

Correcting for the Absence of a Gold Standard Improves Diagnostic Accuracy of Biomarkers in Alzheimer's Disease

Peer-reviewed author version

Coart, Els; GARCIA BARRADO, Leandro; Duits, Flora H.; Scheltens, Philip; van der Flier, Wiesje M.; Teunissen, Charlotte E.; van der Vies, Saskia M. & BURZYKOWSKI, Tomasz (2015) Correcting for the Absence of a Gold Standard Improves Diagnostic Accuracy of Biomarkers in Alzheimer's Disease. In: JOURNAL OF ALZHEIMERS DISEASE, 46 (4), p. 889-899.

DOI: 10.3233/JAD-142886

Handle: <http://hdl.handle.net/1942/19148>

# Correcting for the Absence of a Gold Standard Improves Diagnostic Accuracy of Biomarkers in Alzheimer's Disease

Els Coart<sup>a,\*</sup>, Leandro García Barrado<sup>b</sup>, Flora H. Duits<sup>c</sup>, Philip Scheltens<sup>c</sup>, Wiesje M. van der Flier<sup>c,d</sup>, Charlotte E. Teunissen<sup>e</sup>, Saskia M. van der Vies<sup>f</sup>, Tomasz Burzykowski<sup>a,b</sup> and for the Alzheimer's Disease Neuroimaging Initiative<sup>1</sup>

<sup>a</sup>*International Drug Development Institute (IDDI), Louvain-la-Neuve, Belgium*

<sup>b</sup>*Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BioStat), Hasselt University, Diepenbeek, Belgium*

<sup>c</sup>*Alzheimer Center & Department of Neurology, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, The Netherlands*

<sup>d</sup>*Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands*

<sup>e</sup>*Neurochemistry Laboratory and Biobank, Department of Clinical Chemistry, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, The Netherlands*

<sup>f</sup>*Department of Pathology, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, The Netherlands*

Handling Associate Editor: Henrik Zetterberg

Accepted 26 March 2015

## Abstract.

**Background:** Studies investigating the diagnostic accuracy of biomarkers for Alzheimer's disease (AD) are typically performed using the clinical diagnosis or amyloid- $\beta$  positron emission tomography as the reference test. However, neither can be considered a gold standard or a perfect reference test for AD. Not accounting for errors in the reference test is known to cause bias in the diagnostic accuracy of biomarkers.

**Objective:** To determine the diagnostic accuracy of AD biomarkers while taking the imperfectness of the reference test into account.

**Methods:** To determine the diagnostic accuracy of AD biomarkers and taking the imperfectness of the reference test into account, we have developed a Bayesian method. This method establishes the biomarkers' true value in predicting the AD-pathology status by combining the reference test and the biomarker data with available information on the reliability of the reference test. The new methodology was applied to two clinical datasets to establish the joint accuracy of three cerebrospinal fluid biomarkers (amyloid- $\beta_{1-42}$ , Total tau, and P-tau<sub>181</sub>) by including the clinical diagnosis as imperfect reference test into the analysis.

<sup>1</sup>Part of the data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://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: [https://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNIAcknowledgement\\_List.pdf](https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNIAcknowledgement_List.pdf)

\*Correspondence to: Els Coart, International Drug Development Institute (IDDI), Avenue Provinciale 30, 1340 Louvain-la-Neuve, Belgium. Tel.: +32 10 61 44 44; Fax: +32 10 61 88 88; E-mail: Elisabeth.Coart@iddi.com.

**Results:** The area under the receiver-operating-characteristics curve to discriminate between AD and controls, increases from 0.949 (with 95% credible interval [0.935,0.960]) to 0.990 ([0.985,0.995]) and from 0.870 ([0.817,0.912]) to 0.975 ([0.943,0.990]) for the cohorts, respectively.

**Conclusions:** Use of the Bayesian methodology enables an improved estimate of the exact diagnostic value of AD biomarkers and overcomes the lack of a gold standard for AD. Using the new method will increase the diagnostic confidence for early stages of AD.

Keywords: Alzheimer's disease, Bayesian method, biomarkers, diagnostic test, reference standard

## INTRODUCTION

Biomarkers for Alzheimer's disease (AD) that are linked to the pathological process are of paramount importance for early diagnosis of AD and selection of appropriate patients for clinical trials [1, 2]. Before biomarkers can be used clinically, their diagnostic accuracy needs to be thoroughly ascertained. To this end, a reference test against which the biomarker is verified needs to be selected.

The first choice would be the definite AD diagnosis, provided by postmortem neuropathological analysis. However, autopsy confirmation suffers from considerable between-laboratory differences [3], is by definition *post hoc*, and is only rarely available. In general, the accuracy of early AD biomarkers, or any diagnostic test for AD for that matter, is typically assessed using the clinical diagnosis as the reference test. The latter is imperfect because the clinical diagnosis suffers from classification errors (misdiagnosis) [4] and the onset of the pathogenic process as reflected in biomarker changes can precede the manifestation of clinical symptoms by at least a decade [5]. Hence, a clinical non-AD diagnosis does not exclude underlying AD-pathology and the clinical diagnosis of AD does not predict underlying pathology, as was recently shown in the phase III study with Bapineuzimab [6].

The imperfectness of the clinical diagnosis is usually ignored, resulting in a biased assessment of the diagnostic accuracy of biomarkers and suboptimal biomarker thresholds for clinical applications [7]. If the biomarker and reference test do not tend to misclassify the same patients, the diagnostic accuracy of the biomarker will be underestimated. When the biomarker and the reference test are dependent, the diagnostic accuracy of the biomarker can be either underestimated or overestimated, depending on the strength of the association [8, 9]. Recently, Toledo et al. [10] demonstrated that using the clinical diagnosis as a perfect reference leads to an underestimation of cerebrospinal fluid (CSF) AD biomarker sensitivity and specificity values and shifts the cut-offs compared to using the autopsy confirmed diagnosis as reference test.

Different statistical methods have been developed to correctly estimate diagnostic accuracy when an imperfect reference test is used. Reitsma et al. [7] systematically reviewed the different solutions and provided methodological guidelines depending on the medical test under evaluation and the availability and nature of the data.

To date, these methods have not systematically been applied to estimate the diagnostic accuracy of AD biomarkers. An interesting attempt was undertaken by De Meyer et al. [11], who proposed a method to evaluate the CSF AD biomarkers while completely ignoring the clinical diagnosis.

More recently, positron emission tomography (PET) amyloid imaging was used as reference test for evaluation of the diagnostic accuracy of (mainly CSF) AD biomarkers for brain amyloid- $\beta$  ( $A\beta$ ) deposition [12]. Although this correctly reduces the time-lag in expected onset of changes between biomarkers and reference test, amyloid PET imaging cannot (yet) be considered a gold standard or a perfect reference test for early AD. There is no true *in vivo* gold standard for amyloid burden and there is substantial overlap between the distribution of PET measurements for presumed AD and non-AD groups [13, 14]. In addition, as for all tests, PET analysis is not free from measurement errors, and standardization of different measurement procedures is still ongoing [14].

As an alternative to search for a surrogate gold standard, it has been suggested that the complexity of dementia diagnosis would be best served by integrating multiple sources of information [3]. A Bayesian framework integrates different data sources in a natural way and is most suited for this purpose.

Bayesian methods have become increasingly popular, notably in medical research [15]. A Bayesian approach can include prior information, accommodate adaptive clinical trials (e.g., interim analyses, change to sample size, or change to randomization scheme) and can be useful for analysis of a complex model when a frequentist analysis is difficult to implement or does not exist [16].

115 Recent breakthroughs in computational algorithms  
116 and computing speed have made it possible to carry  
117 out calculations of the often computationally intense  
118 Bayesian analysis. Also the fact that regulatory authori-  
119 ties embrace the use of Bayesian statistics has boosted  
120 its application in medical research. Already in 2003,  
121 the US Food and Drug Administration (FDA) approved  
122 a drug combination (pravastatin and aspirin) based  
123 on a Bayesian analysis [17]. Likewise, the Center for  
124 Devices and Radiological Health of the FDA, that is  
125 among others responsible for clearance of diagnostic  
126 test kits, issued a guideline for the use of Bayesian  
127 statistics and now routinely accepts applications based  
128 on Bayesian trials [18].

129 Bayesian statistics is currently a widespread  
130 approach in oncology. Many leading medical jour-  
131 nals have published original oncology studies using  
132 Bayesian analysis and prominent cancer centers  
133 have implemented several clinical trials, which were  
134 designed using Bayesian methods [19]. In pediatric  
135 science, care providers are accustomed with and often  
136 obliged to rely on evidence from adult studies; bor-  
137 rowing information from adult trials using a Bayesian  
138 approach is common practice [20]. Also in diagnostic  
139 medicine, Bayesian approaches are well-established  
140 and often help to validate diagnostics test with smaller-  
141 sized and shorter-duration pivotal trial [18, 21].

