ELSEVIER

Alzheimer's & Dementia 9 (2013) 262–268

Alzheimer's & Dementia

A prediction model to calculate probability of Alzheimer's disease using cerebrospinal fluid biomarkers

Petra E. Spies[a,b,c,*], Jurgen A. H. R. Claassen[a,b,c], Petronella G. M. Peer[d], Marinus A. Blankenstein[e], Charlotte E. Teunissen[e], Philip Scheltens[f], Wiesje M. van der Flier[f,g], Marcel G. M. Olde Rikkert[a,b,c], Marcel M. Verbeek[b,c,h,i]

[a]*Department of Geriatric Medicine, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands*
[b]*Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands*
[c]*Alzheimer Centre Nijmegen, Nijmegen, The Netherlands*
[d]*Department of Epidemiology, Biostatistics and HTA, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands*
[e]*Department of Clinical Chemistry, Neurological Laboratory, VU University Medical Center, Amsterdam, The Netherlands*
[f]*Department of Neurology and Alzheimer Centre, VU University Medical Center, Amsterdam, The Netherlands*
[g]*Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands*
[h]*Department of Neurology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands*
[i]*Department of Laboratory Medicine, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands*

Abstract

Background: We aimed to develop a prediction model based on cerebrospinal fluid (CSF) biomarkers, that would yield a single estimate representing the probability that dementia in a memory clinic patient is due to Alzheimer's disease (AD).

Methods: All patients suspected of dementia in whom the CSF biomarkers had been analyzed were selected from a memory clinic database. Clinical diagnosis was AD (n = 272) or non-AD (n = 289). The prediction model was developed with logistic regression analysis and included CSF amyloid β_{42}, CSF phosphorylated tau_{181}, and sex. Validation was performed on an independent data set from another memory clinic, containing 334 AD and 157 non-AD patients.

Results: The prediction model estimated the probability that AD is present as follows: $p(AD) = 1/(1 + e^{-[-0.3315 + score]})$, where score is calculated from $-1.9486 \times \ln(\text{amyloid } \beta_{42}) + 2.7915 \times \ln(\text{phosphorylated tau}_{181}) + 0.9178 \times \text{sex}$ (male = 0, female = 1). When applied to the validation data set, the discriminative ability of the model was very good (area under the receiver operating characteristic curve: 0.85). The agreement between the probability of AD predicted by the model and the observed frequency of AD diagnoses was very good after taking into account the difference in AD prevalence between the two memory clinics.

Conclusions: We developed a prediction model that can accurately predict the probability of AD in a memory clinic population suspected of dementia based on CSF amyloid β_{42}, CSF phosphorylated tau_{181}, and sex.

Keywords: Prediction model; Dementia; CSF biomarkers; Differential diagnosis; Validation

1. Background

Cerebrospinal fluid (CSF) biomarkers amyloid β_{42} ($A\beta_{42}$), phosphorylated tau_{181} (p-tau), and total tau (t-tau) are increasingly used in the clinical diagnosis of Alzheimer's disease (AD). A guideline for their interpretation is, however, still lacking. The National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association clinical criteria for dementia due to AD were developed before the biomarker era [1], and although the newly proposed research criteria for AD are more recent, they do not provide a readymade guideline

*Corresponding author. Tel.: +31-24-361-6772; Fax: +31-24-361-7408.
E-mail address: PetraESpies@gmail.com (P.E.S.)

1552-5260/$ - see front matter © 2013 The Alzheimer's Association. All rights reserved.
doi:10.1016/j.jalz.2012.01.010

for implementation [2,3]. A combination of decreased $A\beta_{42}$ and increased t-tau and p-tau concentrations is considered a supportive feature of AD. However, minimal deviations from the cutoff values are intuitively less supportive than concentrations that deviate strongly from these cutoff values. Clinicians who frequently use CSF biomarkers may have developed such an intuitive interpretation of the results, but for clinicians with less experience with CSF biomarkers, it can be more difficult to draw equivocal conclusions from the results provided by the laboratory.

CSF biomarkers provide excellent differentiation between healthy control subjects and AD patients [4,5], but this is not a real challenge in clinical practice. The question that the clinician is mostly faced with is how to differentiate AD from other dementia disorders or from a psychiatric disorder that mimics dementia. Although the reported specificities of CSF biomarkers for other dementias are lower than the specificity for healthy control subjects [6–9], they are often used for differential diagnostic purposes in memory clinics [10]. Previously reported models based on CSF biomarkers do not, however, provide a formula that can be applied in clinical practice for the differentiation of AD from non-AD dementia [11,12].

