-type dementia in subjects with MCI more accurately than each biomarker alone. These previous studies, however, had several limitations. They had a small sample size or were performed in a highly selected research sample. Moreover, some studies included only subjects with aMCI while subjects with MCI due to AD may also present as non-amnestic MCI (naMCI). The use of a broad definition of MCI is also recommended in the new criteria for MCI due to AD of the National Institute on Aging and Alzheimer’s Association workgroup.

Our aim was to identify the best combination of CSF and MRI biomarkers to predict conversion to AD-type dementia after 2 years in a large clinically representative sample of subjects with MCI, such that the number of tests could be reduced and as a consequence reduce patients’ burden and costs. We performed test sequence analyses of CSF and MRI measures and examined how the use of different combinations of biomarkers as inclusion criterion influenced the sample size and number of subjects required for screening for a fictive trial in prodromal AD. Although conversion to AD-type dementia after a 2-year follow-up is a clinically relevant outcome, the potential of CSF and MRI biomarkers to identify prodromal AD may be underestimated as subjects may convert at longer follow-up intervals. Therefore, we additionally performed slope analyses with annual cognitive decline as outcome measure.

METHODS

Subjects

Subjects were recruited from the DESCRIPTA cohort and the Alzheimer Center of the VU University medical center (VUmc) cohort. For the present study, inclusion criteria were baseline diagnosis of MCI, age ≤85 years, availability of data on the CSF ratio of Aβ1-42 to total tau (t-tau) and automatically measured hippocampal volumes (HCVs) on MRI, and being newly referred for assessment of cognitive complaints. Exclusion criteria were diagnosis of dementia at baseline or any other somatic, psychiatric or neurological disorder that might have caused the cognitive impairment.

DESCRIPTA is a European multicenter study performed in a memory clinic setting. Subjects were recruited from 6 centers, as both MRI and CSF data were only available at these centers. Of the 207 eligible subjects enrolled between 2003 and 2005 at these centers, 74 had data for both measures.
The VUmc center was one of the DESCRIPTA centers and contributed an additional sample of subjects that were seen outside the DESCRIPTA inclusion period. Of the 217 additional eligible subjects that were included between 1998 and 2007 at this center, 91 had baseline data for both MRI and CSF. Subjects for whom both MRI and CSF data were available differed from subjects for whom MRI and CSF data were not both available with regard to age (69.3 vs. 71.2 years; p<0.001), MMSE score (26.6 vs. 27, p=0.019), and MCI type (68 vs. 59% aMCI; p=0.016) at baseline. The CSF Aβ1-42/tau ratio and HCV were similar in both groups. The total sample size was 165. The medical ethics committee at each center approved the study. All patients provided informed consent.

Clinical and cognitive assessment
Assessment of the subjects included a clinical history, medical and neurological examination, laboratory tests, functional evaluation with the clinical dementia rating scale, rating scales for neuropsychiatric symptoms, scoring on the MMSE, and neuropsychological assessment. All data were collected by investigators who were blinded to the results of biomarker analyses. Follow-up assessment was performed annually up to 5 years. Primary outcome measure was conversion to AD-type dementia after 2 years according to the DSM-IV and NINCDS-ADRDA criteria. Secondary outcome measure was annual cognitive decline on the MMSE and a cognitive composite score. Baseline diagnosis of MCI was made according to the criteria of Petersen et al. Raw scores on neuropsychological tests were corrected for age, education, and sex, in accordance with locally collected or published normative data and expressed as z-scores; the z-scores were used for further analysis. Subjects with a z-score <-1.5 SD on the learning measure or delayed recall of a word list learning test or equivalent memory test were classified as having aMCI (n=102). Subjects with a z-score <-1.5 SD on the trail making test (TMT) part A, TMT part B, verbal fluency, Rey figure copy test or equivalent test but no memory impairment were classified as having naMCI (n=48; see appendix S1). The MCI subtype diagnosis was missing for 3 subjects, as they did not undergo a neuropsychological assessment. We calculated a cognitive composite score as the average z-score of 5 tests (learning and delayed recall score of the word learning list, TMT A and B, verbal fluency, copy figures), if scores were available for at least 3 cognitive tests.

CSF analyses
CSF was collected by lumbar puncture, centrifuged, and stored at -80°C in polypropylene tubes. One sample was thawed twice but analyses without this sample revealed similar results. CSF Aβ1-42 and t-tau were measured using commercially available sandwich ELISAs (Innotest β-amyloid 1-42; Innotest hTAU-Ag; Innogenetics, Ghent, Belgium), specially constructed to measure Aβ1-42 and t-tau by experienced technicians at the lab in Gothenburg for the DESCRIPTA cohort and in Amsterdam for the VUmc cohort. We corrected for interlaboratory ELISA differences by analyzing 33 samples at both labs and we adjusted VUmc values to those of DESCRIPTA using the following formula: Gothenborg=(SD Gothenborg/SD VUmc)*VUmc+average Gothenborg−((SD Gothenborg/SD VUmc)*average VUmc). As CSF measure, we used the ratio of Aβ1-42 to t-tau, which is the best CSF biomarker according to a recent meta-analysis and accepted by the European Medical Agency for inclusion of subjects for prodromal AD trials. The CSF ratio was dichotomized by defining a cut-off point.
(1,14) that could best predict AD-type dementia after 2 years based on the Youden index from a
time-dependent ROC curve in R,30 including all subjects with CSF data in both cohorts (n=219).
In 5 subjects, CSF was obtained at follow-up. As these subjects had a normal CSF ratio score,
they were considered to have a normal CSF score at baseline as well.