142 In this paper, we present a Bayesian framework  
143 which establishes the diagnostic accuracy of AD  
144 biomarkers by integrating different data sources, with-  
145 out the need for a gold standard or perfect reference  
146 test. We applied the new Bayesian analysis method  
147 to establish the performance of the three CSF AD  
148 biomarkers,  $A\beta_{1-42}$ , Total tau, and P-tau<sub>181p</sub> present  
149 in two datasets, with the clinical diagnosis considered  
150 as an imperfect reference test. We hypothesized that  
151 the diagnostic performance of the CSF AD biomark-  
152 ers would be higher when analyzed with the Bayesian  
153 analysis method that accounts for the imperfectness of  
154 the clinical diagnosis.

## 155 MATERIALS AND METHODS

### 156 *Data sets*

157 We used two independent cohorts. The VUmc (VU  
158 University Medical Center) data set that consists of  
159 patients from the memory-clinic-based Amsterdam  
160 Dementia Cohort who received a diagnosis of either  
161 subjective memory complaints (SMC) or probable  
162 AD. Baseline CSF was collected between October  
163 1999 and November 2011. All patients underwent

164 standard dementia screening at baseline, including  
165 physical and neurological examination, EEG, MRI,  
166 and laboratory tests. Cognitive screening included a  
167 Mini-Mental State Examination (MMSE) and a com-  
168 prehensive neuropsychological test battery. Diagnoses  
169 were made by consensus in a multidisciplinary team  
170 without knowledge of CSF results. The label of SMC  
171 was given when results of all clinical examinations  
172 were normal, and there was no psychiatric diagnosis.  
173 Patients with subjective complaints were considered  
174 as controls, but were only included when the diagno-  
175 sis was confirmed at follow-up visits. This resulted  
176 in 251 SMC subjects. Probable AD ( $n=631$ ) was  
177 diagnosed according to the criteria of the National  
178 Institute of Neurological and Communicative Dis-  
179 orders and Stroke-Alzheimer's Disease and Related  
180 Disorders association (NINCDS-ADRDA), and all  
181 patients met the core clinical NIA-AA criteria [22].  
182 More details about this cohort have been provided  
183 elsewhere [23]. All subjects gave written informed  
184 consent for the use of their clinical data for research  
185 purposes. The current study was approved by the local  
186 ethical review board. CSF levels of  $A\beta_{1-42}$ , Total tau,  
187 and P-tau<sub>181p</sub> were determined using commercially  
188 available single-parameter ELISA kits (respectively,  
189 INNOTEST<sup>®</sup> AMYLOID(1-42), INNOTEST<sup>®</sup> hTAU  
190 Ag, INNOTEST<sup>®</sup> PHOSPHOTAU(181P)) and were  
191 not used for diagnosis.

192 The second data set consisted of Alzheimer's Dis-  
193 ease Neuroimaging Initiative (ADNI)-I patients. ADNI  
194 was launched in 2003 by the National Institute on  
195 Aging (NIA), the National Institute of Biomed-  
196 ical Imaging and Bioengineering (NIBIB), the FDA,  
197 private pharmaceutical companies, and non-profit  
198 organizations. ADNI-I subjects who (i) agreed to  
199 undergo a lumbar puncture, (ii) had results for all three  
200 CSF biomarkers at baseline, and (iii) belonged to either  
201 the control or AD group at baseline, were selected for  
202 the current study. This selection resulted in a dataset  
203 including 96 AD and 109 control subjects. The CSF  
204 biomarker data were obtained using the xMAP plat-  
205 form (Luminex Corp, Austin, Texas) and INNO-BIA  
206 AlzBio3 research-use-only reagents.

207 Table 1 provides baseline characteristics for the two  
208 study populations.

### 209 *Statistical methodology*

#### 210 *Measure of diagnostic accuracy*

211 To establish the joint diagnostic accuracy of the  
212 AD biomarkers, the biomarkers were combined into  
213 a diagnostic score (see below). As a measure of the

Table 1  
Baseline characteristics of the study populations (mean  $\pm$  SD)

Dataset	Group	n	Age (y)	Female (%)	MMSE	A $\beta$ 42* (pg/mL)	Tau* (pg/mL)	Ptau-181* (pg/mL)
VUmc	SMC	251	64 $\pm$ 6.6	104 (41)	28 $\pm$ 1.5	874 $\pm$ 251.0	302 $\pm$ 197.7	52 $\pm$ 24.0
	AD	631	68 $\pm$ 7.5	326 (52)	21 $\pm$ 5.0	465 $\pm$ 161.6	690 $\pm$ 415.4	89 $\pm$ 39.2
ADNI	Control	109	76 $\pm$ 5.3	55 (50)	29 $\pm$ 1.0	206 $\pm$ 54.4	69 $\pm$ 30.2	25 $\pm$ 14.8
	AD	96	75 $\pm$ 8.0	40 (42)	24 $\pm$ 1.9	142 $\pm$ 4.0	122 $\pm$ 57.0	42 $\pm$ 19.8

\*CSF levels of A $\beta$ 1-42, Total tau, and P-tau181p were determined using commercially available single-parameter ELISA kits (INNOTEST<sup>®</sup> AMYLOID(1-42), INNOTEST<sup>®</sup> hTAU Ag, INNOTEST<sup>®</sup> PHOSPHOTAU(181P)) and using the xMAP platform (Luminex Corp, Austin, Texas) and INNO-BIA AlzBio3 reagents at VUmc and ADNI, respectively.

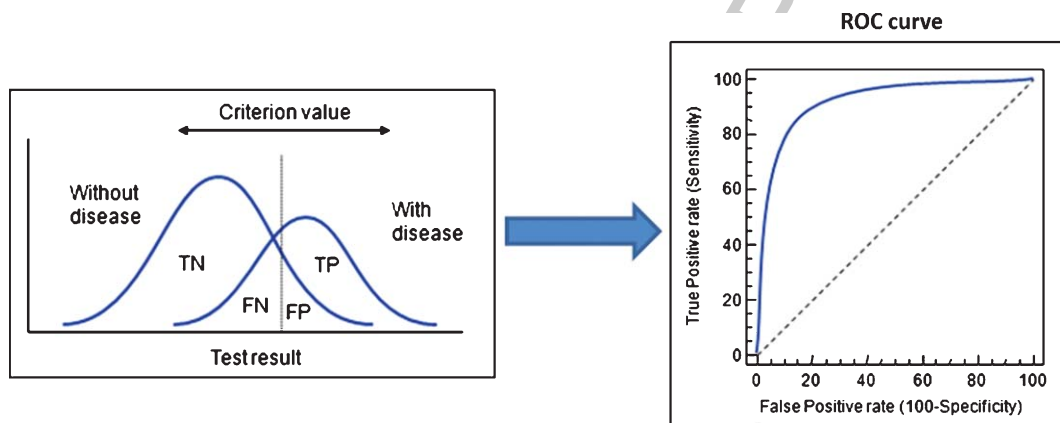


Fig. 1. Schematic summary on the construction of a receiver operating characteristic (ROC) curve and interpretation of the area under the ROC curve (AUC). The ROC curve is a plot of the sensitivity and (1-specificity) for each value of a continuous diagnostic marker. AUC can be interpreted as the probability that, for a randomly selected pair of non-AD and AD subjects, the value of the score for the AD subject will be larger than the value of the non-AD subject. For a score that perfectly separates non-AD and AD populations, the value of AUC is equal to 1, corresponding to the ROC curve passing through the (0,1) point, i.e., the point corresponding to a diagnostic test with 100% sensitivity and 100% specificity. For a score that has no discriminative ability, the value of AUC is equal to 0.5, corresponding to a ROC curve along the diagonal line. TP, true positive; FP, false positive; TN, true negative; FN, false negative.

214 diagnostic performance of this score, the area under the  
215 receiver-operating-characteristics (ROC) curve (AUC)  
216 was used (Fig. 1).

### 217 *AD biomarker performance using a Bayesian* 218 *framework that accounts for an imperfect clinical* 219 *diagnosis*

220 To account for possible errors in the clinical diagno-  
221 sis, both the AD biomarkers AND the clinical diagnosis  
222 were considered as data sources carrying information  
223 about the (unknown) disease status of the subjects.  
224 Note that, in a classical analysis, the clinical diagno-  
225 sis would be taken as the correct disease status, which  
226 does not reflect reality.

227 A Bayesian framework integrates different data  
228 sources in a natural way and is hence most suited for our  
229 purpose. At the core of the Bayesian approach lays the  
230 use of prior information [15]. The information (here-  
231 after also termed ‘prior opinion’ or ‘prior information’)  
232 is provided in the form of probability distributions for  
233 the parameters of a model. The distribution indicates

234 which (sets of) values of the parameters are considered  
235 to be (relatively) more likely than others. In particular,  
236 uninformative distributions (e.g., a normal distribution  
237 with a huge variance) can be used in the data analysis  
238 to imply the absence of any information, i.e., the fact  
239 that all values of a particular parameter are equally  
240 likely. If some information is available, informative  
241 prior distributions are used.