We aimed to facilitate the use of CSF biomarkers in clinical practice by developing a prediction model for AD that yields a single estimate that would represent the probability that AD is the cause of the dementia in a memory clinic patient. The prediction model was validated on an independent database.

2. Methods

2.1. Development of the prediction model

All patients with a presumptive diagnosis of dementia who had visited the Radboud University Nijmegen Medical Centre (RUNMC) memory clinic between 1993 and 2008 and for whom CSF was available were selected. Lumbar punctures had been performed for research purposes or to exclude diseases such as neurolues. The CSF biomarkers were determined later and their interpretation was not part of the diagnostic process. The clinical diagnosis was made by a multidisciplinary panel that used all available clinical data, including neuropsychological assessment and magnetic resonance imaging. Diagnosis was AD (n = 272) or non-AD (n = 289). AD patients fulfilled the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association criteria and *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition, text revision, criteria for dementia [1]. Non-AD patients were defined as patients in whom AD was part of the differential diagnosis, but who were eventually diagnosed with another diagnosis, such as another type of dementia or a psychiatric disorder causing the cognitive disorder. We did not include patients diagnosed with subjective cognitive complaints or mild cognitive impairment (MCI), as only some of these patients will progress to develop dementia due to AD, and hence they cannot be classified as AD or non-AD. Levels of $A\beta_{42}$, t-tau, and p-tau were measured at the RUNMC laboratory using enzyme-linked immunosorbent assays (Innogenetics, Ghent, Belgium). Patient characteristics are shown in Table 1.

2.2. Statistical analysis

The CSF biomarkers were log-transformed to reduce strong influence of outliers. The prediction model was developed with logistic regression analysis, using group membership (AD or non-AD) as dependent variable. Sex and age were included in the analysis as readily available patient characteristics. The three log-transformed CSF biomarkers, sex, and age were entered with backward stepwise selection. Significant predictors were retained in the model. A significance level of 0.10 was used. The ability of the model to discriminate between AD and non-AD was quantified by the area under the receiver operating characteristic curve (AUC). Calibration of the model, that is, the agreement between the probability of AD predicted by the model and the observed frequency of AD diagnoses, was assessed by plotting the mean predicted probability against the observed frequency of AD for each probability class (0–0.1, 0.1–0.2,…, 0.9–1.0).

2.3. Validation of the prediction model

The data set used for validation of the prediction model consisted of all patients with a presumptive diagnosis of dementia who had visited the VU University Medical Center's (VUmc) Alzheimer Center between 2005 and 2008. Clinical diagnosis was made by a multidisciplinary panel, and was AD (n = 334) or non-AD (n = 157). To allow research on CSF biomarkers, diagnoses were made and recorded in the database without knowledge of the CSF biomarkers. All other available data, including neuropsychological assessment and magnetic resonance imaging, were used for diagnosis. Levels of $A\beta_{42}$, t-tau, and p-tau of these patients were measured at the VUmc laboratory using enzyme-linked immunosorbent assays (Innogenetics, Ghent, Belgium) (Table 1).

Before applying the prediction model on this validation data set, a shrinkage factor was applied, which is a methodological procedure in model validation that accounts for the expected overfitting of the developed model [13]. The amount of shrinkage was based on the *P* value of the model and the number of variables analyzed. Another adjustment of the prediction model was based on the need to correct for the difference in AD prevalence between the centers. Prevalence was defined as the number of AD patients among the total number of patients in the data set. If the prevalence in a new population is higher than the prevalence in the population that was used to develop the model, the predicted probabilities will underestimate the actual probabilities. A lower prevalence in the new population will lead to overestimation of the actual probabilities. Ideally, the mean

Table 1
Characteristics of patients in the development and validation data sets

	Development data set		Validation data set	
Parameter	AD (n = 272)	Non-AD (n = 289)	AD (n = 334)	Non-AD (n = 157)
M/F (n)	111/161	180/109	162/172	97/60
Age (year)	69.4 ± 9.2	68.9 ± 11.1	67.3 ± 9.2	66.1 ± 9.4
MMSE	19 ± 5	20 ± 7	21 ± 5	23 ± 5
CSF $A\beta_{42}$ (pg/mL)	410 (128–1162)	553 (118–1362)	455 (213–1335)	739 (201–1419)
CSF p-tau (pg/mL)	95 (23–370)	53 (15–378)	82 (16–224)	46 (17–133)
CSF t-tau (pg/mL)	558 (75–2497)	264 (74–13,440)	619 (91–3150)	347 (83–1285)
Diagnosis (n)				
VaD		69 (24%)		20 (13%)
FTLD		76 (26%)		79 (50%)
DLB		50 (17%)		32 (20%)
Other dementia*		42 (15%)		26 (17%)
Psychiatric		12 (4%)		0
Other†		40 (14%)		0

Abbreviations: AD, dementia due to Alzheimer's disease; non-AD, non-Alzheimer's disease; $A\beta_{42}$, amyloid β_{42}; t-tau, total tau; p-tau, phosphorylated tau_{181}; VaD, vascular dementia; FTLD, frontotemporal lobe degeneration; DLB, dementia with Lewy bodies; M, male; F, female; MMSE, Mini-Mental State Examination; CSF, cerebrospinal fluid.