MRI analyses
For the DESCRIPA cohort, subjects were scanned according to the routine MRI protocol at each site
(appendix S2). Scanning was performed at 1.0 or 1.5 T and included a three-dimensional T1-weighted
gradient echo sequence and a fast fluid attenuated inversion recovery (FLAIR) sequence.31
As measure of medial temporal lobe atrophy, we used HCV measured centrally at the
Department of Computing at Imperial College London by experienced technicians, using
LEAP.32 LEAP is an automated structural segmentation, performed by propagating multiple
manually segmented atlas images to a diverse image set in a stepwise fashion, ensuring that
image registration needed to be performed between similar images only. All labels propagated
to a target image were considered together with a structures intensity distribution to estimate
the final segmentation (e.g. supplemental figure S1). Quality control was performed by visual
inspection of the segmentation on transverse, coronal and sagittal slices. HCVs were normalized
to intracranial volume using MNI scaling. The sum of left and right HCV was dichotomized by
defining a cut-off (5.3 cm$^3$) that could best predict AD-type dementia after 2 years based on the
Youden index from a time-dependent ROC curve in R,30 including all subjects with HCV data in
both cohorts (n=339). In 6 subjects, MRI was performed at follow-up. As these subjects had a
normal HCV score, they were considered to have a normal HCV score at baseline as well.

APOE genotype
APOE genotype was determined by polymerase chain reaction of genomic DNA extracted from EDTA
anticoagulated blood in 138 subjects. Subjects were classified as APOE-ε4 carriers or non-carriers.

Statistical analyses
Statistical analyses were done with SPSS version 16.0 (Chicago, IL, USA) and statistical software
package R version 2.10.1.33 Significance was set at p<0.05 and 95% confidence intervals (CI) were
calculated. Differences between groups were analyzed using a t-test for continuous variables
and chi-square test for categorical variables.
Univariable and multivariable logistic regression analyses were performed for dichotomized
CSF and MRI markers to assess whether a combination of markers was better predictive for
AD-type dementia. Sensitivity, specificity, positive predictive value, negative predictive value,
Youden index (sensitivity+specificity-1) and odds ratio (OR) were calculated. Multivariable
analyses with correction for center yielded similar findings. Therefore, we did not correct for
center in the final analyses.
Change on the MMSE and cognitive composite score were assessed by use of slope analyses
with mixed models. The analyses included the baseline score and available follow-up scores up
to 5 years after baseline. We used an unstructured covariance structure with center as a random
effect as this model provided the best -2 log likelihood compared to models with simpler
covariance structures. We examined whether the slopes of cognition were different for subjects with and without abnormal CSF or MRI biomarker scores and whether the combination of both biomarkers increased predictive accuracy for cognitive decline. Test sequence analyses were performed to investigate the added predictive value of a second biomarker for AD-type dementia or cognitive decline in specific subgroups. We identified the best test sequence of CSF Aβ1-42/tau ratio and HCV assessment and visualized it using probability-modifying plots. The relation between biomarkers, the sample size, and number of subjects required for screening, was tested for a fictive placebo-controlled AD drug trial targeting amyloid pathology in subjects with MCI. The number of subjects was calculated in such a way that the study could detect a relative decrease of 25% in conversion rate to AD-type dementia over a 2-year period in subjects with AD pathology, which was defined as an abnormal CSF Aβ1-42/tau ratio, with a power of 90%, a 2-sided alpha of 5%, a drop-out rate of 30% and equally sized treatment groups. The prevalence of abnormal biomarkers and conversion rate were based on the observed data in our study.

RESULTS
Sample characteristics
153 subjects had at least one follow-up assessment (average 3.7 years (SD=1.4)). 48 (31%) subjects converted to AD-type dementia after 2 years, 6 (4%) subjects converted to other types of dementia and were included in the non-AD group. Subjects with at least one follow-up had similar baseline characteristics as those without follow-up (supplemental table S1).

Predictors of AD-type dementia
Univariable predictors
Subjects with AD-type dementia at 2-year follow-up had lower MMSE scores (p=0.014), lower cognitive composite score (p=0.009), lower CSF levels of Aβ1-42 (p<0.001), higher CSF levels of t-tau (p<0.001), a lower CSF Aβ1-42/tau ratio (p<0.001), smaller HCV (p<0.001), a higher frequency of APOE-ε4 alleles (p=0.011), and tended to have more often aMCI (75 vs. 65%, p>0.10), compared to subjects without AD-type dementia (table 1). Scores on verbal fluency (p=0.052) and delayed recall (p=0.069) tended to be lower for subjects who progressed to AD-type dementia compared to those who did not progress. The overall predictive accuracy (OR) was more than twice as high for the CSF Aβ1-42/tau ratio (OR=9.2, 95% CI 3.9-22, p<0.001) than for HCV (OR=3.7, 95% CI 1.8-7.6, p<0.001, table 2). Of the subjects without dementia at follow-up, 14 had a follow-up shorter than 1.5 years. When we repeated analyses after exclusion of these subjects, results were similar.