242 By combining the prior distribution with the data,  
243 a posterior distribution for the parameter of interest is  
244 obtained. The posterior distribution reflects the change  
245 of the opinion induced by the data, as compared to  
246 the prior opinion (see Fig. 2). When uninformative  
247 prior distributions are used, the data is used as the only  
248 source of information. In Bayesian analysis, it is best  
249 practice to perform a ‘sensitivity analysis’ using differ-  
250 ent priors to disentangle the effect of prior information  
251 and the analysis dataset on the reported results.

252 In our analyses, we made the ‘conditional inde-  
253 pendence assumption’, i.e., we assumed that AD  
254 biomarkers and clinical diagnosis do not misclassify

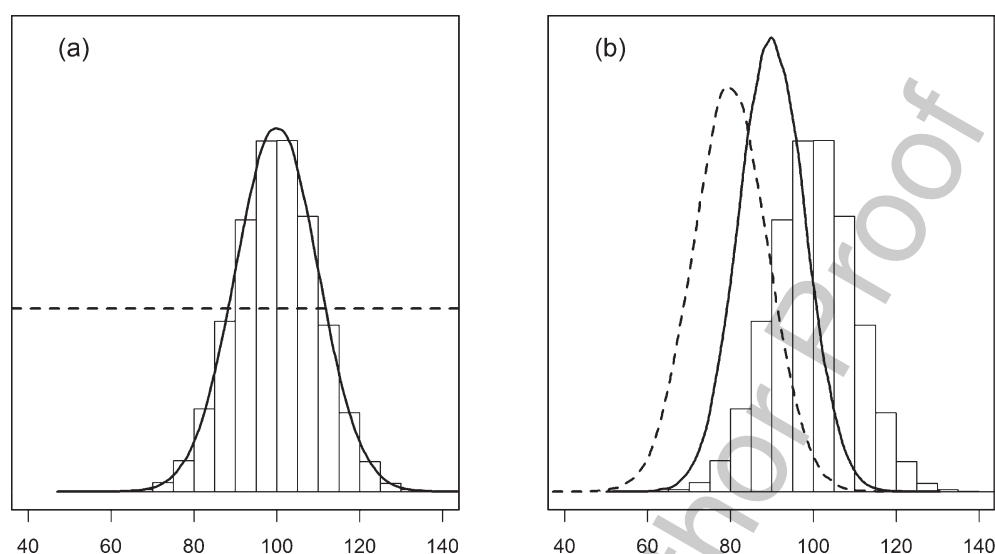


Fig. 2. Schematic illustration of the Bayesian ‘Change of Opinion’ approach. X-axis: Parameter of interest (e.g., average biomarker concentration in  $\text{pg}/\mu\text{l}$ ). Y-axis: Probability of occurrence. Histogram: observed data. Dashed lines: prior opinion (‘prior distribution’). Solid lines: opinion after obtaining data (‘posterior distribution’). Panel (a): Application of an uninformative prior amounts to forming an opinion based solely on the observed data. The horizontal (uninformative) prior distribution indicates that, before data collection, each value is considered equally likely to occur. As a result, the posterior distribution coincides with the observed-data histogram. Panel (b): Application of an informative prior amounts to forming an opinion based on combining the prior information and the observed data. The bell-shaped (informative) prior distribution indicates that, before data collection, the parameter of interest lies, with 95% probability, within the range of 62 to 98. The posterior distribution combines the prior information with the observed data. As a result, the obtained posterior distribution is different from the prior distribution and from the histogram, and it indicates that the value of the parameter lies, with 95% probability, within the range of 75 to 105.

the same individuals. The diagnostic score was constructed by using a linear combination of the biomarkers that maximizes AUC for normally-distributed biomarkers [24].

We used prior distributions for the following parameters: the AUC of a combination of biomarkers, the mean value for each biomarker in the non-AD population, the variances and correlations between all biomarkers in both populations, the prevalence of diseased cases, the sensitivity of the clinical diagnosis, and the specificity of clinical diagnosis. We used uninformative prior distributions for the biomarkers’ means, variances and correlations, and for the disease prevalence.

For the AUC of the linear combination of AD biomarkers, we used more informative priors based on a paper containing data from 12 publications that reported a joint AUC for CSF biomarkers [25]. The lowest reported joint AUC was equal to 0.90 (no standard error provided) [26] and the highest value was equal to 0.997 (95% CI 0.926–1) [27]. Based on those data, we formulated two prior distributions for the joint AUC (Fig. 3a). The first prior distribution implied that the probability that the AUC was larger than 0.7 and 0.9 was equal to 90% and 30%, respectively. This

prior was labeled as ‘optimistic’ in the sense that it pointed toward a high diagnostic accuracy. The second prior distribution choice was labeled as ‘skeptical’ as it suggested that the AUC was around 0.75, with only 5% probability that it exceeded 0.90, the lowest value reported [25].

Also for the specificity and sensitivity of the clinical diagnosis, we used informative priors. Three studies [4, 28, 29] reported high sensitivity of the clinical AD diagnosis (ranging from 81.8% to 100%) in a mixed dementia setting; another study [30] reported much worse sensitivities ranging from 39% to 95% and specificities ranging from 33% to 100%. Based on those data, we formulated two prior distributions (Fig. 3b). The first, ‘optimistic’ prior in accordance with [4, 28, 29], suggested a sensitivity and specificity of about 90%, with 5% probability that sensitivity and specificity were below 80%. The second, more ‘skeptical’ prior, in accordance with [30], was centered at 59%, with a 95% probability that sensitivity and specificity were larger than 25%. The ‘skeptical’ prior assumed less information about the performance of the clinical diagnosis and allowed more flexibility for the biomarkers to ‘override’ the clinical diagnosis, as compared to the ‘optimistic’ prior distribution.

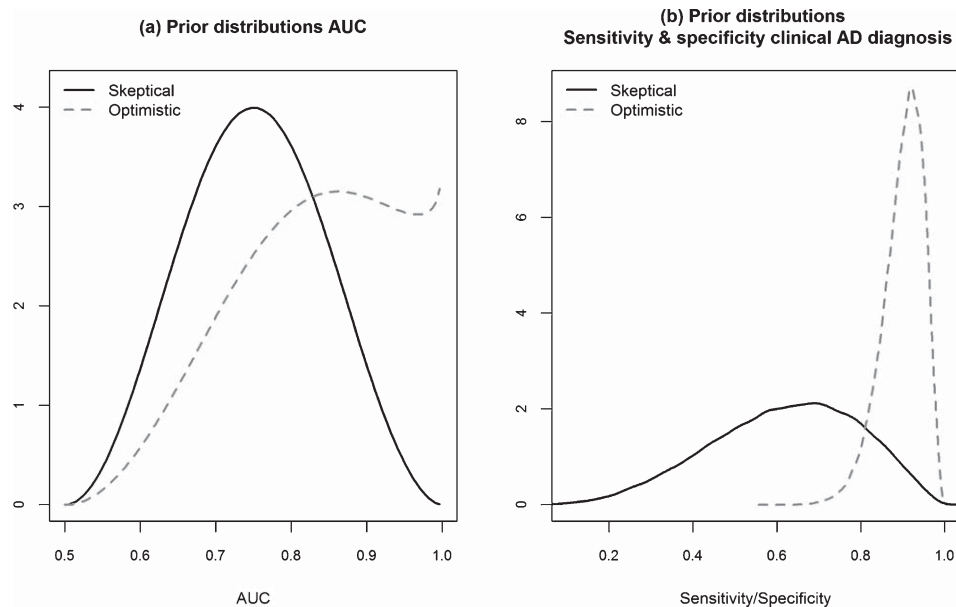


Fig. 3. Prior distributions for the AUC (a) and sensitivity/specificity of the clinical diagnosis (b).

305 If we treat the clinical diagnosis as an imperfect ref-  
 306 erence test, the true disease status of the subjects is  
 307 unknown. It is hence not possible to use a binary clas-  
 308 sifier to establish a ROC curve. Informally speaking,  
 309 the model we use predicts the disease status of the indi-  
 310 viduals that best fits the biomarker and clinical data.  
 311 At the same time, the parameters of a multivariate nor-  
 312 mal distribution for the biomarkers are estimated for  
 313 each group, defined by the predicted disease status of  
 314 the individuals. Based on the estimated distributional  
 315 parameters, a 'bi-normal ROC-curve' [9] is obtained,  
 316 providing estimates of sensitivity and specificity. More  
 317 details on the Bayesian methodology can be found in  
 318 the Supplementary Material.

#### 319 *AD biomarker performance assuming that the* 320 *clinical diagnosis is a perfect reference test*

321 To evaluate the impact of allowing for errors in the  
 322 clinical diagnosis, we also performed two analyses that  
 323 assumed that the clinical diagnosis indicates the correct  
 324 disease status.