NOTE. Age and MMSE are expressed as mean ± SD, CSF biomarkers are expressed as median (range).

*Other dementia comprised progressive supranuclear palsy, corticobasal degeneration, Parkinson dementia, dementia due to alcohol abuse.

†Other diagnoses comprised, among others, delirium, normal pressure hydrocephalus, vitamin B12 deficiency.

probability predicted by the model should be equal to the observed prevalence. The model is adjusted accordingly by correcting the intercept of the logistic regression formula [14]. We examined two types of adjustment. One is a straightforward adjustment based on the discrepancy between the prevalence of AD in the validation and in the development data sets: correction factor = ln([prevalence in validation data set/{1 − prevalence in validation data set}]/[prevalence in development data set/{1 − prevalence in development data set}]). This correction factor can be used under the condition that the difference in prevalence between the data sets is not caused by any of the variables included in the model. It is added to the intercept of the original model [15,16]. This resulted in adjusted model A. The other type of adjustment requires information of individual cases in the new population on the variables that are included in the model. The original prediction model without the intercept is used to calculate a score per individual by multiplying the coefficient for each variable in the model by the individual's value for that variable. A logistic regression model is then fitted with these individual scores as linear predictor and their diagnosis (AD or non-AD) as dependent variable. The slope of this model is fixed at 1. The intercept that is obtained is the correction factor, which is added to the intercept of the original model [17,18]. This resulted in adjusted model B.

To be able to take into account interlaboratory variation, both laboratories (that participate actively in the Alzheimer's Association CSF Quality Control program [19,20]) analyzed $A\beta_{42}$, t-tau, and p-tau in 32 randomly chosen CSF samples from the RUNMC biobank. For this purpose, the CSF biomarkers were measured in two aliquots of the same CSF sample that were both kept frozen at –80°C and were thawed only once for analysis in either center. With linear regression, a correction factor was calculated (Figure S1) that could be applied to the CSF values in the validation data set.

Discrimination of adjusted model A and B was quantified by the AUC. The AUC is based on the ranking of the predicted probabilities of the subjects, and is not influenced by correction factors. Therefore, it is the same for both adjusted models. Calibration was assessed by plotting the predicted probabilities against the observed frequency of AD in the validation data set. Statistical analyses were carried out using SAS version 9.2 (SAS Institute Inc., Cary, North-Carolina, USA).

3. Results

3.1. Development of the prediction model

The prevalence of AD in the development data set was 48%. Ln($A\beta_{42}$), ln(p-tau), and female sex were significant predictors of a diagnosis of AD, all $P < .0001$ (Table 2). Age had no influence. Because of a strong correlation with p-tau, the addition of t-tau was not significant. The prediction model estimates the probability that AD is present by the formula shown in Table 2. The AUC of this model was 0.87, indicating very good discrimination between AD and non-AD. For comparison, we calculated the AUCs of the individual biomarkers, which were all lower: the AUC of $A\beta_{42}$ was 0.70, of p-tau was 0.82, and of t-tau was 0.79.

The agreement between the probability of AD predicted by the model and the observed frequency of AD diagnoses was also good (Figure 1); for example, of the subjects for whom the predicted probability of AD was ≤0.1 (n = 86, mean predicted probability of 0.056, i.e., AD expected in 5.6%), 3.5% had a diagnosis of AD. Likewise, of the subjects

Table 2
Intercept and coefficients of the original prediction model and of the adjusted models A and B

Variable	Original model	A: shrinkage + simple prevalence correction	B: shrinkage + prevalence correction based on individual data
Intercept	−0.3315	0.4840	1.1528
ln($A\beta_{42}$)	−1.9486	−1.9122	−1.9122
ln(p-tau)	2.7915	2.7393	2.7393
Sex*	0.9178	0.9006	0.9006

NOTE. $p(AD) = 1/(1 + e^{-[intercept + score]})$; score = coefficient × ln($A\beta_{42}$) + coefficient × ln(p-tau) + coefficient × sex.
*Male = 0, female = 1.

for whom the predicted probability of AD was >0.9 (n = 67, mean predicted probability of 0.943, i.e., AD expected in 94.3%), 91.0% had a diagnosis of AD.