Multivariable predictors
When both the CSF Aβ1-42/tau ratio and HCV were entered in the same model, the markers predicted AD-type dementia independently from each other (OR CSF ratio=7.9, 95% CI 3.3-19, p<0.001; OR HCV=2.8, 95% CI 1.3-6.2, p=0.01). The combination increased the predictive accuracy relative to the model with only the CSF Aβ1-42/tau ratio (χ²=6.6, df=1, p=0.01). Next, we
Table 1. Baseline characteristics of subjects with MCI by outcome at 2-year follow-up

<table>
<thead>
<tr>
<th></th>
<th>MCI-non AD</th>
<th>MCI-AD</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>105</td>
<td>48</td>
</tr>
<tr>
<td>Age</td>
<td>68.8 (7.3)</td>
<td>70.4 (7.2)</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>43 (41.0)</td>
<td>24 (50.0)</td>
</tr>
<tr>
<td>aMCI, n (%)</td>
<td>67 (65.0)</td>
<td>35 (74.5)</td>
</tr>
<tr>
<td>Years of education</td>
<td>10.7 (3.2)</td>
<td>11.2 (3.0)</td>
</tr>
<tr>
<td>MMSE</td>
<td>26.9 (2.6)</td>
<td>25.8 (2.8)</td>
</tr>
<tr>
<td>Delayed recall, z-score</td>
<td>-1.34 (1.1)</td>
<td>-1.76 (1.0)</td>
</tr>
<tr>
<td>Learning, z-score</td>
<td>-1.15 (1.1)</td>
<td>-1.44 (1.1)</td>
</tr>
<tr>
<td>Verbal fluency, z-score</td>
<td>-0.66 (1.1)</td>
<td>-1.04 (0.9)</td>
</tr>
<tr>
<td>TMT A, z-score</td>
<td>-0.44 (1.6)</td>
<td>-0.87 (1.8)</td>
</tr>
<tr>
<td>TMT B, z-score</td>
<td>-0.60 (1.3)</td>
<td>-0.85 (1.3)</td>
</tr>
<tr>
<td>Visuo-construction, z-score</td>
<td>0.26 (1.2)</td>
<td>0.18 (1.2)</td>
</tr>
<tr>
<td>Cognitive composite score, z-score</td>
<td>-0.63 (0.7)</td>
<td>-0.95 (0.7)</td>
</tr>
<tr>
<td>APOE ε4 carrier, n (%)</td>
<td>42 (44.2)</td>
<td>29 (67.4)</td>
</tr>
<tr>
<td>Hippocampal volume, cm³</td>
<td>5.7 (0.7)</td>
<td>5.3 (0.6)</td>
</tr>
<tr>
<td>Aβ1-42, pg/ml</td>
<td>682 (309)</td>
<td>470 (145)</td>
</tr>
<tr>
<td>T-tau, pg/ml</td>
<td>409 (219)</td>
<td>686 (377)</td>
</tr>
<tr>
<td>Ratio Aβ1-42/t-tau</td>
<td>2.3 (1.7)</td>
<td>0.8 (0.5)</td>
</tr>
</tbody>
</table>

Data are mean (SD), unless specified otherwise. *p<0.05; †p<0.01; ‡p<0.001 compared to MCI-non AD. Group size is different for cognitive tests, APOE, hippocampal volume, and CSF measures. MCI-nonAD=subjects with MCI who did not convert to AD-type dementia at follow-up; MCI-AD=subjects with MCI who converted to AD-type dementia at follow-up; aMCI=amnestic mild cognitive impairment; MMSE=Mini-Mental State Examination; TMT=trail making test; APOE=apolipoprotein E genotype; Aβ1-42=beta amyloid 1-42; t-tau=total tau.

tested the added diagnostic value of CSF Aβ1-42/t-tau ratio and HCV compared to models with other markers that could predict AD-type dementia in the univariable analyses and age. When age, MMSE score, and APOE genotype were entered in the first step, CSF Aβ1-42/t-tau ratio and HCV increased the predictive accuracy in the second step ($\chi^2=24.1$, df=2, p<0.001). Both markers were significant predictors (OR CSF ratio=7.9, 95% CI 2.8-22, p<0.001; OR HCV=2.8, 95% CI 1.1-6.9, p=0.03), whereas age, MMSE score, and APOE genotype no longer predicted outcome (p>0.42). When we tested a multivariable model in which the cognitive composite score was entered in the first step (p=0.012) and CSF Aβ1-42/t-tau ratio and HCV in the second step, both biomarkers predicted AD-type dementia (p<0.05) but the cognitive composite score not (p=0.13).

**Test sequence**

In further analyses, we only included the CSF and HCV measures because other covariates did not contribute to the overall predictive accuracy.

The added predictive value of the CSF Aβ1-42/t-tau ratio and HCV to the other marker is presented in supplemental table S2 and figure 1. If the CSF Aβ1-42/t-tau ratio was analyzed...
first, assessment of HCV increased predictive accuracy for AD-type dementia in subjects with
a normal CSF Aβ1-42/tau ratio (OR=7.1, p=0.014), but not in subjects with an abnormal CSF
Aβ1-42/tau ratio (OR=2.0, p=0.13). If HCV was analyzed first, assessment of CSF Aβ1-42/tau ratio
increased predictive accuracy both in subjects with normal HCV (OR=14.4, p<0.001) and with
abnormal HCV (OR=4.1, p=0.025).