325 First, the data were analyzed using logistic regres-  
 326 sion, a methodology that is often applied to evaluate  
 327 AD biomarkers' performance [31, 32]. A diagnostic  
 328 score was calculated with the regression parameters  
 329 and the diagnostic performance of this score was eval-  
 330 uated against the clinical diagnosis.

331 Second, we analyzed the AD biomarkers' perfor-  
 332 mance with the new Bayesian method (see above),  
 333 assuming that the clinical diagnosis is a perfect ref-

334 erence test. Toward this end, sensitivity and specificity  
 335 of clinical diagnosis in the Bayesian model were set to  
 336 1 (i.e., 'extremely' informative priors were used) and  
 337 the prevalence of AD was estimated by the proportion  
 338 of clinical AD subjects in the datasets.

339 By comparison of the results obtained for the latter  
 340 two analyses the effect of the methodology (Bayesian  
 341 method versus classical logistic regression) could be  
 342 evaluated. In addition, the comparison of the results of  
 343 the two Bayesian analyses allowed the evaluation of the  
 344 effect of handling the clinical diagnosis data (perfect  
 345 versus imperfect reference test) on the assessment of  
 346 the diagnostic performance of the AD biomarkers.

#### 347 *Model fitting*

348 The proposed Bayesian method assumed that all  
 349 biomarkers display a normal distribution. To conform  
 350 to this assumption, Total tau and Ptau-181p values were  
 351 log transformed for all analyses. The analyses were  
 352 performed using R [33], version 3.0.1 and OpenBUGS  
 353 [34]. More information on model fitting is provided in  
 354 the Supplementary Material.

355 After fitting the models, the median AUC was  
 356 obtained from the posterior distribution, together with  
 357 a 95% credible interval (CrI), the Bayesian counterpart  
 358 of the 'classical' confidence interval (CI). CrI provides  
 359 the range of values that are expected with 95% proba-  
 bility according to the (posterior) distribution.

## RESULTS

Figure 4 shows the ROC curves for different analyses of the VUmc data (grey) and ADNI data (black). In particular, it shows the curves for the analysis using the logistic regression (dotted), for the Bayesian model obtained by assuming a perfect reference test (dashed), and by assuming an imperfect reference test (solid). Note that the latter were obtained by using the ‘skeptical’ AUC prior and ‘optimistic’ priors for sensitivity and specificity of the clinical diagnosis.

The ROC curves for the logistic regression are close to the curves corresponding to the Bayesian model that also assumed that the clinical diagnosis is a perfect reference test. These results show that the Bayesian method in principle yields the same results as the ‘classical’ logistic regression, proving confidence in our approach. Consequently, we have further focused on the Bayesian methodology.

When assuming that the clinical diagnosis is an imperfect reference test, the ROC curves are higher compared to the corresponding curves obtained when assuming that the reference test is perfect. This shows that, by assuming that the clinical diagnosis flawlessly indicates the pathophysiological AD status, one underestimates the joint diagnostic performance of the biomarkers.

In particular, for the VUmc dataset, the median AUC was equal to 0.949 with 95% CrI [0.935,0.960] when the diagnosis was treated as a perfect reference test

and 0.990 with 95% CrI [0.985,0.995] when treated as an imperfect reference test. For the ADNI data, the corresponding values were equal to 0.870 (95% CrI: [0.817,0.912]) and 0.975 (95% CrI: [0.943,0.990]), respectively.

Figure 5 shows the results of analyses with different prior distributions. The difference between the ROC curves (and hence, AUC) obtained with different combinations of the ‘optimistic’ and ‘skeptical’ prior distributions for the AUC and sensitivity and specificity of the clinical diagnosis was minimal (Fig. 5).

## DISCUSSION

By applying the newly developed Bayesian method to the two datasets, we were able to show that the AUC to discriminate between subjects with AD pathology and controls, increases from 0.949 (with 95% credible interval [0.935,0.960]) to 0.990 ([0.985,0.995]) and from 0.870 ([0.817,0.912]) to 0.975 ([0.943,0.990]) for the VUmc and ADNI cohorts, respectively.

This effect can be intuitively explained as follows. With an imperfect clinical diagnosis, some individuals will be diagnosed as non-AD, while their AD biomarkers may be indicative of existing AD pathophysiology, as biomarker abnormalities can occur decades before clinical symptoms become apparent [35]. For these individuals, the AD biomarkers will be considered as ‘incorrect’ if the clinical diagnosis is regarded as the perfect reference test. Consequently, the performance of the biomarkers will be underestimated. It is in this complex situation that our proposed approach is most useful [7, 36], enabling an estimation of the biomarkers’ performance by objectively examining the strength of statistical relationships among variables.

We applied a Bayesian approach because this allowed integrating different sources of information, while taking into account the absence of a perfect reference test. In Bayesian inference, the specification of prior distributions for the model parameters is needed. It is good practice to perform a sensitivity analysis to check the influence of the choice of the prior distributions on the results and to disentangle the effect of the prior distributions and of the data on the reported results. Toward this end, ‘skeptical’ and ‘optimistic priors’ for the biomarkers’ AUC and sensitivity and specificity of the clinical diagnosis were used in our analysis. The ‘skeptical’ priors were only weakly informative (containing little prior information) while the ‘optimistic’ priors contained more information that pointed to a better diagnostic

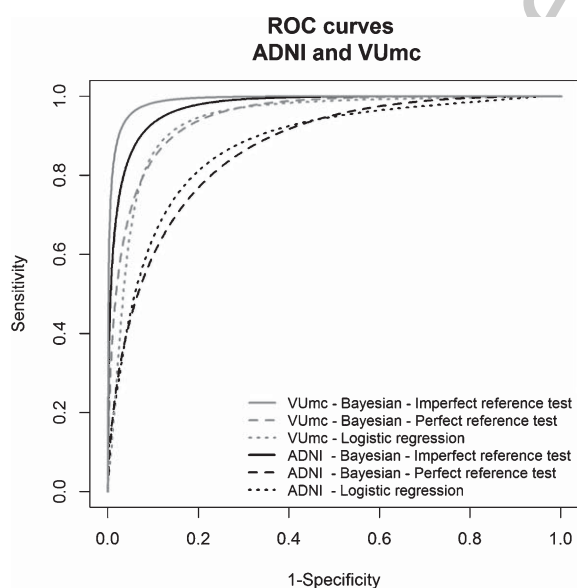


Fig. 4. ROC curves for different analyses for VUmc (grey) and ADNI (black) dataset.



