Biomarker based prognosis for patients with mild cognitive impairment: the ABIDE project

Ingrid S. van Maurik, MSc[1,2] Stephanie J. Vos, PhD[3], Isabelle Bos, PhD[1,3], Femke H. Bouwman, MD[1], Prof Charlotte E. Teunissen, PhD[4], Prof Philip Scheltens, MD1, Prof Frederik Barkhof, MD[5,6] , Prof Lutz Frolich, MD [7] , Prof Johannes Kornhuber, MD[8], Prof Jens Wiltfang, MD[9], Prof Wolfgang Maier, MD[12], Oliver Peters, MD[13,14], Prof Eckart Rüther, MD[15], Prof Flavio Nobili, MD[16,17], Prof Giovanni B. Frisoni, MD[18], Prof Luiza Spiru, MD[19,20], Yvonne Freund-Levi, MD[21,22,23], Asa K. Wallin, PhD[24], Prof Harald Hampel, MD[25,26,27,28], Prof Hilkka Soininen, MD[29], Prof Magda Tsolaki, PhD[30], Prof Frans Verhey, MD[3], Prof Iwona Kłoszewska, MD[31], Prof Patrizia Mecocci, MD[32], Prof Bruno Vellas, MD [33], Prof Simon Lovestone, PhD[34], Samantha Galluzzi, MD[35], Sanna-Kaisa Herukka, MD[29], Isabel Santana, PhD[36], Ines Baldeiras, PhD[36], Alexandre de Mendonça, MD[37], Prof Dina Silva, PhD[38,39], Gael Chetelat, PhD[39], Stephanie Egret, MSc[39], Sebastian Palmqvist, MD[24,40], Prof Oskar Hansson, MD[24,41], Pieter Jelle Visser, MD[1,3], Prof Johannes Berkhof, PhD[2], and Prof Wiesje M. van der Flier, PhD[1,2]

Affiliations

1. Alzheimer Center Amsterdam, Department of Neurology, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands.
2. Department of Epidemiology and Biostatistics, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands.
3. Department of Psychiatry and Neuropsychology, Maastricht University, School for Mental Health and Neuroscience, Alzheimer Centre Limburg, Maastricht, The Netherlands
4. Neurochemistry Laboratory, Department of Clinical Chemistry, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands.
5. Department of Radiology and Nuclear Medicine, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands.
6. Institutes of Neurology and Healthcare Engineering, University College London, London, England.
7. Department of Geriatric Psychiatry, Zentralinstitut für Seelische Gesundheit, Medical Faculty Mannheim University of Heidelberg, Germany.
8. Department of Psychiatry and Psychotherapy, Friedrich-Alexander University of Erlangen-Nürnberg, Erlangen, Germany
9. Department of Psychiatry and Psychotherapy, University Medical Center (UMC), Georg-August-University, Göttingen, Germany.
10. German Center for Neurodegenerative Diseases (DZNE), Göttingen, Germany
11. iBiMED, Medical Sciences Department, University of Aveiro, Aveiro, Portugal
12. Department of Neurodegenerative Diseases and Gerotopsychiatry, University of Bonn, German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany.
13. Department of Psychiatry, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany
14. German Center for Neurodegenerative Diseases (DZNE), Berlin, Germany
15. Department of Psychiatry and Psychotherapy, University of Göttingen, Göttingen, Germany.
16. On behalf of the EADC-PET consortium; Clinical Neurology, Department of Neurosciences (DINOGMI), University of Genoa and IRCCS AOU San Martino-IST, Genoa, Italy.
17. IRCCS Ospedale Policlinico San Martino, Genoa, Italy.
18. Memory Clinic, University Hospital and University of Geneva, Geneva, Switzerland.
19. Geriatrics, Gerontology and Old Age Psychiatry Clinical Department, Carol Davila University of Medicine and Parmacy – "Elias " Emergency Clinical Hospital, Bucharest, Romania.
20. Memory Clinic and Longevity Medicine, Ana Aslan International Foundation, Romania.
21. School of medical sciences, Örebro University, Örebro, Sweden
22. Department of Neurobiology, Care Sciences and Society, Division of Clinical Geriatrics, Karolinska Institutet Center for Alzheimer Research, Stockholm, Sweden

23. Deptartment of Old Age Psychiatry, Psychology and Neuroscience, King's College London, London, UK
24. Clinical Memory Research Unit, Department of Clinical Sciences, Malmö, Lund University, Lund, Sweden.
25. AXA Research Fund & Sorbonne University Chair, Paris, France
26. Sorbonne University, GRC n° 21, Alzheimer Precision Medicine (APM), AP-HP, Pitié-Salpêtrière Hospital, Boulevard de l'hôpital, F-75013, Paris, France
27. Brain & Spine Institute (ICM), INSERM U 1127, CNRS UMR 7225, Boulevard de l'hôpital, F-75013, Paris, France
28. Institute of Memory and Alzheimer's Disease (IM2A), Department of Neurology, Pitié-Salpêtrière Hospital, AP-HP, Boulevard de l'hôpital, F-75013, Paris, France
29. Institute of Clinical Medicine, Neurology, University of Eastern Finland and Neurocenter, Neurology, Kuopio University Hospital, Kuopio, Finland.
30. 1st Department of Neurology, Aristotle University of Thessaloniki, Memory and Dementia Center, "AHEPA" General Hospital, Thessaloniki, Greece.
31. Medical University of Lodz, Lodz, Poland.
32. Institute of Gerontology and Geriatrics, Department of Medicine, University of Perugia, Perugia, Italy.
33. UMR INSERM 1027, CHU Toulouse, Toulouse, France.
34. University of Oxford, Department of Psychiatry, Oxford, UK.
35. Lab Alzheimer's Neuroimaging and Epidemiology, IRCCS San Giovanni di Dio Fatebenefratelli, Brescia, Italy.
36. Center for Neuroscience and Cell Biology; Faculty of Medicine, University of Coimbra; Department of Neurology, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal.
37. Faculty of Medicine, University of Lisbon, Portugal.
38. Centre for Biomedical Research, Universidade do Algarve, Faro, Portugal.
39. Université Normandie, Inserm, Université de Caen-Normandie, Inserm UMR-S U1237, GIP Cyceron, Caen, France
40. Department of Neurology, Skåne University Hospital, Lund, Sweden.
41. Memory Clinic, Skåne University Hospital, Malmö, Sweden.

*Corresponding author: Ingrid S. van Maurik, Alzheimer Center Amsterdam, Department of Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC, De Boelelaan 1118, 1081 HZ Amsterdam, The Netherlands. Telephone: +31204440816. Fax: +31204448529. Email: i.vanmaurik@vumc.nl

[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Abstract

Background Biomarker-based risk predictions in individual patients with MCI are highly relevant in light of care planning and future disease modifying drugs. We aimed to establish robust, prediction models in a multi-center, multi-cohort design.

Methods We included 2611 patients with mild cognitive impairment (MCI) (age=70±8, 44%F) via the European Medical Information Framework for Alzheimer's Disease (EMIF-AD, n=883), Alzheimer's Disease Neuroimaging Initiative (ADNI, n=829), Amsterdam Dementia Cohort (ADC, n=666) and Swedish BioFINDER study (n=233). Primary end-point was clinical progression to dementia. We evaluated performance of our risk prediction models (demographic model, Hippocampal volume (HCV) model, cerebrospinal fluid (CSF) model) by evaluating them across cohorts, incorporating different measurement methods, determining prognostic performance, updating the models by re-estimating parameters and evaluating calibration. Finally, we constructed a model combining markers for amyloid deposition (A), tauopathy (T) and neurodegeneration (N), in accordance with the AD research framework .