3.2. Validation of the prediction model

A shrinkage factor of 0.9813 was applied by multiplying the coefficients in the prediction model by this factor (Table 2). The prevalence of AD in the validation data set was 68%, whereas it was 48% in the development data set. Two types of correction for this difference in prevalence were examined, resulting in two adjusted models: model A with simple prevalence correction and model B with prevalence correction based on individual data.

When using the simple correction factor based only on the prevalence of AD in the validation data set, the intercept was adjusted to 0.4840 (Table 2). This resulted in adjusted model A: $p(AD) = 1/(1 + e^{-[0.4840 + score]})$, where score is calculated from $-1.9122 \times \ln(A\beta_{42}) + 2.7393 \times \ln(p\text{-tau}) + 0.9006 \times sex$. The agreement between the probabilities predicted by adjusted model A and the observed frequency of AD in the validation data set was very good in the low and high prediction classes (Figure 2A and Table 3). For example, of the subjects for whom the predicted probability of AD was ≤0.1 (n = 52, mean predicted probability of 0.060, i.e., AD expected in 6.0%), 9.6% had a diagnosis of AD, and of the subjects for whom the predicted probability of AD was >0.9 (n = 99, mean predicted probability of 0.940, i.e., AD expected in 94.0%), 91.9% had a diagnosis of AD. In the middle classes, the observed frequency of AD was higher than predicted; for example, of the subjects for whom the predicted probability of AD was 0.4 to 0.5 (n = 22, mean predicted probability of 0.440, i.e., AD expected in 44.0%), 68.1% had a diagnosis of AD. Mean predicted probability, which translates to the predicted prevalence, was 0.58.

For the second, more extensive correction resulting in adjusted model B, the scores of the individual cases in the validation data set, as obtained with the original prediction model, were fitted in a logistic regression model with a slope fixed at 1. The intercept of this logistic regression model was 1.4843, and this was added to the intercept of our original prediction model, resulting in an adjusted intercept of 1.1528 (Table 2). Adjusted model B was: $p(AD) = 1/(1 + e^{-[1.1528 + score]})$, score = $-1.9122 \times \ln(A\beta_{42}) + 2.7393 \times \ln(p\text{-tau}) + 0.9006 \times sex$ (Table 2).

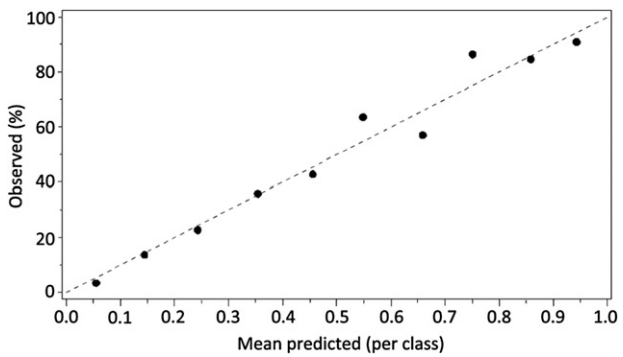

Fig. 1. Calibration plot of predicted probability against observed frequency of Alzheimer's disease (AD) in the development data set using the original prediction model. Horizontal axis: probability of AD as estimated by the prediction model (range: 0–1). Patients are divided into classes of probabilities (0.0–0.1, 0.1–0.2, and so on). Each dot represents the mean probability of patients within one class. Vertical axis: actual percentage of AD patients. Each dot represents the percentage of AD patients within one probability class. Dotted line represents line of perfect agreement between predicted probability and observed percentage of diagnoses.