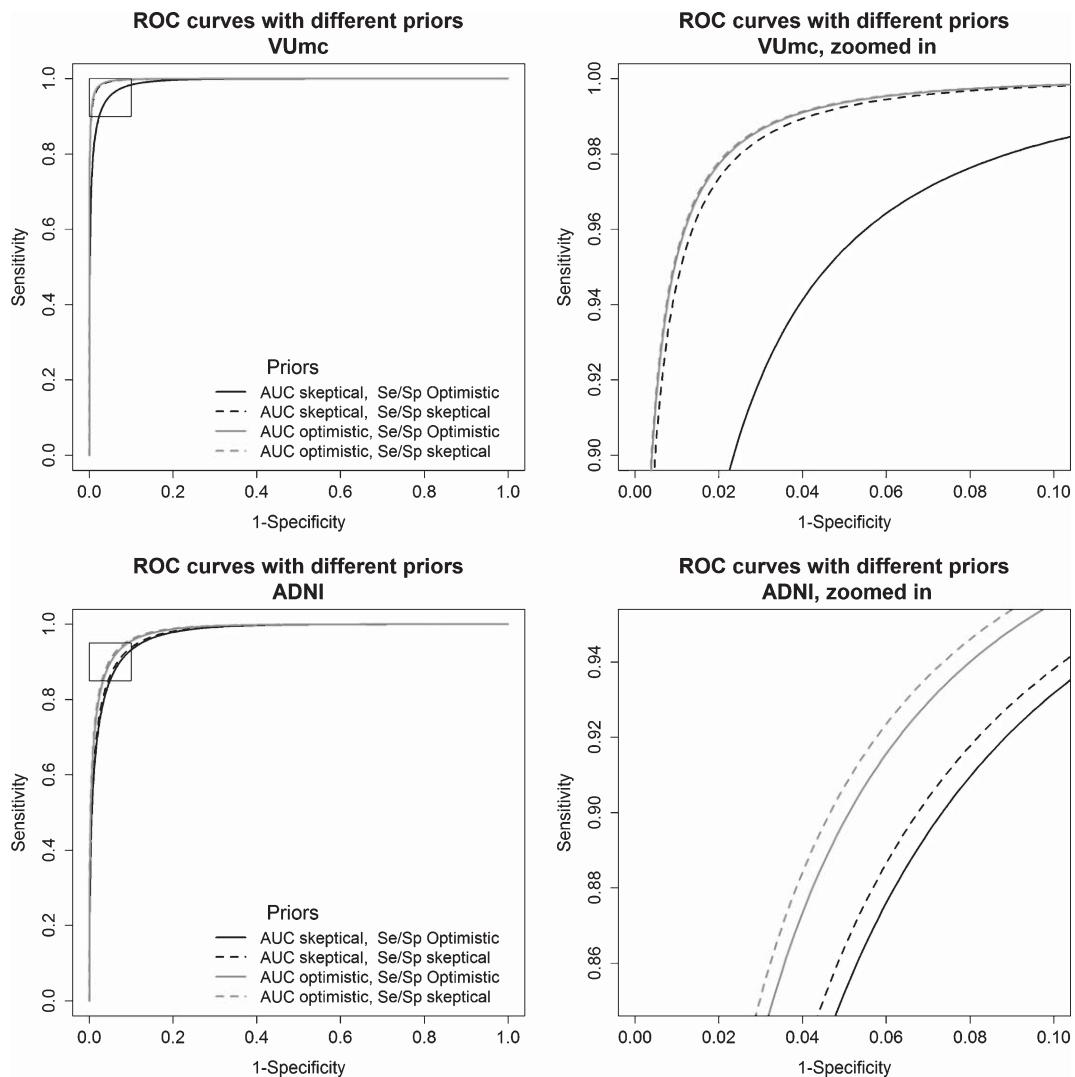


Fig. 5. Sensitivity analysis: ROC curves when using different priors for VUmc (top row) and ADNI dataset (bottom row). The graphs on the right represent the same ROC curves as the graphs on the left, but are zoomed in to the rectangle in the upper left corner. Note that, for VUmc, the solid and dashed grey line overlap.

438 performance of the biomarkers or clinical diagnosis  
 439 as found in literature. Figure 5 shows that the differ-  
 440 ent combinations of these prior distributions resulted  
 441 in minimally different ROC-curves. This implies that  
 442 our conclusions are robust to reasonable changes in  
 443 prior distributions for the diagnostic performance of  
 444 the biomarkers and clinical diagnosis. Put differently,  
 445 the results presented in Fig. 4 are mainly driven by the  
 446 data and not by the prior information.

447 All statistical analyses rely on assumptions.  
 448 Bayesian statistics has the advantage to encour-  
 449 age a thorough consideration and presentation of  
 450 the assumptions underlying the performed analysis.  
 451 We have avoided the assumption that the refer-

452 ence test is perfect, because this has been reported  
 453 to cause biased diagnostic accuracy results [7, 36,  
 454 37]. The validity of the presented approach relies  
 455 on the assumption that the clinical diagnosis and  
 456 AD biomarkers do not misclassify the same sub-  
 457 jects (the ‘conditional independence assumption’).  
 458 At this point, mainly heuristic arguments can be  
 459 offered for the plausibility of this assumption. As  
 460 long as the clinical diagnosis is not based on the  
 461 CSF biomarkers, we can assume that the biomarkers  
 462 and clinical diagnosis do not tend to misclassify the  
 463 same subjects. Furthermore, our findings are in line  
 464 with the reports on lower diagnostic performance of  
 465 CSF biomarkers when evaluated against the clinical

466 diagnosis instead of the pathology confirmed diagno- 518  
467 sis [10]. 519

468 There is no gold standard for a complex disease 520  
469 like AD [3]. We show that this is no longer an 521  
470 issue as the developed Bayesian methodology can 522  
471 deal with the absence of a perfect reference test. The 523  
472 new approach is constructed by assembling compo- 524  
473 nents of methods that have been proposed for the 525  
474 evaluation of the diagnostic performance of a combi- 526  
475 nation of markers [39] when no perfect reference 527  
476 test is available [36]. To our knowledge, this is the 528  
477 first report of the use of a Bayesian approach to 529  
478 define the diagnostic performance of AD biomarkers 530  
479 that acknowledges the absence of a perfect reference 531  
480 test. 532

481 The new methodology is based on well-established 533  
482 statistical concepts, but is more complicated than 534  
483 a simple comparison with the clinical diagnosis or 535  
484 dichotomized PET data as outcome. It is, however, 536  
485 the complexity of a dementia diagnosis that calls for 537  
486 appropriate, more advanced analysis methods. 538

487 The reported diagnostic accuracy results are relevant 539  
488 only for discrimination between the two well-defined 540  
489 groups in this study namely AD versus SMC/control. 541  
490 These estimates of diagnostic accuracy are often 542  
491 higher than expected in the target patient popula- 543  
492 tion which contains difficult-to-diagnose subjects (e.g., 544  
493 MCI patients) [7]. This is not an issue for the pur- 545  
494 pose of our manuscript, as our goals were to develop 546  
495 a new method that allows for an imperfect reference 547  
496 test and to compare the resulting estimates of diagnos- 548  
497 tic accuracy with those obtained by currently applied 549  
498 methodologies. In practice, these extremely high accu- 550  
499 racy estimates will not be achieved because the target 551  
500 patient population will contain difficult-to-diagnose 552  
501 subjects (such as MCI patients) and patients with dif- 553  
502 ferent types of dementia. 554

503 However, the estimates of diagnostic accuracy are 555  
504 expected to be higher in the target patient population 556  
505 when estimated with the Bayesian analysis as com- 557  
506 pared to a classical analysis with the clinical diagnosis 558  
507 as perfect reference test. Although the patterns of dif- 559  
508 ferences between the results for the different models 560  
509 (Fig. 4) were identical for VUmc and ADNI datasets, 561  
510 the numerical values of the AUC estimates were not. 562  
511 For each of the three models, the combined biomark- 563  
512 ers' AUC was higher for the VUmc data than for the 564  
513 ADNI data. This difference is most likely due to the 565  
514 higher age of the ADNI subjects (on average about 10 566  
515 years older than VUmc subjects), as it is well-known 567  
516 that the diagnostic accuracy of CSF AD biomarkers 568  
517 decreases with age [38]. 569