Findings During 3±2 years follow up, 1007 (39%) MCI patients progressed to dementia. Both demographic (Harrell's C=0·62[0·59-0·65]), HCV (Harrell's C=0·67[0·62-0·72]), and CSF (Harrell's C=0·72[0·68-0·74]) models had adequate prognostic performance across cohorts and were well calibrated. The newly constructed ATN model had highest performance (Harrell's c=0·74[0·71-0·76]). As an example, for a female MCI patient (62yrs, MMSE=26) with abnormal biomarkers (Abeta =112, p-tau=35 (Elecsys values) and HCV= 6.2 (Freesurfer values), the probability of progression to dementia was 40%[33-48] in one year, 88%[82-94] in three years and 97%[94-99] in five years.

Interpretation We generated risk models that are robust across cohorts, which adds to their clinical applicability. The models aid clinicians in the interpretation of CSF and HCV results in individual MCI patients and help prepare for a future of precision medicine in AD.

Funding ZonMW-Memorabel (ABIDE; projectnr. 733050201)

Research in context

Evidence before this study

We searched PubMed without language restriction for articles on prognosis in MCI patients at an individual level based on biomarker evidence using the terms "([mild cognitive impairment] AND [prognosis] OR [prognostic factor] OR [prediction model])". Specifically we focused on prognosis in MCI based on biomarker evidence, i.e. atrophy on magnetic resonance imaging (MRI) and amyloid beta (Abeta), total tau (t-tau) and phosphorylated tau (p-tau) in CSF. For this validation study, we took as a starting our previously constructed biomarker-based prognostic models that allow risk prediction on the individual level. However, these proof of principle models were based on a homogeneous, mono-center cohort and did not accommodate different cohorts and biomarker measurement methods. Moreove, prediction beyond three years was not reliable.

Added value of this study

In the current study of 2611 patients with MCI from mono- and multi center cohorts in Europe and America, we validated and updated, according to the TRIPOD guidelines, multivariable (biomarker-based) models for the prediction of dementia. We showed that the models had good generalizability and were well calibrated up to more than five years of follow-up. Moreover, the models accommodate different biomarker measurement methods. In addition, we constructed a model combining measures of amyloid (A), tau (T), and neurodegeneration (N) to provide predictions in accordance with the most recent research guidelines for AD.

Implications of all the available evidence

We have shown the generalizability and robustness of the predictions and the models are made freely available for academic use by the authors upon request. The models allow clinicians to estimate – for any given combination of biomarker results – the probability of progression to dementia within a given period of time. For example, for a female MCI patient (age=62, MMSE=26) without knowledge of biomarker results the progression probabilities to dementia are 11% [10-12] in one year, 39% [36-42] in three years and 57% [52-61] in five years. When both MRI and CSF would be available and

abnormal (Abeta =112, p-tau=35 (CSF measured with Elecsys) and HCV= 6.2 (calculated with Freesurfer software)), the progression probabilities change to 40%[33-48] at one year, 88%[82-94] at three years and 97%[94-99] at five years. On the other end of the spectrum, a MCI patient (male, age=62, MMSE=29) without knowledge of biomarker results, has progression probabilities to dementia of 7%[6-8] in one year, 26%[23-29] in three years and 40% [44-35] in five years. With normal biomarkers (Abeta=1264, p-tau=12 (Elecsys) and HCV=9.8 (Freesurfer)) this patient would have progression probabilities of in 1%[1-2] one year, 5%[4-7] in three years and 8%[6-11] in five years. The outcomes of this study facilitate a more timely and accurate diagnosis which is of high importance at the individual level even in the absence of specific therapies, as this is the starting point to plan and organize care.

Background

Patients with mild cognitive impairment (MCI) have an increased risk of progressing to dementia, most often due to Alzheimer's Disease (AD).[1] Roughly half of MCI patients develop dementia in the course of three years.[2] The other half remains stable or reverts to normal levels of cognition. As a result, patients live with uncertainty for a long time. In a former study on the communication of diagnosis, patients indicated they preferred more information on the future course of their disease.[3] Diagnostic tests, such as magnetic resonance imaging (MRI) measures and/or biomarkers in cerebrospinal fluid (CSF), could help to establish a more accurate prognosis.[4-7]

Practice guidelines for MCI from the American Academy of Neurology (AAN) acknowledge that biomarker research in AD is a rapidly moving field and that biomarker evidence in MCI may be particularly important for prognosis.[8] At the same time, these guidelines state that biomarkers are not yet ready for clinical implementation. This is also confirmed by the Geneva Roadmap.[9] Although there is a wealth of literature showing the prognostic value of CSF and MRI biomarkers on a group level[5-7], these studies do not allow direct translation to the individual. For example, the prognostic value of biomarkers may be influenced by patient characteristics such as age, sex and cognitive status. To extract maximal information from each biomarker, the results should be interpreted in the context of these characteristics. However, these characteristics are often omitted in prognostic research. Furthermore, recommendations on how to handle conflicting and borderline results are lacking.[9] In this context, the novel National Institute on Aging and Alzheimer's Association (NIA-AA) research framework that defines AD as a biological construct is of great interest. The research framework proposes to use biomarkers for amyloid (A), tau (T), and neurodegeneration (N) to classify patients. For MCI, it is unknown how the use of this framework informs predictions.[10]

In a previous study we constructed, as a proof of principle, biomarker-based prognostic models that allow risk prediction on the individual level.[4] These models, which were based on a homogeneous, mono-center cohort, provide probabilities of progression to (AD)dementia in the course of one or three years of follow-up for any given value of each biomarker. To successfully enter clinical practice however, generalizability has to be demonstrated by extensive external validation.[11] A prerequisite for

generalizability is that the models are able to accommodate different biomarker measurement methods and have value for different cohorts, beyond the ones they were initially developed in.[9] Taking our initially developed risk prediction models as a starting point, the aim of this study was to establish robust prediction models in a multi-center, multi-cohort design. In addition, we constructed an ATN-model allowing the use of this framework to inform predictions.

Methods

Study design and participants

MCI patients were included from mono- and multi-center cohorts in Europe and America; the Amsterdam Dementia Cohort (ADC[12], n=666), the Alzheimer's Disease Neuroimaging Initiative (ADNI[13], n=829), the Biomarkers For Identifying Neurodegenerative Disorders Early and Reliably Study (BioFINDER[14], n=233) and the collaborative cohorts of the European Medical Information Framework for AD (EMIF-AD, n=883) composed of the following (multi)center studies: DESCRIPA[15], AddNeuroMed[16], German Dementia Competence Network (DCN[17]), IMAP [18], European Alzheimer's Disease Consortium (EADC)-PET[19], Brescia [20], Coimbra[21], Kuopio[22] and Lisbon[23]. In Table 1 the cohort characteristics are summarized. The cohort characteristics of the separate EMIF cohorts are shown in Supplemental Table 1 (page 2-3).

Inclusion criteria of the present study were a baseline diagnosis of MCI, at least 6 months of follow-up and availability of a baseline MMSE and MRI or CSF biomarkers. All participant gave written informed consent and institutional review boards approved the study. This study is reported in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guideline.[24]

Original prediction models

The original prediction models were constructed using Cox proportional hazards modelling in the Amsterdam Dementia Cohort.[4] In the current study, we validated the following previously published models*:* Demographic only model, HCV model and CSF model.[4] Variables included in the models and corresponding estimates are shown in Supplemental table 2 (page 4). In short, the demographic model included age, sex and MMSE, the HCV model included HCV (cm3), age and MMSE, and the CSF model included Abeta (1-42), total tau, MMSE and an interaction between Abeta and total tau. As whole brain volume was not available in one cohort (EMIF), we were unable to assess the performance of a model combining CSF and MRI features (i.e. combined model). In the original paper, the prognostic models showed moderate to good discrimination (Harrell's C's demographic model=0.59 [0.54-0.64], HCV model=0·73 [0·66-0·80] and CSF model=0·67 [0·67-0·81]). External

validation in ADNI-2 showed robustness of all models (Harrell's C's demographic model=0·67 [0·60-0·74], HCV model=0·73 [0·66-0·80] and CSF model=0·74 [0·67-0·81]).[4]

Part of the ANDI sample was used in the original paper, therefore we excluded these patients from the validation analyses, but included them in the model update (Figure 1).