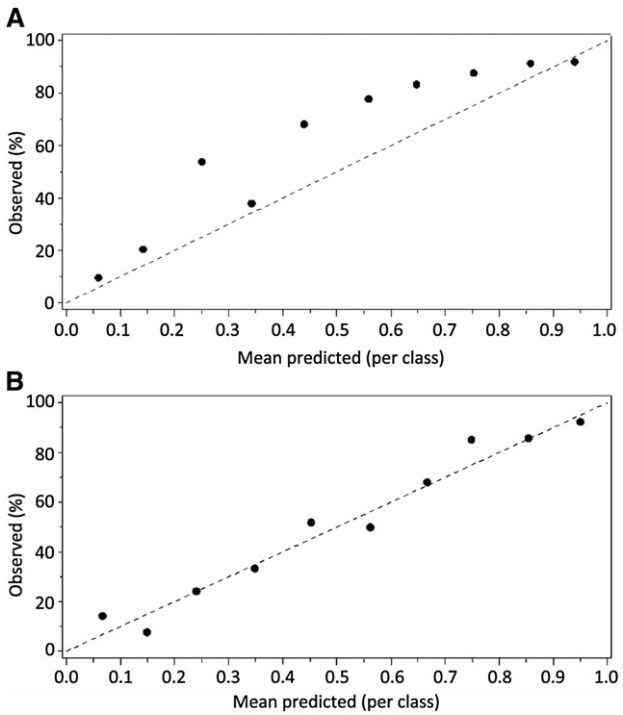

Fig. 2. Calibration plot of predicted probability against observed frequency of AD in the validation data set. (A) adjusted model A with simple correction for difference in prevalence, and (B) adjusted model B with extensive correction for difference in prevalence. Horizontal axis: probability of AD as estimated by the prediction model (range: 0–1). Patients are divided into classes of probabilities (0.0–0.1, 0.1–0.2, and so on). Each dot represents the mean probability of patients within one class. Vertical axis: actual percentage of AD patients. Each dot represents the percentage of AD patients within one probability class. Dotted line represents line of perfect agreement between predicted probability and observed percentage of diagnoses.

Table 3
Mean predicted probability of AD and observed number of AD and non-AD cases per class in the validation data set, using model A or B

	Model A: simple correction for difference in prevalence				Model B: extensive correction for difference in prevalence			
Class*	Mean predicted probability of AD	n per class	Observed AD cases, n (%)	Observed non-AD cases, n (%)	Mean predicted probability of AD	n per class	Observed AD cases, n (%)	Observed non-AD cases, n (%)
0–0.1	5.96%	52	5 (9.62%)	47 (90.38%)	6.72%	21	3 (14.29%)	18 (85.71%)
0.1–0.2	14.22%	44	9 (20.45%)	35 (79.55%)	14.92%	39	3 (7.69%)	36 (92.31%)
0.2–0.3	25.08%	26	14 (53.85%)	12 (46.15%)	24.06%	29	7 (24.14%)	22 (75.86%)
0.3–0.4	34.25%	29	11 (37.93%)	18 (62.07%)	34.84%	21	7 (33.33%)	14 (66.67%)
0.4–0.5	43.98%	22	15 (68.18%)	7 (31.82%)	45.26%	27	14 (51.85%)	13 (48.15%)
0.5–0.6	55.92%	45	35 (77.78%)	10 (22.22%)	56.24%	28	14 (50.00%)	14 (50.00%)
0.6–0.7	64.77%	36	30 (83.33%)	6 (16.67%)	66.74%	25	17 (68.00%)	8 (32.00%)
0.7–0.8	75.33%	57	50 (87.72%)	7 (12.28%)	74.92%	54	46 (85.19%)	8 (14.81%)
0.8–0.9	85.82%	81	74 (91.36%)	7 (8.64%)	85.48%	77	66 (85.71)	11 (4.29)
0.9–1.0	93.97%	99	91 (91.92%)	8 (8.08%)	95.02%	170	157 (92.35)	13 (7.65)

*Class refers to a class of probability; the classes are the same as the ones that are used on the horizontal axis of the calibration plots (Figs. 1 and 2). To facilitate interpretation, mean predicted probability of AD is expressed in %.

The agreement between the probabilities of AD predicted by adjusted model B and the observed frequency of AD was very good in all classes (Figure 2B and Table 3). For example, of the subjects for whom the predicted probability of AD was $\leq$0.1 (n = 21, mean predicted probability of 0.067, i.e., AD expected in 6.7%), 14.3% had a diagnosis of AD; of the subjects for whom the predicted probability of AD was 0.4 to 0.5 (n = 27, mean predicted probability of 0.453, i.e., AD expected in 45.3%), 51.9% had a diagnosis of AD; and of the subjects for whom the predicted probability of AD was >0.9 (n = 170, mean predicted probability of 0.950, i.e., AD expected in 95%), 92.4% had a diagnosis of AD. Mean predicted probability was 0.68, equal to the prevalence of AD in this data set.

When correction for interlaboratory variation was applied to the CSF biomarker values in the validation data set, the performances of both adjusted models (A and B) remained largely the same; mean predicted probability was 0.57 and 0.68, respectively (Figure S2).

The AUC of adjusted model A and B was 0.85, indicating that the ability of the model to discriminate between AD and non-AD was very good.