518 The new methodology can now be used for re- 519  
520 investigation of the clinical value of existing AD 521  
522 biomarkers to determine which CSF biomarkers are 523  
524 needed for maximum discriminate between stable and 525  
526 progressing MCI patients or for a differential demen- 527  
528 tia diagnosis. The cut-offs that would be derived from 529  
530 the ROC-curve of the new method will be differ- 531  
532 ent from the current cut-offs values that are set with 533  
534 the clinical diagnosis as perfect reference test. Also 535  
536 the comparison of the clinical value between CSF 537  
538 biomarkers measured using different platforms or A $\beta$  539  
540 PET deposition measured with different tracers could 541  
542 be addressed. Importantly, the new analysis method 543  
544 also supports the direct comparison of the diagnostic 545  
546 value of CSF and imaging biomarkers for A $\beta$  depo- 547  
548 sition. In this way, the interchangeability (assumed in 548  
549 the (preclinical) AD criteria [1, 5]) or complementarity 549  
550 (as suggested by the reported proportion of discordant 550  
551 cases [12–14]) of the two *in vivo* biomarkers could 551  
552 be determined. We anticipate that the use of the new 552  
553 Bayesian framework will lead to a more accurate diag- 553  
554 nosis based on biomarkers and hence more diagnostic 554  
555 confidence in early stages of AD. 555  
556 557  
558 559  
560 561  
562 563  
564 565  
566 567  
568 569  
569 570  
570 571  
571 572  
572 573  
573 574  
574 575  
575 576  
576 577  
577 578  
578 579  
579 580  
580 581  
581 582  
582 583  
583 584  
584 585  
585 586  
586 587  
587 588  
588 589  
589 590  
590 591  
591 592  
592 593  
593 594  
594 595  
595 596  
596 597  
597 598  
598 599  
599 600  
600 601  
601 602  
602 603  
603 604  
604 605  
605 606  
606 607  
607 608  
608 609  
609 610  
610 611  
611 612  
612 613  
613 614  
614 615  
615 616  
616 617  
617 618  
618 619  
619 620  
620 621  
621 622  
622 623  
623 624  
624 625  
625 626  
626 627  
627 628  
628 629  
629 630  
630 631  
631 632  
632 633  
633 634  
634 635  
635 636  
636 637  
637 638  
638 639  
639 640  
640 641  
641 642  
642 643  
643 644  
644 645  
645 646  
646 647  
647 648  
648 649  
649 650  
650 651  
651 652  
652 653  
653 654  
654 655  
655 656  
656 657  
657 658  
658 659  
659 660  
660 661  
661 662  
662 663  
663 664  
664 665  
665 666  
666 667  
667 668  
668 669  
669 670  
670 671  
671 672  
672 673  
673 674  
674 675  
675 676  
676 677  
677 678  
678 679  
679 680  
680 681  
681 682  
682 683  
683 684  
684 685  
685 686  
686 687  
687 688  
688 689  
689 690  
690 691  
691 692  
692 693  
693 694  
694 695  
695 696  
696 697  
697 698  
698 699  
699 700  
700 701  
701 702  
702 703  
703 704  
704 705  
705 706  
706 707  
707 708  
708 709  
709 710  
710 711  
711 712  
712 713  
713 714  
714 715  
715 716  
716 717  
717 718  
718 719  
719 720  
720 721  
721 722  
722 723  
723 724  
724 725  
725 726  
726 727  
727 728  
728 729  
729 730  
730 731  
731 732  
732 733  
733 734  
734 735  
735 736  
736 737  
737 738  
738 739  
739 740  
740 741  
741 742  
742 743  
743 744  
744 745  
745 746  
746 747  
747 748  
748 749  
749 750  
750 751  
751 752  
752 753  
753 754  
754 755  
755 756  
756 757  
757 758  
758 759  
759 760  
760 761  
761 762  
762 763  
763 764  
764 765  
765 766  
766 767  
767 768  
768 769  
769 770  
770 771  
771 772  
772 773  
773 774  
774 775  
775 776  
776 777  
777 778  
778 779  
779 780  
780 781  
781 782  
782 783  
783 784  
784 785  
785 786  
786 787  
787 788  
788 789  
789 790  
790 791  
791 792  
792 793  
793 794  
794 795  
795 796  
796 797  
797 798  
798 799  
799 800  
800 801  
801 802  
802 803  
803 804  
804 805  
805 806  
806 807  
807 808  
808 809  
809 810  
810 811  
811 812  
812 813  
813 814  
814 815  
815 816  
816 817  
817 818  
818 819  
819 820  
820 821  
821 822  
822 823  
823 824  
824 825  
825 826  
826 827  
827 828  
828 829  
829 830  
830 831  
831 832  
832 833  
833 834  
834 835  
835 836  
836 837  
837 838  
838 839  
839 840  
840 841  
841 842  
842 843  
843 844  
844 845  
845 846  
846 847  
847 848  
848 849  
849 850  
850 851  
851 852  
852 853  
853 854  
854 855  
855 856  
856 857  
857 858  
858 859  
859 860  
860 861  
861 862  
862 863  
863 864  
864 865  
865 866  
866 867  
867 868  
868 869  
869 870  
870 871  
871 872  
872 873  
873 874  
874 875  
875 876  
876 877  
877 878  
878 879  
879 880  
880 881  
881 882  
882 883  
883 884  
884 885  
885 886  
886 887  
887 888  
888 889  
889 890  
890 891  
891 892  
892 893  
893 894  
894 895  
895 896  
896 897  
897 898  
898 899  
899 900  
900 901  
901 902  
902 903  
903 904  
904 905  
905 906  
906 907  
907 908  
908 909  
909 910  
910 911  
911 912  
912 913  
913 914  
914 915  
915 916  
916 917  
917 918  
918 919  
919 920  
920 921  
921 922  
922 923  
923 924  
924 925  
925 926  
926 927  
927 928  
928 929  
929 930  
930 931  
931 932  
932 933  
933 934  
934 935  
935 936  
936 937  
937 938  
938 939  
939 940  
940 941  
941 942  
942 943  
943 944  
944 945  
945 946  
946 947  
947 948  
948 949  
949 950  
950 951  
951 952  
952 953  
953 954  
954 955  
955 956  
956 957  
957 958  
958 959  
959 960  
960 961  
961 962  
962 963  
963 964  
964 965  
965 966  
966 967  
967 968  
968 969  
969 970  
970 971  
971 972  
972 973  
973 974  
974 975  
975 976  
976 977  
977 978  
978 979  
979 980  
980 981  
981 982  
982 983  
983 984  
984 985  
985 986  
986 987  
987 988  
988 989  
989 990  
990 991  
991 992  
992 993  
993 994  
994 995  
995 996  
996 997  
997 998  
998 999  
999 1000

## 541 ACKNOWLEDGMENTS 542

543 This research was conducted within the framework 544  
545 of the European EUROTRANS-BIO – ERA-NET 546  
547 project ‘B4AD’, a collaborative project of Interna- 548  
549 tional Drug Development Institute (Louvain-la-Neuve, 549  
550 Belgium), PamGene International (Den Bosch, The 550  
551 Netherlands) and the VU University medical center 551  
552 and the Alzheimer center (Amsterdam, The Nether- 552  
553 lands). We thank Riet Hilhorst, Faris Naji, Rik de Wijn, 553  
554 and Rinie van Beuningen (PamGene) for helpful dis- 554  
555 cussions. 555  
556 557  
558 559  
559 560  
560 561  
561 562  
562 563  
563 564  
564 565  
565 566  
566 567  
567 568  
568 569  
569 570  
570 571  
571 572  
572 573  
573 574  
574 575  
575 576  
576 577  
577 578  
578 579  
579 580  
580 581  
581 582  
582 583  
583 584  
584 585  
585 586  
586 587  
587 588  
588 589  
589 590  
590 591  
591 592  
592 593  
593 594  
594 595  
595 596  
596 597  
597 598  
598 599  
599 600  
600 601  
601 602  
602 603  
603 604  
604 605  
605 606  
606 607  
607 608  
608 609  
609 610  
610 611  
611 612  
612 613  
613 614  
614 615  
615 616  
616 617  
617 618  
618 619  
619 620  
620 621  
621 622  
622 623  
623 624  
624 625  
625 626  
626 627  
627 628  
628 629  
629 630  
630 631  
631 632  
632 633  
633 634  
634 635  
635 636  
636 637  
637 638  
638 639  
639 640  
640 641  
641 642  
642 643  
643 644  
644 645  
645 646  
646 647  
647 648  
648 649  
649 650  
650 651  
651 652  
652 653  
653 654  
654 655  
655 656  
656 657  
657 658  
658 659  
659 660  
660 661  
661 662  
662 663  
663 664  
664 665  
665 666  
666 667  
667 668  
668 669  
669 670  
670 671  
671 672  
672 673  
673 674  
674 675  
675 676  
676 677  
677 678  
678 679  
679 680  
680 681  
681 682  
682 683  
683 684  
684 685  
685 686  
686 687  
687 688  
688 689  
689 690  
690 691  
691 692  
692 693  
693 694  
694 695  
695 696  
696 697  
697 698  
698 699  
699 700  
700 701  
701 702  
702 703  
703 704  
704 705  
705 706  
706 707  
707 708  
708 709  
709 710  
710 711  
711 712  
712 713  
713 714  
714 715  
715 716  
716 717  
717 718  
718 719  
719 720  
720 721  
721 722  
722 723  
723 724  
724 725  
725 726  
726 727  
727 728  
728 729  
729 730  
730 731  
731 732  
732 733  
733 734  
734 735  
735 736  
736 737  
737 738  
738 739  
739 740  
740 741  
741 742  
742 743  
743 744  
744 745  
745 746  
746 747  
747 748  
748 749  
749 750  
750 751  
751 752  
752 753  
753 754  
754 755  
755 756  
756 757  
757 758  
758 759  
759 760  
760 761  
761 762  
762 763  
763 764  
764 765  
765 766  
766 767  
767 768  
768 769  
769 770  
770 771  
771 772  
772 773  
773 774  
774 775  
775 776  
776 777  
777 778  
778 779  
779 780  
780 781  
781 782  
782 783  
783 784  
784 785  
785 786  
786 787  
787 788  
788 789  
789 790  
790 791  
791 792  
792 793  
793 794  
794 795  
795 796  
796 797  
797 798  
798 799  
799 800  
800 801  
801 802  
802 803  
803 804  
804 805  
805 806  
806 807  
807 808  
808 809  
809 810  
810 811  
811 812  
812 813  
813 814  
814 815  
815 816  
816 817  
817 818  
818 819  
819 820  
820 821  
821 822  
822 823  
823 824  
824 825  
825 826  
826 827  
827 828  
828 829  
829 830  
830 831  
831 832  
832 833  
833 834  
834 835  
835 836  
836 837  
837 838  
838 839  
839 840  
840 841  
841 842  
842 843  
843 844  
844 845  
845 846  
846 847  
847 848  
848 849  
849 850  
850 851  
851 852  
852 853  
853 854  
854 855  
855 856  
856 857  
857 858  
858 859  
859 860  
860 861  
861 862  
862 863  
863 864  
864 865  
865 866  
866 867  
867 868  
868 869  
869 870  
870 871  
871 872  
872 873  
873 874  
874 875  
875 876  
876 877  
877 878  
878 879  
879 880  
880 881  
881 882  
882 883  
883 884  
884 885  
885 886  
886 887  
887 888  
888 889  
889 890  
890 891  
891 892  
892 893  
893 894  
894 895  
895 896  
896 897  
897 898  
898 899  
899 900  
900 901  
901 902  
902 903  
903 904  
904 905  
905 906  
906 907  
907 908  
908 909  
909 910  
910 911  
911 912  
912 913  
913 914  
914 915  
915 916  
916 917  
917 918  
918 919  
919 920  
920 921  
921 922  
922 923  
923 924  
924 925  
925 926  
926 927  
927 928  
928 929  
929 930  
930 931  
931 932  
932 933  
933 934  
934 935  
935 936  
936 937  
937 938  
938 939  
939 940  
940 941  
941 942  
942 943  
943 944  
944 945  
945 946  
946 947  
947 948  
948 949  
949 950  
950 951  
951 952  
952 953  
953 954  
954 955  
955 956  
956 957  
957 958  
958 959  
959 960  
960 961  
961 962  
962 963  
963 964  
964 965  
965 966  
966 967  
967 968  
968 969  
969 970  
970 971  
971 972  
972 973  
973 974  
974 975  
975 976  
976 977  
977 978  
978 979  
979 980  
980 981  
981 982  
982 983  
983 984  
984 985  
985 986  
986 987  
987 988  
988 989  
989 990  
990 991  
991 992  
992 993  
993 994  
994 995  
995 996  
996 997  
997 998  
998 999  
999 1000