Predictors

The following baseline predictors were available in all cohorts: patient characteristics (age, sex), Mini-Mental State Examination (MMSE) score, CSF biomarkers (Abeta, total tau (t-tau), phosphorylated tau (p-tau)) and Hippocampal Volume (HCV). In supplemental Figure 1 (page 5) the distributions of these predictors are shown across the different cohorts.

As absolute values of both CSF concentrations and volumetric MRI measures varied across methods, we bridged CSF and volumetric MRI data where possible. A detailed description of this bridging analysis is provided in the supplement (Supplemental text 1 (page 6) and Supplemental Figure 2 (page 7)).

Outcome

Clinical progression to any type of dementia was used as primary outcome. In a secondary analysis, we validated all models with AD-dementia as outcome (Supplemental Figure 3 (page 8) and Supplemental Table 3 (page 9)). The ADC and BioFINDER are memory clinic-based cohorts and patients were re-evaluated on a yearly basis (EMIF-AD substudy follow-up is reported in Supplemental Table 1 (page 2-3)). ADNI is a research cohort and diagnosis is evaluated on a 3 to 12-month interval.

Statistical analysis

We took the following four steps to validate and update our biomarker-based prediction models. First, model performances of the originally developed models were assessed in all cohorts with Harrell's concordance statistic. Second, we updated the models by 1) re-estimating parameters with and without center specific effects, in order to evaluate whether we could safely omit the adjustment for center which would increase generalizability and 2) replacing t-tau by p-tau in the CSF model according to

NIA-AA criteria. Third, we estimated a model including both amyloid, p-tau and HCV in accordance with the recently proposed ATN framework[10]. Lastly, we assessed calibration (concordance predicted with observed outcome) of the models by superimposing observed and expected survival predicted by the models. A detailed description of these steps can be found in the supplement (Supplemental text 2, page 10-12). Analyses were performed in STATA SE 14 and were based on complete cases and therefore the number of patients varies across models (Figure 1).

Role of the funding source

Funders of this study had no involvement in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Participants

We included 2611 subjects of which 1007 (39%) subjects progressed to dementia. Mean age was 70±8, mean MMSE was 27±2 and 44% (n=1153) were female.

Model Performance

Figure 2 shows the pooled model performance for each model. The original model performances are shown in grey. For the demographics only (0·62 [0·59-0·65]), HCV model (0·67 [0·62-0·72]) and the CSF model (0·67 [0·64-0·71] the pooled Harrell's C's are similar to those found in the original study. The results for AD-dementia as outcome are shown in Supplemental figure 3 (page 8). For AD-dementia as outcome, the pooled Harrell's C of the CSF model is lower than in the original study, indicating possible misfit (0·69 [0·65-0·72]).

Model update

Re-estimating the parameters did not increase model fit for the demographics only and HCV model, both with and without center specific effects (Table 3). For the CSF model, re-estimating the parameters increased model fit. In none of the models, inclusion of center specific effects improved the models relative to the models without center specific effects (Table 3). Results from an additional set of analyses further supported this finding, as we found that center specific effects were not confounded by measurement methods for MRI and CSF (supplementary data). Therefore, the models without center specific effects are favored over a model with center specific effects, as this increases generalizability. Replacing t-tau by p-tau in the CSF model did not affect model performance (Table 3).

ATN model

In Table 4 we present a novel ATN model (validation procedure is shown in Supplemental table 4 (page 13), results for AD-dementia are shown in Supplemental table 5 (page 14)). Main effects of Abeta, p-tau, HCV, age and MMSE retained ($p<0.10$). Moreover, interaction effects between Abeta and p-tau, Abeta and age, and p-tau and MMSE were included ($p<0.10$). The interactions indicate that

the prognostic value of Abeta was stronger in younger patients. The prognostic value of p-tau was most pronounced in patients with higher (normal) Abeta values and lower (abnormal) MMSE values. Harrell's C-statistic was 0·74 [0·71-0·76].

Calibration

Figure 3 shows the superimposed expected and observed survival curves for different risk groups. In general, for all models, the lines of the observed and expected survival appear to be similar, indicating good calibration. For the HCV model, visual inspection suggests a degree of misfit for very long term predictions (>5 years), as the model tends to overestimate survival in the good group and underestimate survival in the poor group. Of note, all models are well calibrated up to five years of follow up.

Clinical use

As all models are well calibrated up to five years of follow-up, we now updated the models to provide five year risk estimates in addition to the one and three year risk estimates provided in the original paper. A spreadsheet calculator can be provided by the authors on request. This calculator allows the user to select which platform was used for CSF analysis (Innotest, Luminex or Elecsys) and which method was used to calculate hippocampal volume (FSL FIRST or Freesurfer). After selecting the appropriate methods for CSF and MRI, clinicians can easily fill in patient specific values.

For example, for a female MCI patient (age=62, MMSE=26) without knowledge of biomarker results the progression probabilities to dementia are 11% [10-12] in one year, 39% [36-42] in three years and 57% [52-61] in five years. When both MRI and CSF would be available, the progression probabilities change to 40%[33-48] at one year, 88%[82-94] at three years and 97%[94-99] at five years if abnormal biomarkers are abnormal (Abeta =112, p-tau=35 (CSF measured with Elecsys) and HCV= 6.2 (calculated with Freesurfer software).

On the other end of the spectrum, a male MCI patient (age=62, MMSE=29) without knowledge of biomarker results, has progression probabilities to dementia of 7%[6-8] in one year, 26%[23-29] in three years and 40% [44-35] in five years. With normal biomarkers (Abeta=1264, p-tau=12 (Elecsys)

and HCV=9.8 (Freesurfer)) this patient would have progression probabilities of in 1%[1-2] one year, 5%[4-7] in three years and 8%[6-11] in five years. Particularly this strong negative predictive value of these biomarker results (i.e. reassure when normal) may have immediate clinical relevance.

Discussion

We have constructed and validated biomarker-based models, including an ATN-model, to provide predictions for dementia in individual MCI patients. We showed that the models had strong external validity, being generalizable across continents and memory clinic cohorts. Moreover, the models accommodate different assays which further increases their generalizability. Depending on the preferences and needs of the clinicians and patients, the models can be used to extract individually tailored prognostic information from the tests which have been performed in the diagnostic set-up. By doing so, we take the first crucial steps on the road towards a precision medicine approach.