4. Discussion

We developed a prediction model that yields the probability that a patient from the memory clinic population suspected of dementia has AD. The probability is calculated from the patient's sex combined with the levels of CSF $A\beta_{42}$ and CSF p-tau. The model was built using a data set containing >500 patients and was validated on an independent and comparably large data set. An important feature of this model is that it takes into account the differential diagnostic dilemma that is often the reality of clinical practice. When a previously described model based on CSF $A\beta_{42}$ and tau aimed to discriminate AD from non-AD dementia, almost half of the non-AD patients were incorrectly classified as AD [12]. Another model based on $A\beta_{42}$ and p-tau could very well discriminate between AD and control subjects and correctly identified MCI patients who would progress to AD, but did not offer a formula that can be used by others [11]. We have provided the details of our model to allow others to use it and adjust it, and in doing so, to retain the information of >1000 patients. The complex formula of the prediction model can be programmed in an application that will only require the clinician to enter the three parameters: sex, $A\beta_{42}$, and p-tau.

The variables that were selected in our model are in line with previous research. Epidemiological studies indicate that AD is more frequent in women [21,22]. CSF $A\beta_{42}$ and p-tau were found to identify AD in a group of control subjects and MCI patients [11]. In addition, it has been shown that this combination of biomarkers may help to differentiate AD from vascular dementia [23], dementia with Lewy bodies [24], and frontotemporal lobe degeneration [25]. Total-tau, on the other hand, is considered a general marker of neurodegeneration and the least specific of the three CSF biomarkers [7].

An important step in creating a prediction model is its validation. We validated our model using an independent data set and showed that our model accurately predicted the probability of AD. This proof of principle may encourage other centers to develop their own prediction model based on CSF biomarkers, if they have a data set available. Alternatively, our model can be adjusted for use in other centers. This is an important advantage of our approach, as existing knowledge can be incorporated into the model without loss of information [14,18]. We examined two different types of adjustment for a difference in prevalence. The minimum requirement for model A is to know the prevalence of AD in that center [14]. Calibration can further be fine-tuned with individual data (model B), but in that case, a complete data set is required, which might be more difficult to acquire. However, the prediction model can be used regardless of which of these two adjustments is applied, as with both adjustments, the discrimination between AD and

non-AD was very good (AUC = 0.85) and agreement between predicted probability and observed frequency of AD was good to excellent, especially when the predicted probabilities for AD were low or high. The low and high probabilities are also clinically the most meaningful. For example, even with perfect agreement between predicted probability of AD and observed frequency of AD, a calculated probability of AD of 50% will be of no aid in the differential diagnosis between AD or non-AD dementia. In contrast, calculated probabilities of, for instance, <20% or >80% will truly inform decision making and will provide an accurate prediction of the probability of AD in our models.

Remarkably, correction for interlaboratory variation did not influence our results. Differences between the laboratories may have been diminished by inclusion of large groups of patients and because the laboratories used the same enzyme-linked immunosorbent assays. Interlaboratory variation was substantial in a multicenter analysis, but it decreased with the use of the same assay [26]. Apparently, the quality of the CSF analyses performed by the two laboratories in this study is comparable. However, substantial interlaboratory variation may still exist when the results of our laboratories are compared with other laboratories, especially if different assays are used. To take this interlaboratory variation into account, a correction factor can be calculated and applied to the values of CSF $A\beta_{42}$ and p-tau. Such a correction factor may be derived from the ongoing international Quality Control program by the Alzheimer's Association [19,20].

Our model provides a continuous outcome. We decided not to calculate an optimal cutoff point that yields the best ratio between true positives and false negatives. Defining the "best" ratio is arbitrary, and although a dichotomous outcome would appear to make clinical diagnosing easier, it would be at the cost of loss of information. As an example, let us consider a 70% or higher probability of AD as supporting a diagnosis of AD and a probability of <70% as supportive of a non-AD diagnosis. Sensitivity would be 80.5%, specificity 79.6%, positive predictive value 89.4%, negative predictive value 65.8%, and 80.2% of cases would be correctly classified (data not shown). Although these percentages appear promising, such a dichotomous cutoff value means that when a patient has a 69.9% probability of AD and another has a 70.1% probability of AD, the former patient would be classified as non-AD and the latter as AD, whereas, objectively, their chance of having AD would be equal.

How should the probability of AD as derived from our model be interpreted and implemented? For a clinically meaningful interpretation, this estimate should not be interpreted on its own, but should be instead weighed against the previous probability of AD, which the clinician derives from interpretation of a combination of history taking, physical examination, neuropsychological evaluation, and neuroimaging. Theoretically, all these factors could be included in a prediction model. It is, however, difficult to specify which of the items of, for example, history taking or neuropsychological testing should be included, especially because neither has reached any form of standardization across centers.

It is important to emphasize that the prediction model is applicable in the memory clinic population suspected of dementia. The model is not yet validated for MCI patients or patients for whom AD is not part of the differential diagnosis. However, both in multicenter studies and in meta-analyses, it has repeatedly been shown that combinations of CSF biomarkers can identify incipient AD in MCI patients [27–30]. We, therefore, have reason to speculate that our prediction model will be of value in MCI patients as well, to discriminate MCI patients with AD pathology from those without AD pathology; however, to confirm this assumption, the prediction model should first be validated in a group of MCI patients.