Figure 1 shows probability plots of AD-type dementia for different sequences of CSF Aβ1-42/tau
ratio and HCV assessment, thereby visualizing the change in AD probability after adding the
results for each biomarker.

<table>
<thead>
<tr>
<th>Markers</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>Youden Index</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF Aβ1-42/tau ratio</td>
<td>0.83</td>
<td>0.65</td>
<td>0.52</td>
<td>0.89</td>
<td>0.48</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>(0.73-0.94)</td>
<td>(0.56-0.74)</td>
<td>(0.41-0.63)</td>
<td>(0.83-0.96)</td>
<td>(0.34-0.62)</td>
<td>(3.9-22)</td>
</tr>
<tr>
<td>HCV</td>
<td>0.56</td>
<td>0.74</td>
<td>0.50</td>
<td>0.79</td>
<td>0.31</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>(0.42-0.70)</td>
<td>(0.66-0.83)</td>
<td>(0.37-0.63)</td>
<td>(0.71-0.87)</td>
<td>(0.14-0.47)</td>
<td>(1.8-7.6)</td>
</tr>
<tr>
<td>CSF Aβ1-42/tau ratio</td>
<td>0.94</td>
<td>0.52</td>
<td>0.47</td>
<td>0.95</td>
<td>0.46</td>
<td>17</td>
</tr>
<tr>
<td>or HCV abnormal</td>
<td>(0.87-1.00)</td>
<td>(0.43-0.62)</td>
<td>(0.37-0.57)</td>
<td>(0.89-1.00)</td>
<td>(0.34-0.58)</td>
<td>(4.8-56)</td>
</tr>
<tr>
<td>CSF Aβ1-42/tau ratio</td>
<td>0.46</td>
<td>0.87</td>
<td>0.61</td>
<td>0.78</td>
<td>0.33</td>
<td>5.5</td>
</tr>
<tr>
<td>and HCV abnormal</td>
<td>(0.32-0.60)</td>
<td>(0.80-0.93)</td>
<td>(0.45-0.77)</td>
<td>(0.70-0.85)</td>
<td>(0.17-0.48)</td>
<td>(2.5-12)</td>
</tr>
</tbody>
</table>

*All p-values <0.001. Results are presented with 95 % confidence intervals. Youden index=(sensitivity+specificity-1). CSF=cerebrospinal fluid; MRI=magnetic resonance imaging; MCI=mild cognitive impairment; HCV=hippocampal volume; PPV=positive predictive value; NPV=negative predictive value.

Figure 1. Probability of AD-type dementia in 2 years related to the sequence of CSF and HCV assessment. The graph shows the change in AD probability according to biomarker scores. The starting point represents the AD conversion rate in the total MCI population. Dotted lines represent negative (normal) test results and full lines represent positive (abnormal) test results. (A) First CSF assessment then HCV assessment; (B) First HCV assessment then CSF assessment. 36 subjects had both and abnormal CSF ratio and HCV (with 22 converters to AD-type dementia), 41 subjects had an abnormal CSF ratio and a normal HCV (18 converters), 18 subjects had a normal CSF ratio and an abnormal HCV (5 converters), 58 subjects had both a normal CSF ratio and HCV (3 converters). AD=Alzheimer’s disease; CSF=cerebrospinal fluid; HCV=hippocampal volume. *p<0.05; †p<0.001: difference in AD probability in 2 years between subjects with normal and abnormal biomarkers after second biomarker assessment.
Predictors of cognitive decline

Although conversion to AD-type dementia after 2-year follow-up is a clinically relevant outcome measure, the potential of CSF and MRI biomarkers to identify prodromal AD may be underestimated as subjects may convert at longer follow-up intervals. Therefore, we additionally performed slope analyses for annual cognitive decline over the 5-year follow-up period.

**Univariable predictors**
The CSF Aβ1-42/tau ratio at baseline predicted cognitive decline on the MMSE and cognitive composite score (all p<0.001), while HCV only predicted decline on the MMSE (p=0.008; table 3). The CSF Aβ1-42/tau ratio was a better predictor for decline on the MMSE compared to HCV (difference -2 log likelihood (LL)=15.4; CSF F=27.6, p<0.001; MRI F=7.5, p=0.008).

**Multivariable predictors**
Multivariable analyses were performed with decline on the MMSE as outcome, since only for this outcome measure both biomarkers predicted decline in univariable analyses. When the CSF Aβ1-42/tau ratio was entered first, addition of HCV significantly increased the overall predictive accuracy (decrease in -2LL=6.5, F=4.72, p=0.011).