Our study has important clinical implications. Patients and caregivers become increasingly assertive in their need for (prognostic) information. In clinical practice however, risk communication in MCI patients is only sparsely observed and if communicated, these are mostly group averages; "being an MCI patient, you have a fifty-fifty percent risk of progression to dementia". With biomarker results available, this fifty-fifty situation for most patients is not true however. As with abnormal biomarkers, the risk of progression may be higher, while with normal biomarkers this risk can be (far) lower. With our validated, biomarker-based prediction models, a prognosis for an individual patient can be estimated in the context of their own characteristics, showing that precision medicine for AD may be on the horizon. The models are easy to use and a calculator (simple excel sheet) for academic use can be provided by the authors upon request. To further facilitate clinical use, we incorporated the models in an easy to use online tool (ADappt; https://www.alzheimercentrum.nl/professionals/adappt-contact/).[25]

However, there are also arguments against the disclosure of risk in clinical practice. A recent review on the disclosure of amyloid PET results in pre-dementia patients showed that these arguments are to a large extent theoretical in nature and relate mostly to the principle of non-maleficence (i.e. do no harm).[26] Empirical evidence for this is largely lacking and the effect on psychological harm is not known. In a previous ABIDE study, patients and caregivers expressed their need for risk communication in early phases of AD and anxiety or uncertainty did not increase after disclosure of

amyloid PET.[3, 27] This suggests that it is conceivable that models such as developed in this manuscript could be used in clinical practice. Nonetheless before this type of model could be implemented in daily practice, there are some important next steps to take, particularly to determine clinical utility. In the current study, we used retrospective data to construct the models. As a first next step, the models should be evaluated prospectively, ideally in a phase 3 RCT design. This RCT should provide important answers on clinical utility of the models, particularly if their use has impact on a patients' understanding, emotional wellbeing, and behaviour (e.g. lifestyle changes).

In parallel, studies should focus on the optimal way to disclose risks to non-demented individuals, and it is conceivable that clinicians should receive training on how best to disclose this type of probabilistic information. Moreover, before initiating biomarker testing, it is of utmost importance that realistic expectations are set with regard to what kind of results can be anticipated. One could also imagine a different scenario, where the risk prediction models would actually be used before initiating biomarker testing. By filling in hypothetical biomarker results and comparing these to the results of the demographic only model, the clinician can evaluate whether or not these results would add prognostic value. The clinician could also engage the patient and caregiver in this discussion on different biomarker scenario's and potential outcomes. In this light, the models could serve as a decision support tool and could even enhance shared decision making.

We included multiple mono- and multicenter cohorts, both from Europe and the USA. Although we did not find heterogeneity in the baseline hazard and baseline survivor function (data not shown), differences inevitably exist between cohorts. For that reason, we thoroughly tested for center-effects. We found that adding center-effects did not improve the performance of the models, nor did it result in a difference in progression probabilities on an individual level (data not shown). According to the principle of parsimony, a model without center specific effects is preferable, as this allows the clinician to use the model without further adjusting the model to their own memory clinic. Moreover, this indicates that our models are also applicable for MCI patients in other memory clinics which were not included in the development or validation phase of our study. Secondly, it appeared that in the original CSF model, the parameters of Abeta and t-tau were overestimated, leading to less optimal

model performance in other cohorts. Re-estimating the parameters resulted in an increase in model performance. As a measure of amyloid, we used CSF concentrations rather than amyloid PET. Although of interest, amyloid PET is currently less often used in clinical practice, and is usually evaluated in a dichotomous fashion, while in the current models we include all biomarkers as continuous measures, with the objective to make it readily available for clinicians. We have developed amyloid PET based models in an earlier study however, and are therefore confident that results would generalize to amyloid PET as well.[28] In our updated models, we replaced t-tau by p-tau to improve alignment with the recently published NIA-AA criteria.[10] As CSF t-tau and p-tau are very highly correlated, this replacement did not influence the model performance. One could debate whether APOE would have been a helpful addition to the models. Although APOE e4 is the strongest genetic risk factor for AD, we decided not to include this genetic characteristic however, as APOE is currently not used in clinical practice, and likewise is mentioned in none of the diagnostic guidelines. Of note, in a former paper, we found that including APOE e4 status as an additional variable in biomarker-based models to predict dementia in MCI did not increase prognostic performance or alter the predictions on an individual level.[28]

The recently launched NIA-AA research framework states that by coding research participants according to the AT(N) system, the field moves in the direction of precision medicine.[10] This coding system highly depends on cut-off values, as a patient is either positive or negative for a specific biomarker. As a consequence of this dichotomy, the AT(N) system comprises eight categories. For clinical practice however, the use of eight categories is rather complex. But simultaneously, reality may be even more complicated than eight categories as the dichotomization does not include information on extent of abnormality. In the current study we present a model in which A, T and N biomarkers are simultaneously taken into account, yet can be entered in a continuous fashion, to yield risk estimation of disease progression to dementia in individual MCI patients. By doing so, every combination is possible and maximum information from each biomarker is exploited. To further foster clinical usefulness, our models provide probabilities of progression within a specific time frame, while taking patient characteristics into account. The NIA-AA coding scheme does not provide this type of

information yet. From risk communication literature we know that a numeric format of risk communication is preferred above verbal formats (high, intermediate, low), as verbal formats are sensitive to a high degree of variability in interpretation.[29] Accompanying the risk estimate with a time frame is considered best practice, ideally with a visual representation.[29]

The current models are updated to allow the use of raw values of different platforms for CSF and two widely used methods of hippocampal volume calculation, further promoting generalizability. With regards to CSF, the field is currently shifting away from manual assays like Luminex xMAP and Innotest ELISA, towards automated platforms like Elecsys and Lumipulse. In the current study, we used a recently published equation to bridge Innotest values to Elecsys values.[30] We used the same method to bridge Luminex to Elecsys values. For the calculation of brain volumes, there is more variation in software. We were able to bridge FSL FIRST data to Freesurfer. These two software packages are widely used, easily available and have a clear pipeline.

A potential limitation of bridging different types of data, is that it may cause additional noise on the risk prediction. However, this did not negatively affect the prognostic performance. Another limitation is that we used complete cases only in the analyses, resulting in sample size variations and might introduce a degree of bias.[24] Lastly, the cohorts used in this study inevitably differed not only in the definition of the predictors, but also in the outcome of AD-dementia. In validating prediction models, such differences may be intentional for two reasons.[24] If we would like our models to have clinical usability, we should align with clinical practice. And in clinical practice, differences in the definition of AD dementia are inevitable. Second, using different definitions in the outcome measure of our analysis will give us information on whether the models can be extrapolated to different populations.

Among the strengths of our study is the size and heterogeneity of the cohorts used. Moreover, prediction models, especially when constructed with Cox proportional hazards analysis, are often not validated to the extent that we did.[11] We thoroughly tested for center-effects and concluded that adjustment for center could safely be omitted. This finding greatly enhances the generalizability and therefore the clinical applicability of our models, since it implies that the models can also be applied to

centers not included in the current study. Of note, the models are specifically of relevance for memory clinics and perhaps in a trial setting, and cannot (yet) be extrapolated to for example general practitioners. For risk stratification purposes, discrimination between those who will and those who won't progress to dementia is clearly the key indicator of model success or failure. But for a model to be used in clinical practice and to provide patients with probabilistic information, calibration (i.e. concordance between predicted and observed outcome) is very important as well. In the evaluation of prediction models, this aspect is often neglected as studies do not report the baseline survival function. As we ultimately want our study to support clinical practice, we performed a strict type of calibration assessment, leading us to conclude that the models are well calibrated for predictions well beyond five years.[11]

In conclusion, we have constructed and validated biomarker-based models for prediction of progression to dementia in MCI patients. We have shown the generalizability and robustness of the predictions and the models are made freely available by the authors upon request. The outcomes of this study facilitate a more timely and accurate diagnosis which is of high importance at the individual level even in the absence of specific therapies, as this is the starting point to plan and organize care.

Funding

Research of the Alzheimer center Amsterdam is part of the neurodegeneration research program of Amsterdam Neuroscience. The Alzheimer Center Amsterdam is supported by Stichting Alzheimer Nederland and Stichting VUmc fonds. W.M.vd.F. is recipient of a grant by Stichting LSH-TKI (ABIDE-clinical utility: LSHM18075, cofunded by Life-MI) This study is funded by ZonMW-Memorabel (ABIDE; project no 733050201), a project in the context of the Dutch Deltaplan Dementie.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.;Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.;Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

This study resulted from a collaboration between centers of the European Alzheimer's Disease Consortium (EADC). The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under EMIF grant agreement n° 115372, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.