A limitation of our model is one that is common in dementia research: we used the clinical diagnosis as a reference standard, and some misclassifications are bound to have occurred. This limitation is not easily overcome. Not only is it difficult to obtain neuropathological information of this many patients, but even a neuropathological diagnosis is not as unequivocal as it is sometimes presented, especially when the clinical picture is unclear [31]. More recent developments such as amyloid imaging by 11C-Pittsburgh Compound B cannot serve as a reference standard, as positive scans are found in 33% of healthy control subjects [32] and it is as yet unknown whether these represent very early AD or false positives. Diagnoses in our data sets were made by multidisciplinary panels highly experienced in diagnosing dementia, and although some misclassifications are not ruled out, our model reflects current clinical practice in the best possible way.

To conclude, we developed a prediction model that can serve as a first step in the implementation of CSF biomarker analysis in diagnosing dementia due to AD in the memory clinic population suspected of dementia.

Acknowledgments

The authors thank the technicians of the Department of Laboratory Medicine of the RUNMC and the Neurological Laboratory of the Department of Clinical Chemistry VU University Medical Center for CSF analysis, and Wim Lemmens of the Department of Epidemiology, Biostatistics and Health Technology Assessment of the RUNMC for statistical analysis.

This research was performed within the framework of the Center for Translational Molecular Medicine, www.ctmm.nl, project LeARN grant 02N-101—a Dutch public–private partnership on the development and valorization of innovative diagnostic techniques for early detection of AD. LeARN is a consortium of Philips, Merck, Virtual Proteins, Bio Affinity Company, Cyclotron BV, Leiden University Medical Center, VUmc, Maastricht University Medical Center, and RUNMC.

J.A.H.R.C. is supported by grants from Netherlands Heart Foundation and NWO (Netherlands Organisation for

Scientific Research); M.A.B. has received payment for lectures from Ferring Netherlands and Abbott Diagnostics; C.E.T. is supported by grants from European Union FP7 and Dutch MS Society, and received an Ectrims conference grant as speaker and chair; P.S. is a consultant for Pfizer Inc, Genentech, Danone Research, Lundbeck, GE Healthcare, Roche AG, Novartis AG, and Jansen AI, and has received payment for lectures from Lundbeck and Danone Research; M.G.M.O.R. is a consultant for Janssen, Numico, and Schering-Plough; M.M.V. is supported by grants from Stichting Internationaal Parkinson Fonds, Van Alkemade–Keuls Fonds, American Alzheimer's Association, Alzheimer Drug Discovery Foundation, Aromatic Amino Acid Decarboxylase Deficiency Research Trust, Hersenstichting Nederland.

References

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

[2] Dubois B, Feldman HH, Jacova C, DeKosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol 2007;6:734–46.

[3] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr., Kawas CH, et al. The diagnosis of dementia 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;7:263–9.

[4] Blennow K. Cerebrospinal fluid protein biomarkers for Alzheimer's disease. NeuroRx 2004;1:213–25.

[5] de Jong D, Kremer BP, Olde Rikkert MG, Verbeek MM. Current state and future directions of neurochemical biomarkers for Alzheimer's disease. Clin Chem Lab Med 2007;45:1421–34.

[6] Vanmechelen E, Vanderstichele H, Hulstaert F, Andreasen N, Minthon L, Winblad B, Davidsson P, Blennow K. Cerebrospinal fluid tau and beta-amyloid(1–42) in dementia disorders. Mech Ageing Dev 2001;122:2005–11.

[7] Verbeek MM, de Jong D, Kremer HP. Brain-specific proteins in cerebrospinal fluid for the diagnosis of neurodegenerative diseases. Ann Clin Biochem 2003;40:25–40.

[8] Spies PE, Slats D, Sjogren JM, Kremer BP, Verhey FR, Rikkert MG, Verbeek MM. The cerebrospinal fluid amyloid beta42/40 ratio in the differentiation of Alzheimer's disease from non-Alzheimer's dementia. Curr Alzheimer Res 2010;7:470–6.

[9] Mulder C, Verwey NA, van der Flier WM, Bouwman FH, Kok A, van Elk EJ, Scheltens P, Blankenstein MA. Amyloid-beta(1–42), total tau, and phosphorylated tau as cerebrospinal fluid biomarkers for the diagnosis of Alzheimer disease. Clin Chem 2010;56:248–53.