**Table 3.** Annual cognitive decline over 5 years of follow-up according to CSF and MRI biomarker results at baseline

<table>
<thead>
<tr>
<th></th>
<th>MMSE</th>
<th>Cognitive composite score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline score</td>
<td>Slope</td>
</tr>
<tr>
<td></td>
<td>Baseline score</td>
<td>Slope</td>
</tr>
<tr>
<td>CSF Aβ1-42/tau ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>27.5 (0.82)¹</td>
<td>-0.28” (0.13)</td>
</tr>
<tr>
<td>Abnormal</td>
<td>26.6 (0.83)</td>
<td>-1.24” (0.12)</td>
</tr>
<tr>
<td></td>
<td>-0.70 (0.10)¹</td>
<td>0.001¹ (0.02)</td>
</tr>
<tr>
<td></td>
<td>-0.93 (0.11)</td>
<td>-0.25¹ (0.04)</td>
</tr>
<tr>
<td>HCV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>27.3 (0.73)²</td>
<td>-0.61&quot; (0.13)</td>
</tr>
<tr>
<td>Abnormal</td>
<td>26.1 (0.80)</td>
<td>-1.17&quot; (0.16)</td>
</tr>
<tr>
<td></td>
<td>-0.76 (0.09)</td>
<td>-0.05 (0.02)</td>
</tr>
<tr>
<td></td>
<td>-0.89 (0.11)</td>
<td>-0.14² (0.04)</td>
</tr>
</tbody>
</table>

Data are mean (SE). Slope refers to annual change on the test. A negative slope indicates cognitive decline. CSF=Cerebrospinal fluid, HCV=Hippocampal volume, MMSE=Mini Mental State Examination. ¹p<0.05, ²p<0.001 for slope different from zero (this means a statistically significant change over time in test score); ³p<0.001, ⁴p<0.01, ⁵p<0.05 for baseline score or slope compared to baseline score or slope in abnormal biomarker group.

**Test sequence**
The added predictive value for decline on the MMSE of the CSF Aβ1-42/tau ratio and HCV to the other marker is presented in supplemental table S3 and figure 2. If the CSF Aβ1-42/tau ratio was analyzed first, assessment of HCV increased predictive accuracy for cognitive decline in subjects with a normal CSF Aβ1-42/tau ratio (p=0.001), but not in subjects with an abnormal CSF Aβ1-42/tau ratio (p=0.3). If HCV was analyzed first, assessment of CSF Aβ1-42/tau ratio increased predictive accuracy both in subjects with normal HCV (p<0.001) and with abnormal HCV (p=0.038).

**Biomarkers as inclusion criterion for prodromal AD trial**
We used the different combinations of the CSF Aβ1-42/tau ratio and HCV shown in table 2 to select subjects with MCI for a fictive amyloid targeting AD trial with conversion to AD-type dementia.
dementia after 2 years as outcome. The sample sizes, numbers required for screening, and screening failure rates varied considerably between the different combinations (table 4). The sample size was smallest for the strategy that required both an abnormal CSF $\alpha\beta_1$-42/tau ratio and abnormal HCV at baseline (642 vs. 880-1381). The number required for screening was lowest if only the CSF $\alpha\beta_1$-42/tau ratio was required to be abnormal (1761 vs. 2228-3931). The screening failure rate was lowest if either the CSF $\alpha\beta_1$-42/tau ratio or HCV was required to be abnormal (38% vs. 50-76%). The strategy that required only abnormal CSF markers had the best trade-off between sample size (n=880) and number required for screening (n=1761), as numbers were best (screening) or second best (sample size).

Table 4. Number of subjects required for AD trial design based on biomarker inclusion criteria

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>Number required for screening</th>
<th>Sample size for trial</th>
<th>Screening failure rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF $\alpha\beta_1$-42/tau ratio abnormal</td>
<td>1761</td>
<td>880</td>
<td>0.50</td>
</tr>
<tr>
<td>HCV abnormal</td>
<td>3931</td>
<td>1376</td>
<td>0.65</td>
</tr>
<tr>
<td>CSF $\alpha\beta_1$-42/tau ratio or HCV abnormal</td>
<td>2228</td>
<td>1381</td>
<td>0.38</td>
</tr>
<tr>
<td>CSF $\alpha\beta_1$-42/tau ratio and HCV abnormal</td>
<td>2675</td>
<td>642</td>
<td>0.76</td>
</tr>
</tbody>
</table>

The assumptions with regard to trial design are described in the methods section. Screening failure rate = 1-sample size divided by number required for screening, i.e. the proportion of subjects that were screened but not included in the AD trial, as they did not have abnormal biomarkers. Note that this failure rate does not take into account failure rates of CSF collection. AD = Alzheimer’s disease; CSF = cerebrospinal fluid; HCV = hippocampal volume.
DISCUSSION

We showed that the CSF Aβ1-42/tau ratio was the best predictor for AD-type dementia at follow-up and for cognitive decline in subjects with MCI. HCV could increase the predictive accuracy of the CSF Aβ1-42/tau ratio. This increase in predictive accuracy for AD-type dementia and cognitive decline was because HCV predicted AD-type dementia and cognitive decline in subjects with a normal CSF Aβ1-42/tau ratio.

Test sequence analyses suggested that CSF assessment should be performed first for prediction of AD-type dementia at follow-up. HCV assessment may be considered in subjects who have a normal CSF Aβ1-42/tau ratio. A strategy in which HCV assessment is performed first and CSF assessment next seems less efficient as CSF assessment would further increase predictive accuracy regardless of the outcome of the HCV analysis. However, with respect to feasibility, MRI scanning may be more acceptable than CSF collection in daily practice. MRI is often already routinely done to exclude other diseases. Still, volumetric assessment of the hippocampus is not widely available.