552 Research of the VUmc Alzheimer center and the 552  
553 Department of Pathology is part of the Neurodegen- 553  
554 eration research program of the Neuroscience Campus 554  
555 Amsterdam. The VUmc Alzheimer center is supported 555  
556 by Alzheimer Nederland and Stichting VUmc fonds. 556  
557 The VUmc clinical database structure was developed 557  
558 with funding from Stichting Dioraphte. 558

559 Authors' disclosures available online ([http://j- 559  
560 alz.com/manuscript-disclosures/14-2886r2](http://j-alz.com/manuscript-disclosures/14-2886r2)). 560

561 Data collection and sharing for this project was 561  
562 funded by the Alzheimer's Disease Neuroimaging Ini- 562  
563 tiative (ADNI) (National Institutes of Health Grant 563  
564 U01 AG024904) and DOD ADNI (Department of 564  
565 Defense award number W81XWH-12-2-0012). ADNI 565  
566 is funded by the National Institute on Aging, the 566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972

567 National Institute of Biomedical Imaging and Bio-  
 568 engineering, and through generous contributions from  
 569 the following: Alzheimer's Association; Alzheimer's  
 570 Drug Discovery Foundation; BioClinica, Inc.; Biogen  
 571 Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.;  
 572 Elan Pharmaceuticals, Inc.; Eli Lilly and Company;  
 573 F. Hoffmann-La Roche Ltd and its affiliated com-  
 574 pany Genentech, Inc.; GE Healthcare; Innogenetics,  
 575 N.V.; IXICO Ltd.; Janssen Alzheimer Immunother-  
 576 apy Research & Development, LLC.; Johnson &  
 577 Johnson Pharmaceutical Research & Development  
 578 LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso  
 579 Scale Diagnostics, LLC.; NeuroRx Research; Novar-  
 580 tis Pharmaceuticals Corporation; Pfizer Inc.; Piramal  
 581 Imaging; Servier; Synarc Inc.; and Takeda Pharma-  
 582 ceutical Company. The Canadian Institutes of Health  
 583 Research is providing funds to support ADNI clinical  
 584 sites in Canada. Private sector contributions are  
 585 facilitated by the Foundation for the National Institutes  
 586 of Health (<http://www.fnih.org>). The grantee organiza-  
 587 tion is the Northern California Institute for Research  
 588 and Education, and the study is Rev December 5, 2013  
 589 coordinated by the Alzheimer's Disease Cooperative  
 590 Study at the University of California, San Diego. ADNI  
 591 data are disseminated by the Laboratory for Neuro  
 592 Imaging at the University of Southern California.

## 593 SUPPLEMENTARY MATERIAL

594 The supplementary material is available in the  
 595 electronic version of this article: [http://dx.doi.org/](http://dx.doi.org/10.3233/JAD-142886)  
 596 [10.3233/JAD-142886](http://dx.doi.org/10.3233/JAD-142886).

## 597 REFERENCES

- 598 [1] McKhann GM, Knopman DS, Chertkow H, Hyman BT,  
 599 Jack CR Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly  
 600 JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Schel-  
 601 tens P, Carrillo MC, Thies B, Weintraub S, Phelps CH  
 602 (2011) The diagnosis of dementia due to Alzheimer's dis-  
 603 ease: Recommendations from the National Institute on  
 604 Aging-Alzheimer's Association workgroups on diagnostic  
 605 guidelines for Alzheimer's disease. *Alzheimers Dement* **7**,  
 606 263-269.
- 607 [2] Vos SJB, Xiong C, Visser PJ, Jasielc MS, Hassenstab J,  
 608 Grant EA, Cairns NJ, Morris JC, Holtzman DM, Fagan AM  
 609 (2013) Preclinical Alzheimer's disease and its outcome: A  
 610 longitudinal cohort study. *Lancet Neurol* **12**, 957-965.
- 611 [3] Scheltens P, Rockwood K (2011) How golden is the gold  
 612 standard of neuropathology in dementia? *Alzheimers Dement*  
 613 **7**, 486-489.
- 614 [4] Beach TG, Monsell SE, Phillips LE, Kukull W (2012) Accu-  
 615 racy of the clinical diagnosis of Alzheimer disease at National  
 616 Institute on Aging Alzheimer Disease Centers, 2005-2010.  
 617 *J Neuropathol Exp Neurol* **71**, 266-273.
- [5] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S,  
 618 Fagan AM, Iwatsubo T, Jack CR Jr, Kaye J, Montine TJ, Park  
 619 DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K,  
 620 Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV,  
 621 Phelps CH (2011) Toward defining the preclinical stages of  
 622 Alzheimer's disease: Recommendations from the National  
 623 Institute on Aging and the Alzheimer's Association work-  
 624 group. *Alzheimers Dement* **7**, 280-292.
- 625 [6] Salloway S, Sperling R, Fox NC, Blennow K, Klunk W,  
 626 Raskind M, Sabbagh M, Honig LS, Porsteinsson AP, Ferris S,  
 627 Reichert M, Ketter N, Nejadnik B, Guenzler V, Miloslavsky  
 628 M, Wang D, Lu Y, Lull J, Tudor IC, Liu E, Grundman M,  
 629 Yuen E, Black R, Brashear HR; Bapineuzumab 301 and 302  
 630 Clinical Trial Investigators (2014) Two phase 3 trials of bap-  
 631 ineuzumab in mild-to-moderate Alzheimer's disease. *N Engl*  
 632 *J Med* **370**, 322-333.
- 633 [7] Reitsma JB, Rutjes AWS, Khan KS, Bossuyt PM (2009) A  
 634 review of solutions for diagnostic accuracy studies with an  
 635 imperfect or missing reference standard. *J Clin Epidemiol*  
 636 **62**, 797-806.
- 637 [8] Valenstein PN (1990) Evaluating diagnostic tests with imper-  
 638 fect standards. *Am. J Clin Pathol* **93**, 252-258.
- 639 [9] Zhou X-H, Obuchowski NA, McClish DK (2002) *Statistical*  
 640 *Methods in Diagnostic Medicine*, Wiley Inc., New York.
- 641 [10] Toledo JB, Brettschneider J, Grossman M, Arnold SE,  
 642 Hu WT, Xie SX, Lee VM, Shaw LM, Trojanowski JQ  
 643 (2012) CSF biomarkers cutoffs: The importance of coin-  
 644 cident neuropathological diseases. *Acta Neuropathol* **124**,  
 645 23-35.
- 646 [11] De Meyer G, Shapiro F, Vanderstichele H, Vanmechelen  
 647 E, Engelborghs S, De Deyn PP, Coart E, Hansson O,  
 648 Minthon L, Zetterberg H, Blennow K, Shaw L, Trojanowski  
 649 JQ, Alzheimer's Disease Neuroimaging Initiative (2010)  
 650 Diagnosis-independent Alzheimer disease biomarker signa-  
 651 ture in cognitively normal elderly people. *Arch Neurol* **67**,  
 652 949-956.
- 653 [12] Fagan AM, Shaw LM, Xiong C, Vanderstichele H, Mintun  
 654 MA, Trojanowski JQ, Coart E, Morris JC, Holtzman DM  
 655 (2011) Comparison of analytical platforms for cerebrospinal  
 656 fluid measures of  $\beta$ -amyloid 1-42, Total tau, and P-tau181  
 657 for identifying Alzheimer disease amyloid plaque pathology.  
 658 *Arch Neurol* **68**, 1137-1144.
- 659 [13] Palmqvist S, Zetterberg H, Blennow K, Vestberg S, Andreasson  
 660 U, Brooks DJ, Owenius R, Hägerström D, Wollmer P,  
 661 Minthon L, Hansson O (2014) Accuracy of brain amyloid  
 662 detection in clinical practice using cerebrospinal fluid  $\beta$ -  
 663 amyloid 42. A cross-validation study against amyloid positron  
 664 emission tomography. *JAMA Neurol* **71**, 1282-1289.
- 665 [14] Landau SM, Lu M, Joshi AD, Pontecorvo M, Mintun MA,  
 666 Trojanowski JQ, Shaw LM, Jagust WJ, Alzheimer's Disease  
 667 Neuroimaging Initiative (2013) Comparing positron emission  
 668 tomography imaging and cerebrospinal fluid measurements of  
 669  $\beta$ -amyloid. *Ann Neurol* **74**, 826-836.
- 670 [15] Spiegelhalter DJ, Abrams KR, Myles JP (2004) *Bayesian*  
 671 *Approaches to Clinical Trials and Health-Care Evaluations*.  
 672 Wiley Inc., New York.
- 673 [16] Gelman A, Carlin JB, Stern HS, Rubin DB (2003) *Bayesian*  
 674 *Data Analysis, second edition*. Chapman & Hall/CRC,  
 675 New York.
- 676 [17] Berry SM, Berry DA, Natarajana K, Lina C-S, Hennekens  
 677 CH, Belder R (2004) Bayesian survival analysis with nonpro-  
 678 portional hazards. *J Am Stat Assoc* **99**, 36-44.
- 679 [18] Anonymous (2010) Guidance for Industry and FDA Staff;  
 680 Guidance for the Use of Bayesian Statistics in Medical  
 681 Device Clinical Trials. <http://www.fda.gov/downloads/>  
 682