The DESCRIPA study was funded by the European Commission within the 5th framework program (QLRT-2001- 2455).

The Coimbra Centre was funded by the European Regional Development Fund (ERDF), through the Centro 2020 Regional Operational Programme under project CENTRO-01-0145-FEDER-000008:BrainHealth 2020, and through the COMPETE 2020 - Operational Programme for Competitiveness and Internationalisation and Portuguese national funds via FCT – Fundação para a Ciência e a Tecnologia, I.P., under project[s] POCI-01-0145-FEDER-007440.

The Memory Clinic (Centre de la mémoire) of Geneva University Hospitals is funded by; A.P.R.A. - Association Suisse pour la Recherche sur Alzheimer, Genève; Fondation Segré, Genève; Ivan Pictet, Genève; Fondazione Agusta, Lugano; Fondation Chmielewski, Genève; Fondation VELUX; Swiss National Science Foundation.

The IMAP Study (G Chételat, Caen, France) was funded by Programme Hospitalier de Recherche Clinique (PHRCN 2011-A01493-38 and PHRCN 2012 12-006-0347); Agence Nationale de la

Recherche (LONGVIE 2007); Région Basse-Normandie and Fondation Plan Alzheimer (Alzheimer Plan 2008-2012).

The Dementia Competence Network (DCN) has been supported by a grant from the German Federal Ministry of Education and Research (BMBF): Kompetenznetz Demenzen (01GI0420).

The BioFINDER study was supported by the European Research Council, the Swedish Research Council, the Knut and Alice Wallenberg foundation, the Marianne and Marcus Wallenberg foundation, the Strategic Research Area MultiPark (Multidisciplinary Research in Parkinson's disease) at Lund University, the Swedish Alzheimer Foundation, the Swedish Brain Foundation, The Parkinson foundation of Sweden, the Skåne University Hospital Foundation, and the Swedish federal government under the ALF agreement.

Declaration of interests

Dr. Teunissen has functioned in advisory boards of Fujirebio and Roche, received nonfinancial support in the form of research consumables from ADxNeurosciences and Euroimmun, performed contract research or received grants from Probiodrug, Janssen Prevention Center, Boehringer, Brains Online, Axon Neurosciences, EIP Pharma, and Roche.

Dr. Scheltens has acquired grant support (for the institution) from GE Healthcare, Danone Research, Piramal, and MERCK. In the past 2 years, he has received consultancy/speaker fees (paid to the institution) from Lilly, GE Healthcare, Novartis, Sanofi, Nutricia, Probiodrug, Biogen, Roche, Avraham, and EIP Pharma.

Dr. Barkhof is supported by the NIHR biomedical research centre at UCLH.

Dr. Frolich: research funding, consultancy fees or speech honoraria from Allergan, Avid-Eli Lilly, Avanir, Avraham Pharmaceuticals, Axon Neuroscience, Axovant, Biogen, Boehringer Ingelheim, Eisai, Functional Neuromodulation, GE Health Care, Lundbeck, MerckSharpe&Dohme, Novartis, Pfizer, Piramal Imaging, Roche, Schwabe Pharma.

Dr. Wiltfang has acquired research support (for the institution) from Immungenetics, TECAN-IBL. Consultancy fees from Lilly, Roche Pharma, MSD SHARP & DOHME, Boehringer-Ingelheim, Abbot EPD. Speech honoraria Roche Pharma, Helios Klinikum Wuppertal, Vitos Kurhessen-Bad Emstal, Pfizer, Janssen, AGNP, Actelion. Patents: PCT/EP 2011 001724: New formulations for diagnosis of Alzheimer's disease . PCT/EP 2015 052945: Biosensor for conformation and secondary structure analysis.

Dr. Peters reports grants and personal fees from Roche, grants and personal fees from Biogen, grants from Novartis, grants from Eisai, grants from Lilly, grants from Pharmatrophix, outside the submitted work.

Dr. Harald Hampel serves as Senior Associate Editor for the Journal Alzheimer's & Dementia; he received lecture fees from Biogen, Roche, Eisai, research grants from Pfizer, Avid, and MSD Avenir (paid to the institution), travel funding from Functional Neuromodulation, Axovant, Eli Lilly and company, Takeda and Zinfandel, GE-Healthcare and Oryzon Genomics, consultancy fees from Qynapse, Jung Diagnostics, Cytox Ltd., Axovant, Anavex, Takeda and Zinfandel, GE Healthcare and Oryzon Genomics, and Functional Neuromodulation, and participated in scientific advisory boards of Functional Neuromodulation, Axovant, Eli Lilly and company, Cytox Ltd., GE Healthcare, Takeda and Zinfandel, Oryzon Genomics and Roche Diagnostics.

H.H. is supported by the AXA Research Fund, the “Fondation partenariale Sorbonne Université” and the “Fondation pour la Recherche sur Alzheimer”, Paris, France. Ce travail a bénéficié d'une aide de l’Etat “Investissements d’avenir” ANR-10-IAIHU-06. The research leading to these results has received funding from the program “Investissements d’avenir” ANR-10-IAIHU-06 (Agence Nationale de la Recherche-10-IA Agence Institut Hospitalo-Universitaire-6).

H.H. is co-inventor in the following patents as a scientific expert and has received no royalties:

• In Vitro Multiparameter Determination Method for The Diagnosis and Early Diagnosis of Neurodegenerative Disorders Patent Number: 8916388

• In Vitro Procedure for Diagnosis and Early Diagnosis of Neurodegenerative Diseases Patent Number: 8298784

• Neurodegenerative Markers for Psychiatric Conditions Publication Number: 20120196300

• In Vitro Multiparameter Determination Method for The Diagnosis and Early Diagnosis of Neurodegenerative Disorders Publication Number: 20100062463

• In Vitro Method for The Diagnosis and Early Diagnosis of Neurodegenerative Disorders Publication Number: 20100035286

• In Vitro Procedure for Diagnosis and Early Diagnosis of Neurodegenerative Diseases Publication Number: 20090263822

• In Vitro Method for The Diagnosis of Neurodegenerative Diseases Patent Number: 7547553

• CSF Diagnostic in Vitro Method for Diagnosis of Dementias and Neuroinflammatory Diseases Publication Number: 20080206797

• In Vitro Method for The Diagnosis of Neurodegenerative Diseases Publication Number: 20080199966

• Neurodegenerative Markers for Psychiatric Conditions Publication Number: 20080131921

Dr. Vellas reports grants and personal fees from Biogen, MSD, Lily and Roche, outside the submitted work.

Dr. Lovestone has, within the past 5 years, held research grants with funding from multiple industry partners through the IMI funding scheme and also with Astra Zeneca. He is currently an employee of Janssen R&D.

Dr. Chetelat reports grants from Institut National de la Santé et de la Recherche Médicale (Inserm), grants from Fondation Plan Alzheimer (Alzheimer Plan 2008-2012), grants from Programme Hospitalier de Recherche Clinique (PHRCN 2011-A01493-38 and PHRCN 2012 12-006-0347), grants from Agence Nationale de la Recherche (LONGVIE 2007), grants from Région Basse-Normandie, grants from Association France Alzheimer et maladies apparentées, grants from Fondation Vaincre Alzheimer, from The European Commission (H2020 PHC22, Medit-Ageing study, grant agreement N667696), outside the submitted work; .

Dr. Hansson has acquired research support (for the institution) from Roche, GE Healthcare, Biogen, AVID Radiopharmaceuticals, Fujirebio, and Euroimmun. In the past 2 years, he has received consultancy/speaker fees (paid to the institution) from Biogen, Roche, and Fujirebio.