Furthermore, the current study showed that for a fictive AD trial targeting amyloid pathology in subjects with MCI, subjects could be best selected based on CSF biomarkers because this provided the best balance between sample size and number of subjects required for screening. Still, the sample size was smaller if both CSF Aβ1-42/tau ratio and HCV were required to be abnormal. However, the screening failure rate was very high for this combination, which may limit its clinical utility. The screening failure rate was lowest if subjects were required to have either an abnormal CSF Aβ1-42/tau ratio or an abnormal HCV. A disadvantage of this approach was the much higher sample size compared to that of the selection based on an abnormal CSF Aβ1-42/tau ratio alone. Selection based on an abnormal HCV alone was least efficient because both the number required for screening and sample size were among the highest of the four approaches. This may be partly because the fictive trial was designed for targeting amyloid pathology such that only subjects with an abnormal CSF Aβ1-42/tau ratio were supposed to respond to therapy.

This is the first large-scale multicenter study conducted in a memory clinic setting of subjects classified according to a broad definition of MCI, including aMCI (68%) and naMCI that compared CSF and MRI markers for prediction of cognitive decline. Our study corroborates the findings of studies of subjects with aMCI,14-16 that the combination of the CSF ratio and measure of medial temporal lobe atrophy predicted AD-type dementia better than each biomarker alone. This also confirms the findings of smaller single-center studies performed in subjects with MCI.10-12

It validates the use of biomarkers for prediction of AD-type dementia in subjects who meet the criteria of a broad definition of MCI, including aMCI and naMCI. This supports the new NIA-AA criteria for MCI due to AD, which uses a broad definition of MCI as well.18 The conversion rate of 31% to AD-type dementia after 2 years in the present study is comparable to conversion rates in previous studies with a similar follow-up period.11,12,15

The overall predictive accuracy (OR) for AD-type dementia of the CSF Aβ1-42/tau ratio was twice as high as that of the HCV. This was mainly due to the higher sensitivity of the CSF Aβ1-42/tau ratio compared to that of the HCV (table 2). This is in line with the hypothetical dynamic model of AD biomarkers,16 stating that CSF biomarkers become abnormal at an earlier stage than MRI.
biomarkers. We also found that the specificity of the HCV was higher than that of the CSF Aβ1-42/tau ratio. This may be considered unexpected as medial temporal lobe atrophy is not specific for AD and also observed in other neurodegenerative and vascular disorders. However, the higher specificity of the HCV could also be explained by the hypothetical dynamic model of AD biomarkers. As CSF markers are expected to become abnormal before the HCV,31 there will be more MCI subjects with prodromal AD with an abnormal CSF Aβ1-42/tau ratio who have not yet converted to AD-type dementia within 2 years than there will be MCI subjects with prodromal AD with an abnormal HCV who have not yet converted to AD-type dementia within 2 years. Test sequence analyses revealed the presence of a small group of subjects (n=18, 12% of the total sample) with a normal CSF Aβ1-42/tau ratio but abnormal HCV and a conversion rate of 28% to AD-type dementia after 2 years. This is an interesting group because it is not in line with the hypothetical dynamic model of AD biomarkers,31 as abnormal CSF biomarkers did not precede medial temporal lobe atrophy. Further research is needed to investigate whether these subjects have an atypical presentation of AD or whether they are misclassified cases either because of the applied cut-offs or because AD is not the cause of the dementia.

MMSE score, a cognitive composite score, and APOE genotype predicted AD-type dementia in univariable analyses, however, in the combined model with CSF and MRI markers, these variables were no longer significant predictors. This indicates that CSF and MRI can be used independent of cognitive scores and APOE genotype for short-term prediction of AD-type dementia. Nevertheless, our findings are inconsistent with some previous studies of subjects with aMCI. One study showed that MRI and APOE genotype did not significantly add to the predictive accuracy of neuropsychological tests.36 Another study found that the optimal biomarker combination and neuropsychological tests were equally predictive for AD-type dementia.14 The more severely impaired cohort that was used in the former study and the inclusion of only subjects with aMCI in both studies could explain the inconsistent findings. Although a 2-year follow-up is clinically relevant, the potential of CSF and MRI biomarkers to identify prodromal AD may be underestimated.29 Therefore, we also performed slope analyses for annual cognitive decline over 5 years. These analyses yielded very similar findings as our analysis on conversion to AD-type dementia after 2 years. The CSF Aβ1-42/tau ratio predicted annual cognitive decline better than HCV. In the sequence analysis with decline on MMSE as outcome, HCV increased predictive accuracy only in subjects with a normal CSF Aβ1-42/tau ratio but not in subjects with an abnormal ratio, while the CSF Aβ1-42/tau ratio improved predictive accuracy regardless of HCV. The lower predictive accuracy of HCV can again in part be explained by a lower sensitivity of this marker for prodromal AD. This is supported by the observation that decline in subjects with a normal HCV is higher than in subjects with a normal CSF ratio in table 3. This higher decline is probably driven by subjects with an abnormal CSF ratio among subjects with a normal HCV.