- 683 MedicalDevices/DeviceRegulationandGuidance/Guidance  
684 Documents/ucm071121.pdf
- 685 [19] Adamina M, Tomlinson G, Guller U (2009) Bayesian statistics  
686 in oncology. *Cancer* **115**, 5371-5381. 732
- 687 [20] Schoenfeld DA, Zheng H, Finkelstein DM (2009) Bayesian  
688 design using adult data to augment pediatric trials. *Clin Trials*  
689 **6**, 297-304. 733
- 690 [21] Broemeling LD (2007) *Bayesian Biostatistics and Diagnostic*  
691 *Medicine*. Chapman & Hall/CRC, New York. 734
- 692 [22] Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-  
693 Gateau P, Cummings J, Delacourte A, Galasko D, Gauthier  
694 S, Jicha G, Meguro K, O'brien J, Pasquier F, Robert P,  
695 Rossor M, Salloway S, Stern Y, Visser PJ, Scheltens P (2007)  
696 Research criteria for the diagnosis of Alzheimer's disease:  
697 Revising the NINCDS-ADRDA criteria. *Lancet Neurol* **6**,  
698 734-746. 735
- 699 [23] Duits FH, Teunissen CE, Bouwman FH, Visser PJ, Matts-  
700 son N, Zetterberg H, Blennow K, Hansson O, Minthon L,  
701 Andreasen N, Marcusson J, Wallin A, Rikkert MO, Tso-  
702 laki M, Parnetti L, Herukka SK, Hampel H, De Leon MJ,  
703 Schröder J, Aarsland D, Blankenstein MA, Scheltens P, van  
704 der Flier WM (2014) The cerebrospinal fluid 'Alzheimer pro-  
705 file': Easily said, but what does it mean? *Alzheimers Dement*  
706 **10**, 713-723. 736
- 707 [24] Su JQ, Liu JS (1993) Linear combinations of multiple diag-  
708 nostic markers. *J Am Stat Assoc* **88**, 1350-1355. 737
- 709 [25] Bloudek LM, Spackman DE, Blankenburg M, Sullivan SD  
710 (2011) Review and meta-analysis of biomarkers and diag-  
711 nostic imaging in Alzheimer's disease. *J Alzheimers Dis* **26**,  
712 627-645. 738
- 713 [26] Ibach B, Binder H, Dragon M, Poljansky S, Haen E, Schmitz  
714 E, Koch H, Putzhammer A, Klauenemann H, Wieland W,  
715 Hajak G (2006) Cerebrospinal fluid tau and beta-amyloid  
716 in Alzheimer patients, disease controls and an age-matched  
717 random sample. *Neurobiol Aging* **27**, 1202-1211. 739
- 718 [27] Kapaki E, Liappas I, Paraskevas GP, Theotoka I, Rabavilas A  
719 (2005) The diagnostic value of tau protein, beta-amyloid (1-  
720 42) and their ratio for the discrimination of alcohol-related  
721 cognitive disorders from Alzheimer's disease in the early  
722 stages. *Int J Geriatr Psychiatry* **20**, 722-729. 740
- 723 [28] Schoonenboom NS, Reesink FE, Verwey NA, Kester MI, Teu-  
724 nissen CE, van de Ven PM, Pijnenburg YA, Blankenstein MA,  
725 Rozemuller AJ, Scheltens P, van der Flier WM (2012) Cere-  
726 brospinal fluid markers for differential dementia diagnosis in  
727 a large memory clinic cohort. *Neurology* **3**, 47-54. 741
- 728 [29] Toledo JB, Cairns NJ, Da X, Chen K, Carter D, Fleisher A,  
729 Householder E, Ayutyanont N, Roontiva A, Bauer RJ, Eisen  
730 P, Shaw LM, Davatzikos C, Weiner MW, Reiman EM, Mor-  
731 ris JC, Trojanowski JQ; Alzheimer's Disease Neuroimaging  
732 Initiative (ADNI) (2013) Clinical and multimodal biomarker  
733 correlates of ADNI neuropathological findings. *Acta Neu-  
734 ropathol Commun* **1**, 65. 735
- [30] Wollman DE, Prohovnik I (2003) Sensitivity and specificity  
736 of neuroimaging. *Dialogues Clin Neurosci* **5**, 89-99. 737
- [31] Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K,  
738 Minthon L (2006) Association between CSF biomarkers and  
739 incipient Alzheimer's disease in subjects with mild cognitive  
740 impairment: A follow-up study. *Lancet Neurol* **5**, 228-234. 741
- [32] Mattsson N, Zetterberg H, Hansson O, Andreasen N, Parnetti  
742 L, Jonsson M, Herukka SK, van der Flier WM, Blankenstein  
743 MA, Ewers M, Rich K, Kaiser E, Verbeek M, Tsolaki M,  
744 Mulugeta E, Rosén E, Aarsland D, Visser PJ, Schröder J,  
745 Marcusson J, de Leon M, Hampel H, Scheltens P, Pirttilä T,  
746 Wallin A, Jönköping ME, Minthon L, Winblad B, Blennow  
747 K (2009) CSF biomarkers and incipient Alzheimer disease in  
748 patients with mild cognitive impairment. *JAMA* **302**, 385-393. 742
- [33] R Core Team (2013) R: A language and environment for stati-  
749 stical computing. R Foundation for Statistical Computing,  
750 Vienna, Austria. <http://www.R-project.org/>  
751 743
- [34] Lunn D, Spiegelhalter D, Thomas A, Best N (2009) The  
752 BUGS project: Evolution, critique and future directions (with  
753 discussion). *Stat Med* **28**, 3049-3082. 744
- [35] Jack CR Jr, Knopman DS, Jagust WJ, Petersen RC, Weiner  
754 MW, Aisen PS, Shaw LM, Vemuri P, Wiste HJ, Weigand SD,  
755 Lesnick TG, Pankratz VS, Donohue MC, Trojanowski JQ  
756 (2013) Tracking pathophysiological processes in Alzheimer's  
757 disease: An updated hypothetical model of dynamic biomark-  
758 ers. *Lancet Neurol* **12**, 207-216. 745
- [36] Scott AN, Joseph L, Bélisle P, Behr MA, Schwartzman K  
759 (2007) Bayesian modeling of tuberculosis clustering from  
760 DNA fingerprint data. *Stat Med* **27**, 140-156. 746
- [37] Lu Y, Dendrukuri N, Schiller I, Joseph L (2010) A Bayesian  
761 approach to simultaneously adjusting for verification and refer-  
762 ence standard bias in diagnostic test studies. *Stat Med* **29**,  
763 2532-2543. 747
- [38] Mattsson N, Rosén E, Hansson O, Andreasen N, Parnetti L,  
764 Jonsson M, Herukka SK, van der Flier WM, Blankenstein  
765 MA, Ewers M, Rich K, Kaiser E, Verbeek MM, Olde Rikkert  
766 M, Tsolaki M, Mulugeta E, Aarsland D, Visser PJ, Schröder  
767 J, Marcusson J, de Leon M, Hampel H, Scheltens P, Wallin  
768 A, Eriksdotter-Jönköping M, Minthon L, Winblad B, Blennow  
769 K, Zetterberg H (2012) Age and diagnostic performance of  
770 Alzheimer disease CSF biomarkers. *Neurology* **78**, 468-476. 748
- [39] O'Malley AJ, Zou KH, Fielding JR, Tempany CMC (2001)  
771 Bayesian regression methodology for estimating a receiver  
772 operating characteristic curve with two radiologic applica-  
773 tions: Prostrate biopsy and spiral CT of uterine stones. *Acad*  
774 *Radiol* **8**, 713-725. 749