Dr. van der Flier performs contract research for Biogen MA Inc. Research programs of W.M.v.d.F. have been funded by ZonMW, Health Holland, Pasman stichting, NWO, EU-FP7, EU-JPND, Alzheimer Nederland, Cardiovasculair Onderzoek Nederland, stichting Dioraphte, Gieskes-Strijbis

fonds, Boehringer Ingelheim, Piramal Neuroimaging, Roche BV, Janssen Stellar, and Combinostics. All funding is paid to her institution.

Van Maurik, Bos. Vos, Bouwman, Kornhuber, Maier, Peters, Rüther, Nobili, Frisoni, Spiru, Freund-Levi, Wallin, Soininen, Tsolaki, Verhey, Kloszewska , Mecocci, Galluzzi, Herukka, Santana, Baldeiras, de Mendonça, Silva, Chetelat, Egret, Palmqvist, Visser and Berkhof report no conflict of interests.

Data sharing

The corresponding author can provide the dataset used and/or documentation on the analysis performed upon reasonable request.

Author contributions

ISvM, JB and WMvdF contributed to the study design. ISvM did the literature search, analyzed the data and created the figures. All authors contributed to the data collection and interpretation of results, reviewed and critically revised the manuscript, and approved the final version for submission.

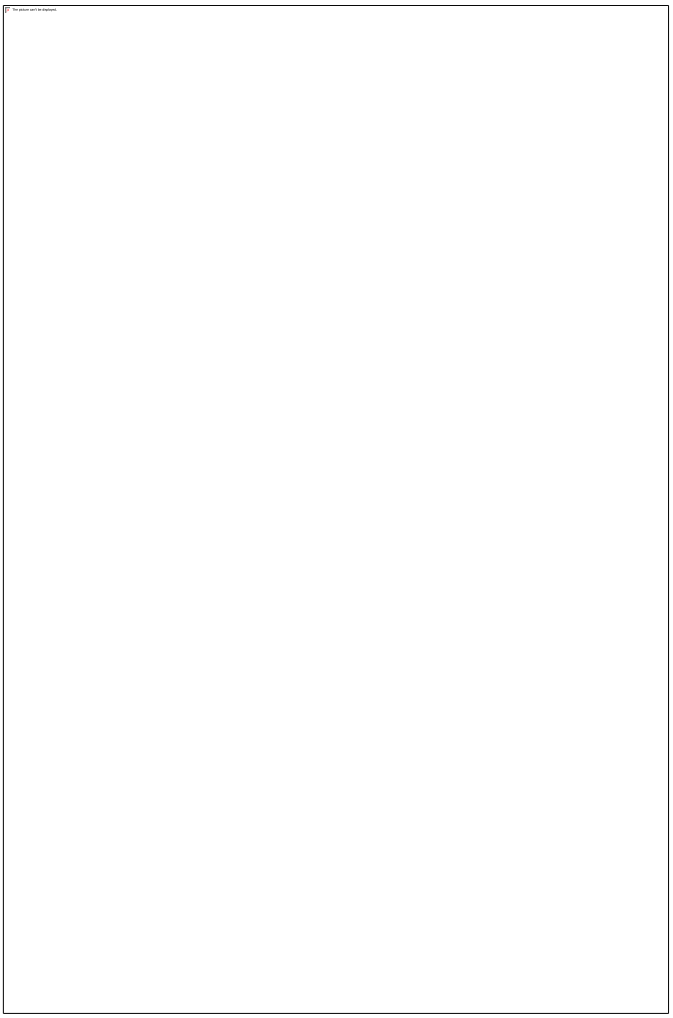
References

1. Scheltens P, Blennow K, Breteler MM, et al. Alzheimer's disease. Lancet. 2016;388(10043):505-517.
2. Vos SJ, Verhey F, Frolich L, et al. Prevalence and prognosis of Alzheimer's disease at the mild cognitive impairment stage. Brain. 2015;138(Pt 5):1327-1338.
3. Kunneman M, Pel-Littel R, Bouwman FH, et al. Patients' and caregivers' views on conversations and shared decision making in diagnostic testing for Alzheimer's disease: The ABIDE project. Alzheimers Dement (N Y). 2017;3(3):314-322.
4. van Maurik IS, Zwan MD, Tijms BM, et al. Interpreting Biomarker Results in Individual Patients With Mild Cognitive Impairment in the Alzheimer's Biomarkers in Daily Practice (ABIDE) Project. JAMA Neurol. 2017;74(12):1481-1491.
5. Handels RLH, Vos SJB, Kramberger MG, et al. Predicting progression to dementia in persons with mild cognitive impairment using cerebrospinal fluid markers. Alzheimers Dement. 2017;13(8):903-912.
6. van Harten AC, Visser PJ, Pijnenburg YA, et al. Cerebrospinal fluid Abeta42 is the best predictor of clinical progression in patients with subjective complaints. Alzheimers Dement. 2013;9(5):481-487.
7. Davatzikos C, Bhatt P, Shaw LM, Batmanghelich KN, Trojanowski JQ. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. Neurobiol Aging. 2011;32(12):2322 e2319-2327.
8. Petersen RC, Lopez O, Armstrong MJ, et al. Practice guideline update summary: Mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology. 2018;90(3):126-135.
9. Frisoni GB, Boccardi M, Barkhof F, et al. Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. Lancet Neurol. 2017;16(8):661-676.
10. Jack CR, Jr., Bennett DA, Blennow K, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimers Dement. 2018;14(4):535-562.
11. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. Bmc Med Res Methodol. 2013;13.
12. van der Flier WM, Pijnenburg YA, Prins N, et al. Optimizing patient care and research: the Amsterdam Dementia Cohort. J Alzheimers Dis. 2014;41(1):313-327.
13. Mueller SG, Weiner MW, Thal LJ, et al. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). Alzheimers Dement. 2005;1(1):55-66.
14. Palmqvist S, Zetterberg H, Mattsson N, et al. Detailed comparison of amyloid PET and CSF biomarkers for identifying early Alzheimer disease. Neurology. 2015;85(14):1240-1249.
15. Visser PJ, Verhey FR, Boada M, et al. Development of screening guidelines and clinical criteria for predementia Alzheimer's disease. The DESCRIPA Study. Neuroepidemiology. 2008;30(4):254-265.
16. Lovestone S, Francis P, Kloszewska I, et al. AddNeuroMed--the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. Ann N Y Acad Sci. 2009;1180:36-46.
17. Kornhuber J, Schmidtke K, Frolich L, et al. Early and differential diagnosis of dementia and mild cognitive impairment: design and cohort baseline characteristics of the German Dementia Competence Network. Dement Geriatr Cogn Disord. 2009;27(5):404-417.
18. Arenaza-Urquijo EM, Bejanin A, Gonneaud J, et al. Association between educational attainment and amyloid deposition across the spectrum from normal cognition to dementia: neuroimaging evidence for protection and compensation. Neurobiol Aging. 2017;59:72-79.
19. Morbelli S, Drzezga A, Perneczky R, et al. Resting metabolic connectivity in prodromal Alzheimer's disease. A European Alzheimer Disease Consortium (EADC) project. Neurobiol Aging. 2012;33(11):2533-2550.
20. Frisoni GB, Prestia A, Zanetti O, et al. Markers of Alzheimer's disease in a population attending a memory clinic. Alzheimers Dement. 2009;5(4):307-317.