This study had several limitations. As the findings were based on memory clinic populations, they may not be generalized to other settings, including the general population. Furthermore, the cut-offs of the CSF Aβ1-42/tau ratio and HCV were determined within a study population that also included the subjects from the present analyses. Although this may have led to an overestimation of the predictive accuracy of the biomarkers, it is unlikely that it influences our findings with
respect to the differences in predictive accuracy between the CSF Aβ1-42/tau ratio and HCV, as we used the same method to define the cut-point for each biomarker. We defined cut-offs regardless of age, although biomarkers may change with age. However, when we repeated analyses using age-adjusted cut-offs (<70 vs. >70 years) analyses yielded similar findings. In addition, there was variability in scanners used, which may have influenced the volumetric measurements and predictive accuracy of the HCV. However, the LEAP approach has been shown to be robust for scanner variability. The follow-up period to AD-type dementia was relatively short but for clinical trials short-term prognosis may be important. However, we used a longer follow-up period to investigate the predictive accuracy of CSF and MRI biomarkers for annual cognitive decline over 5 years. Moreover, the diagnosis of AD-type dementia at follow-up was not neuropathologically validated. This may have led to the misclassification of some cases.

A major strength of this study was the large sample size. The overall multicenter design favors generalizability to other memory clinic settings, although a considerable amount of subjects was recruited from one single center. Furthermore, this is the first study of predictive accuracy of the combination of CSF and MRI markers for AD-type dementia and cognitive decline that investigated the test sequence of CSF and HCV assessments in an unselected population with subjects classified according to a broad definition of MCI.

In sum, test sequence analyses may lead to cost reduction and to a decrease in patients’ burden. Future research is needed to evaluate the ratio of costs and effectiveness for patients regarding biomarker assessment, as this may eventually require adjustments to patient policy. Also the combination of CSF and MRI biomarkers with PIB-PET, and FDG-PET should be investigated, since this may further aid to define an algorithm of markers for the accurate prediction of AD-type dementia in subjects with MCI. While hippocampal atrophy and CSF markers are given equal diagnostic accuracy in recent criteria for prodromal AD or MCI—due to AD, our findings suggest that this may not be the case and a CSF Aβ1-42/tau ratio may be preferred over HCV assessment because of its higher sensitivity.

REFERENCES


30. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival


APPENDIX S1. NEUROPSYCHOLOGICAL TESTS USED TO DIAGNOSE MCI

Neuropsychological tests were used to assess cognitive performance in the domains of memory, language, executive function and attention, and visuoconstruction. Because the DESCRIPA study followed the routine practice at each centre, the tests used to assess each domain varied among centers.

At each center, we selected a primary test for each cognitive domain that was the same as or similar to the tests that were used at the other centers (Visser et al 2008). The primary tests to assess memory were the learning and delayed recall measure of the Rey auditory verbal learning test (4 centers) and the word list of the consortium to establish a registry for AD (CERAD) test battery (2 centers). The primary test to assess language was the 1-min verbal fluency for animals (all centers). The primary test to assess executive function and attention was the trail making test (TMT) parts A and B (all centers). The primary tests to assess visuoconstruction were the copy subtest of the Rey–Osterrieth complex figure (4 centers) or the copy of the CERAD figures (2 centers). If patients could not be classified on the basis of these primary tests because of missing data, we used alternative tests, qualitative rating by a neuropsychologist, or a subscore on the MMSE.
APPENDIX S2. SCAN PARAMETERS AND MRI PROTOCOLS USED AT EACH CENTER.

CENTER AMSTERDAM VUMC
Siemens Magnetom Impact Expert 1.0 T, 3D scan, 168 slices, FOV 250 mm, matrix 256 × 256; slice thickness 1.5 mm, TE: 7 ms, TR: 15 ms, TI 300 ms, flip angle 15°.

CENTER HUDDINGE
Siemens Avanto 1.5 T, 21 slices, FOV 220 mm, FOV phase 87.1, distance factor 30, phase R>L, slice thickness 5.0 mm, TE: 96 ms, TR: 4000 ms, flip angle 150°, number of averages 1.

Siemens Symphony 1.5 T, 21 slices, FOV 220 mm, FOV phase 75.0, distance factor 30, phase R>L, slice thickness 5.0 mm, TE: 99 ms, TR: 4100 ms, flip angle 150°, number of averages 2.

CENTER KUOPIO
Siemens Vision 1.5 T, T1 3D-scan, MPRAGE OBL; COR>TRA, FOV 250, mat 256x256, 128 slices, TR 9.7ms, TE 4 ms, Slice th 2.0mm, no slice gap, flip angle 12°

CENTER MALMO
Siemens Sonata 1.5 T, MPRAGE + lmpr-cor, 144 slices, FOV 250 mm, phase R>L, TR 1970, TE 3.93, distance factor 50, slice thickness 1.5 mm, flip angle 15°.

CENTER MUNICH
Siemens Magnetom Vision; 1.5 T; MPRAGE; Slice thickness 1.05 mm, TR 11.4, TE 4.4, TI 300; FOV 256*256; Flip angle 8graden; number of averages 1.