[10] Spies PE, Slats D, Ramakers I, Verhey FR, Olde Rikkert MG. Experiences with cerebrospinal fluid analysis in Dutch memory clinics. Eur J Neurol 2011;18:1014–6.

[11] De Meyer G, Shapiro F, Vanderstichele H, Vanmechelen E, Engelborghs S, De Deyn PP, et al. Diagnosis-independent Alzheimer disease biomarker signature in cognitively normal elderly people. Arch Neurol 2010;67:949–56.

[12] Hulstaert F, Blennow K, Ivanoiu A, Schoonderwaldt HC, Riemenschneider M, De Deyn PP, et al. Improved discrimination of AD patients using beta-amyloid(1–42) and tau levels in CSF. Neurology 1999;52:1555–62.

[13] Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. Stat Med 1990;9:1303–25.

[14] Janssen KJ, Vergouwe Y, Kalkman CJ, Grobbee DE, Moons KG. A simple method to adjust clinical prediction models to local circumstances. Can J Anaesth 2009;56:194–201.

[15] Smits M, Dippel DW, Steyerberg EW, de Haan GG, Dekker HM, Vos PE, et al. Predicting intracranial traumatic findings on computed tomography in patients with minor head injury: the CHIP prediction rule. Ann Intern Med 2007;146:397–405.

[16] Steyerberg EW. Clinical prediction models. 1st ed. Berlin, Germany: Springer; 2009.

[17] Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. Stat Med 2004;23:2567–86.

[18] Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. J Clin Epidemiol 2008;61:76–86.

[19] Alzheimer's Association Cerebrospinal Fluid Quality Control Program. Available at: http://www.neurophys.gu.se/sektioner/psykiatri_och_neurokemi/neurokem/thealzassqcprogram. Accessed September 1, 2011.

[20] Mattsson N, Andreasson U, Persson S, Arai H, Batish SD, Bernardini S, et al. The Alzheimer's Association external quality control program for cerebrospinal fluid biomarkers. Alzheimers Dement 2011;7:386–95.

[21] Vina J, Lloret A. Why women have more Alzheimer's disease than men: gender and mitochondrial toxicity of amyloid-beta peptide. J Alzheimers Dis 2010;20(Suppl 2):S527–33.

[22] Plassman BL, Langa KM, Fisher GG, Heeringa SG, Weir DR, Ofstedal MB, et al. Prevalence of dementia in the United States: the aging, demographics, and memory study. Neuroepidemiology 2007;29:125–32.

[23] de Jong D, Jansen RW, Kremer BP, Verbeek MM. Cerebrospinal fluid amyloid beta42/phosphorylated tau ratio discriminates between Alzheimer's disease and vascular dementia. J Gerontol A Biol Sci Med Sci 2006;61:755–8.

[24] Mukaetova-Ladinska EB, Monteith R, Perry EK. Cerebrospinal fluid biomarkers for dementia with lewy bodies. Int J Alzheimers Dis 2010;2010:536538.

[25] Schoonenboom NS, Pijnenburg YA, Mulder C, Rosso SM, van Elk EJ, Van Kamp GJ, Van Swieten JC, Scheltens P. Amyloid beta(1–42) and phosphorylated tau in CSF as markers for early-onset Alzheimer disease. Neurology 2004;62:1580–4.

[26] Verwey NA, van der Flier WM, Blennow K, Clark C, Sokolow S, De Deyn PP, et al. A worldwide multicentre comparison of assays for cerebrospinal fluid biomarkers in Alzheimer's disease. Ann Clin Biochem 2009;46:235–40.

[27] Mattsson N, Zetterberg H, Hansson O, Andreasen N, Parnetti L, Jonsson M, et al. CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. JAMA 2009;302:385–93.

[28] Schmand B, Huizenga HM, van Gool WA. Meta-analysis of CSF and MRI biomarkers for detecting preclinical Alzheimer's disease. Psychol Med 2010;40:135–45.

[29] van Rossum I, Vos S, Handels R, Visser PJ. Biomarkers as predictors for conversion from mild cognitive impairment to Alzheimer-type dementia: implications for trial design. J Alzheimers Dis 2010;20:881–91.

[30] Visser PJ, Verhey F, Knol DL, Scheltens P, Wahlund LO, Freund-Levi Y, et al. Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive impairment or mild cognitive impairment in the DESCRIPA study: a prospective cohort study. Lancet Neurol 2009;8:619–27.

[31] Scheltens P, Rockwood K. How golden is the gold standard of neuropathology in dementia? Alzheimers Dement 2011;7:486–9.

[32] Rowe CC, Ellis KA, Rimajova M, Bourgeat P, Pike KE, Jones G, et al. Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. Neurobiol Aging 2010;31:1275–83.