21. Santana I, Baldeiras I, Santiago B, et al. Underlying Biological Processes in Mild Cognitive Impairment: Amyloidosis Versus Neurodegeneration. J Alzheimers Dis. 2018;64(s1):S647-S657.
22. Seppala TT, Koivisto AM, Hartikainen P, et al. Longitudinal changes of CSF biomarkers in Alzheimer's disease. J Alzheimers Dis. 2011;25(4):583-594.
23. Maroco J, Silva D, Rodrigues A, et al. Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. BMC Res Notes. 2011;4:299.
24. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015;350:g7594.
25. van Maurik IS, Visser LNC, Pel-Littel R, et al. Development and usability of ADappt – an online tool to support clinicians, patients and caregivers in the diagnosis of Mild Cognitive Impairment and Alzheimer's disease. JMIR. 2019;unreviewd preprint.
26. de Wilde A, van Buchem MM, Otten RHJ, et al. Disclosure of amyloid positron emission tomography results to individuals without dementia: a systematic review. Alzheimers Res Ther. 2018;10(1):72.
27. Albert MS, DeKosky ST, Dickson D, 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(3):270-279.
28. van Maurik IS, van der Kall LM, de Wilde A, et al. Added value of amyloid PET in individualized risk predictions for MCI patients. Alzheimers Dement (Amst)

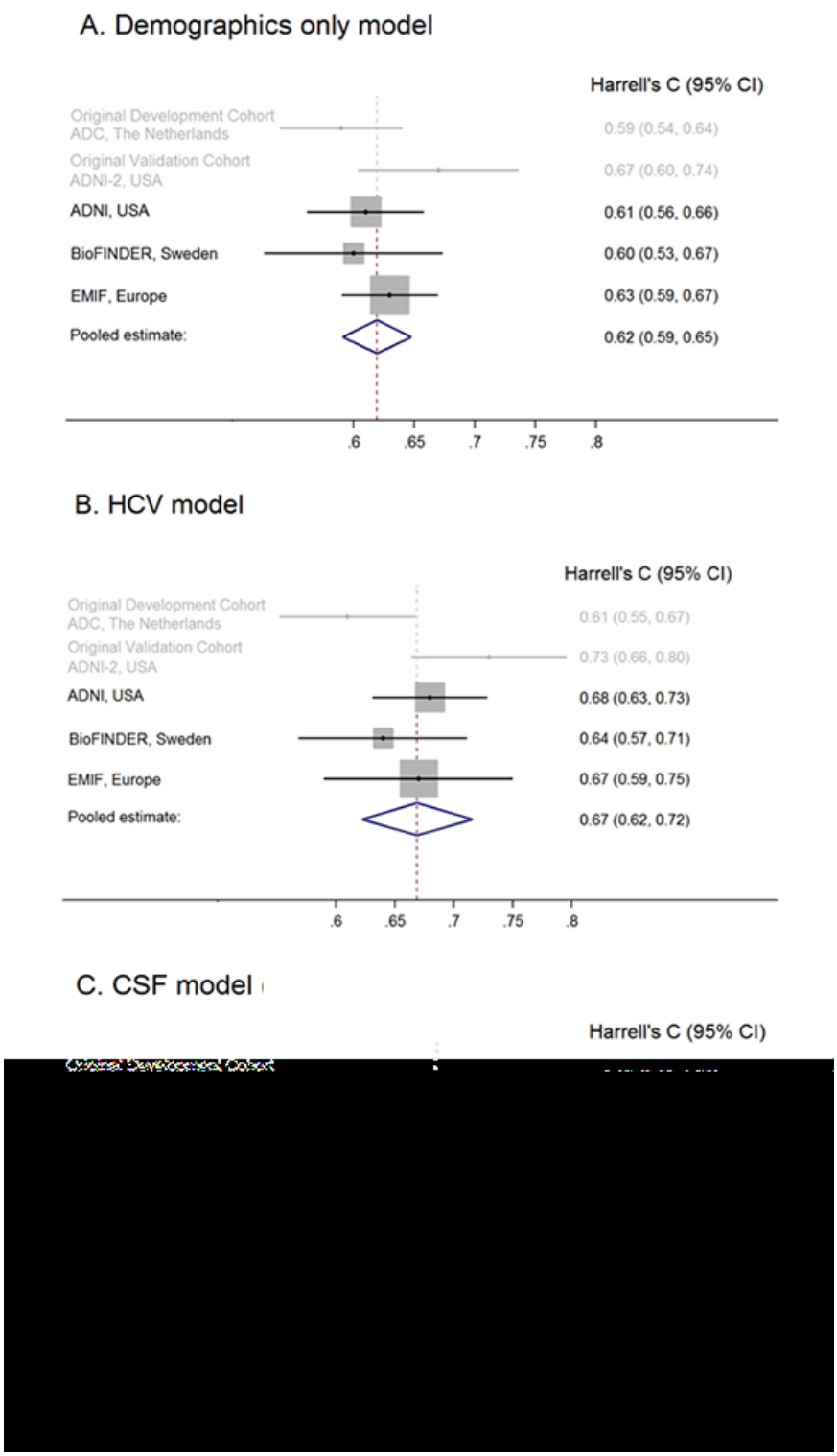
2019;in press.
29. Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. Med Decis Making. 2007;27(5):696-713.
30. Willemse EAJ, van Maurik IS, Tijms BM, et al. Diagnostic performance of Elecsys immunoassays for cerebrospinal fluid Alzheimer's disease biomarkers in a nonacademic, multicenter memory clinic cohort: The ABIDE project. Alzheimers Dement (Amst). 2018;10:563-572.

Figure 1. Flow chart of samples used



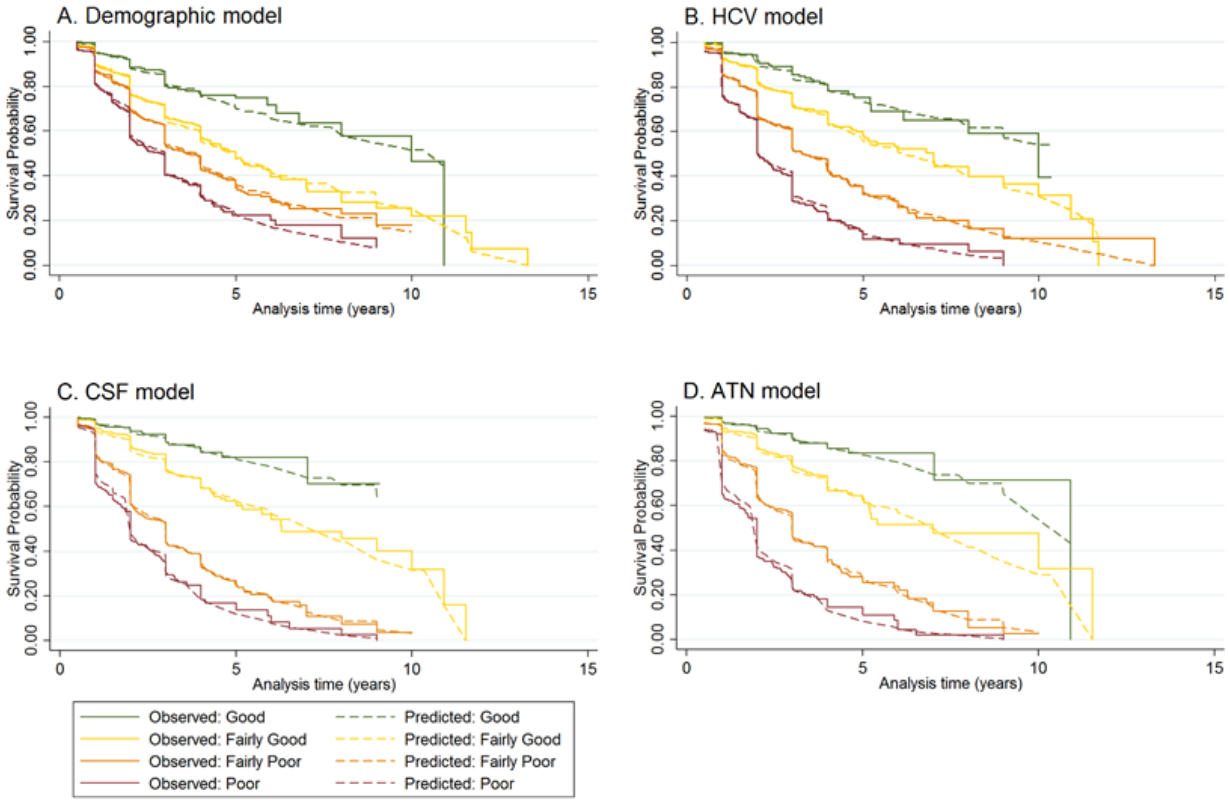
ADC= Amsterdam Dementia Cohort, ADNI= Alzheimer's Disease Neuroimaging Initiative, CSF=cerebrospinal fluid, EMIF-AD= European Medical Information Framework, HCV=hippocampal volume. Flow of subjects included in the validation analyses and model updates.