CENTER THESSALONIKI
Siemens Expert Plus unit 1.0 T, 3D-MPR: 15 (TR) ,7 TE ,8 FLIP ANG., 1,49Ef thick, 168 Partitions, 250 FOV,256x192 Matrix, 1 Aquis. , ACQ TIME 10,21min
Table S1. Baseline characteristics for subjects with and without follow-up

<table>
<thead>
<tr>
<th></th>
<th>MCI all</th>
<th>MCI FU</th>
<th>MCI no FU</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>165</td>
<td>153</td>
<td>12</td>
</tr>
<tr>
<td>Age</td>
<td>69.3 (7.3)</td>
<td>69.3 (7.3)</td>
<td>70.3 (6.8)</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>73 (44.2)</td>
<td>67 (43.8)</td>
<td>6 (50.0)</td>
</tr>
<tr>
<td>aMCI, n (%)</td>
<td>112 (67.9)</td>
<td>102 (66.7)</td>
<td>10 (83.3)</td>
</tr>
<tr>
<td>Years of education</td>
<td>10.8 (3.1)</td>
<td>10.9 (3.2)</td>
<td>10.2 (2.8)</td>
</tr>
<tr>
<td>MMSE</td>
<td>26.6 (2.6)</td>
<td>26.6 (2.7)</td>
<td>26.6 (2.1)</td>
</tr>
<tr>
<td>APOE ε4 carrier, n (%)</td>
<td>75 (50.3)</td>
<td>71 (51.4)</td>
<td>4 (36.4)</td>
</tr>
<tr>
<td>Hippocampal volume, cm³</td>
<td>5.6 (0.7)</td>
<td>5.6 (0.7)</td>
<td>5.8 (0.1)</td>
</tr>
<tr>
<td>Aβ1-42, pg/ml</td>
<td>609 (282)</td>
<td>616 (285)</td>
<td>527 (236)</td>
</tr>
<tr>
<td>T-tau, pg/ml</td>
<td>487 (301)</td>
<td>496 (305)</td>
<td>377 (219)</td>
</tr>
<tr>
<td>Ratio Aβ1-42/t-tau</td>
<td>1.9 (1.6)</td>
<td>1.8 (1.6)</td>
<td>2.0 (1.6)</td>
</tr>
</tbody>
</table>

Data are mean (SD), unless specified otherwise. *Group size is different for APOE, hippocampal volume and CSF measures. No significant statistical differences were found between subjects with and without follow-up. FU=follow-up; aMCI=amnestic mild cognitive impairment; MMSE=Mini-Mental State Examination; APOE=apolipoprotein E genotype; Aβ1-42=beta amyloid 1-42; t-tau=total tau.

Table S2. Added predictive value of a second biomarker for AD-type dementia after 2-year follow-up

<table>
<thead>
<tr>
<th>Biomarker status after first test</th>
<th>Second biomarker tested</th>
<th>N</th>
<th>LR+</th>
<th>LR-</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF Aβ1-42/tau ratio normal</td>
<td>HCV</td>
<td>76</td>
<td>3.3</td>
<td>0.46</td>
<td>7.1*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.6-6.8)</td>
<td>(0.19-1.14)</td>
<td></td>
</tr>
<tr>
<td>CSF Aβ1-42/tau ratio abnormal</td>
<td>HCV</td>
<td>77</td>
<td>1.5</td>
<td>0.72</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.88-2.4)</td>
<td>(0.47-1.11)</td>
<td></td>
</tr>
<tr>
<td>HCV normal</td>
<td>CSF Aβ1-42/tau ratio</td>
<td>99</td>
<td>2.9</td>
<td>0.20</td>
<td>14'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2.0-4.3)</td>
<td>(0.07-0.58)</td>
<td></td>
</tr>
<tr>
<td>HCV abnormal</td>
<td>CSF Aβ1-42/tau ratio</td>
<td>54</td>
<td>1.6</td>
<td>0.38</td>
<td>4.1'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.1-2.4)</td>
<td>(0.16-0.93)</td>
<td></td>
</tr>
</tbody>
</table>

Results are presented with 95% confidence intervals. AD=Alzheimer’s disease; MCI=mild cognitive impairment; CSF=cerebrospinal fluid; HCV=hippocampal volume; LR+=positive likelihood ratio; LR-=negative likelihood ratio. *P<0.05; p<0.001.
Table S3. Added predictive value of a second biomarker for cognitive decline on MMSE

<table>
<thead>
<tr>
<th>Biomarker status after first test</th>
<th>Second biomarker tested</th>
<th>Biomarker status</th>
<th>Slope</th>
<th>P-value difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF Aβ1-42/tau ratio normal</td>
<td>HCV</td>
<td>Normal</td>
<td>-0.27 (0.1)</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
<td>-0.71 (0.1)</td>
<td></td>
</tr>
<tr>
<td>CSF Aβ1-42/tau ratio abnormal</td>
<td>HCV</td>
<td>Normal</td>
<td>-1.06 (0.2)</td>
<td>0.333</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
<td>-1.33 (0.2)</td>
<td></td>
</tr>
<tr>
<td>HCV normal</td>
<td>CSF Aβ1-42/tau ratio</td>
<td>Normal</td>
<td>-0.19 (0.1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
<td>-1.08 (0.1)</td>
<td></td>
</tr>
<tr>
<td>HCV abnormal</td>
<td>CSF Aβ1-42/tau ratio</td>
<td>Normal</td>
<td>-0.57 (0.3)</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
<td>-1.35 (0.2)</td>
<td></td>
</tr>
</tbody>
</table>

Data are slopes (SE) for annual decline on MMSE. A negative slope indicates cognitive decline. As the sample size of the subgroups was different, different slopes may be presented for the same final groups. CSF=Cerebrospinal fluid, HCV=Hippocampal volume, MMSE=Mini Mental State Examination. ‘P<0.001.'
Figure S1. Image of hippocampal LEAP segmentation. This figure shows an example of an image of hippocampal volume segmentation for horizontal, coronal, and sagittal sections using the atlas-based LEAP method.