Figure 2. Model performance of published models



ADC= Amsterdam Dementia Cohort, ADNI= Alzheimer's Disease Neuroimaging Initiative, CSF=cerebrospinal fluid, EMIF-AD= European Medical Information Framework for AD, HCV=hippocampal volume. As a reference, the model performance of the original development and validation cohort are shown in grey. Pooled estimates of model performance for AD-dementia as clinical endpoint are shown in Supplemental Figure 3 (page 8).

Figure 3. Calibration of biomarker-based models



ATN=Amyloid (Abeta), tauopathy (p-tau), neurodegeneration (HCV), CSF=cerebrospinal fluid, HCV=hippocampal volume. Risk groups were made based on the PI determined, resulting in a good (>84th percentile), fairly good (50-84th percentile), fairly poor (16-50th percentile) and poor prognosis (<16th percentile group).
Solid lines: Observed progression rates (Kaplan-Meier), dashed lines: predicted progression by the cox models. Findings are based on data from all four cohorts. Calibration of model performance for dementia as clinical endpoint are shown in Supplemental Figure 4 (page 15).

Table 1. Study Characteristics

Characteristic	**ADC**	**ADNI**	**EMIF-AD**	**BioFINDER**
N	666	829	883	233
Baseline data collection period	1995-2014	2004-2014	Varied per substudy; supplemental table 1 (page 2-3).	2010-2015
Study design	Mono-center cohort study	Multicenter longitudinal cohort study	Multicenter longitudinal cohort study	Multicenter longitudinal cohort study
Setting	Tertiary memory clinic	Research	Memory clinics	Memory clinics
Inclusion criteria	Referred to memory clinic, do not fulfill criteria for dementia.	Memory complaints verified by study partner, Abnormal memory functioning, MMSE between 24-30, CDR=0·5, do not fulfill criteria for dementia	Varied per substudy; supplemental table 1 (page 2-3).	Referred to memory clinic, age between 60-80, baseline MMSE 24-30, do not fulfill criteria for dementia
Patients with outcome dementia	288 (43%)	319 (38%)	272 (31%)	128 (55%)
Follow-up	Clinical follow-up every 12 months	3- to 12 month interval	Varied per substudy; supplemental table 1 (page 2-3).	Every 12 months for at least 6 years
MRI data available	539 (81%)	705 (85%)	727 (82%)	233 (100%)
MRI quantification method	FSL-FIRST, Freesurfer version 5·3	Freesurfer version 5·3	Varied per substudy; supplemental table 1 (page 2-3).	Freesurfer version 5·3
CSF data available	485 (73%)	558 (67%)	366 (41%)	221 (95%)
CSF platform	Innotest	Luminex and Elecsys	Innotest	Innotest

AD= Alzheimer's disease, ADC= Amsterdam Dementia Cohort, ADNI= Alzheimer's Disease Neuroimaging Initiative, CDR=clinical dementia rating scale, CSF= cerebrospinal fluid, EMIF= European Medical Information Framework, MMSE= mini-mental state examination, MRI= magnetic resonance imaging,

Table 2. Cohort characteristics.

	Original sample[4]		New validation sample		
	ADC, Netherlands	***ADNI-2, USA***	**ADNI, USA**	**EMIF-AD, Europe**	**BioFINDER, Sweden**
	n=485	***n=299***	**n=530**	**n=883**	**n=233**
No. Progressors	*243 (50%)*	*88 (29%)*	231 (43%)	272 (31%)	128 (55%)
AD dementia	*195 (40%)*	*85 (28%)*	223 (42%)	218 (25%)	87 (37%)
Other types of dementia	*48 (10%)*	*3 (1%)*	8 (2%)	54 (6%)	41 (18%)
Follow-up time	*2·4±1·6*	*2·6±1·4*	3·3±2·*4*	2·2±1·*1*	2·3±1·*3*
Age	*67±8*	*71±7*	73±8	69±8	71±5
Sex (F)	*192 (40%)*	*132 (45%)*	204 (38%)	461 (52%)	97 (41%)
MMSE	*27±2*	*28±2*	27±2	27±2	27±2
Hippocampal volume (cm3)	*6·9±1·1**	*6·9±1·1*	6·6±1·*1*	0·02±0·*99*[#]	6·7±1·*2*
CSF abeta (pg/mL)	*876±547**	*872±322**	990±571	913±603	635±407
CSF t-tau (pg/mL)	*256±141**	*280±131**	293±126	230±111	222±80
CSF p-tau (pg/mL)	*27±16*	*27±15*	29±15	25±16	25±14

*Values are bridged and do therefore not correspond with the values from the original paper
#HCV in the EMIF cohort was measured with different techniques than FSL-FIRST or Freesurfer, therefore the values were not bridged but converted to z-scores. Note that for the ADC cohort we here present the characteristics of the original sample. For the current study, n=181 (n=45 (25%) progressors) new patients were included, making the total sample size for ADC n=666.
AD=Alzheimer's disease, F=female, MMSE=mini-mental state examination, CSF=Cerebrospinal fluid, ADC=Amsterdam dementia cohort, ADNI=Alzheimer's disease neuroimaging initiative, EMIF-AD=European Medical Information Framework for AD.

Table 3. Harrell's C of published models

Model	Original parameters	Refitted parameters without Center Specific effects	Refitted parameters with Center Specific effects
Demographics only	0.62 [0.59-0.65]	0·63 [0·61-0·65]	0·65 [0·64-0·68]
HCV model	0.67 [0.62-0.72]	0·69 [0·67-0·71]	0·69 [0·67-0·72]
CSF model	0.67 [0.64-0.71]	0·72 [0·68-0·74]	0·72 [0·70-0·74]
CSF model with p-tau	NA	0·72 [0·70-0·74]	0·72 [0·69-0·74]

Presented data are Harrell's concordance statistics [95%CI]. Outcome was progression to any type of dementia. CSF=Cerebrospinal fluid; HCV=hippocampal volume; NA=not applicable. Model performances of the models for AD dementia as clinical endpoint are shown in supplemental Table 3.

Table 4. Partial regression coefficients and model performance of the ATN model

ATN model	Partial Regression Coefficients	95%CI	Harrell's C
Abeta	-0·5187	[-0·633 – -0·405]	
p-tau	0·6207	[0·439 – 0·802]	
HCV	-0·4164	[-0·516 – -0·317]	
Age	-0·0065	[-0·020 – 0·007]	0·74
MMSE	-0·1107	[-0·151 – -0·070]	[0·71-0·76]
Abeta*p-tau	0·1772	[-0·024 – 0·378]	
Abeta*age	0·0166	[-0·002 – 0·035]	
p-tau*MMSE	0·0928	[0·019 – 0.167]	

HCV= hippocampal volume; MMSE=Mini-mental state examination. Model is based on cross validated estimates from all cohorts. ATN model for AD dementia as clinical endpoint is shown in supplemental Table 5.