Unbiased approach to counteract upward drift in cerebrospinal fluid amyloid-β 1–42 analysis results

Betty M. Tijms
Eline A.J. Willemse
Marissa D. Zwan
Sandra D. Mulder
Pieter Jelle Visser
Bart N.M. van Berckel
Wiesje M. van der Flier
Philip Scheltens
Charlotte E. Teunissen

Clinical Chemistry (2018) 64(3):576-585
Abstract
Low cerebrospinal fluid (CSF) amyloid-β 1–42 (Aβ 1–42) concentrations indicate amyloid plaque accumulation in the brain, a pathological hallmark of Alzheimer disease (AD). Innotest assay values of Aβ 1–42 have gradually increased over the past 2 decades, which might lead to misclassification of AD when a single cutpoint for abnormality is used. We propose an unbiased approach to statistically correct for drift.
We determined year-specific cutpoints with Gaussian mixture modeling, based on the cross-section of bimodal distributions of Aβ 1–42 concentrations in 4397 memory clinic patients. This allowed us to realign year-specific cutpoints as an unbiased method to remove drift from the data. Sensitivity and specificity to detect AD dementia were compared between corrected and uncorrected values.
Aβ 1–42 values increased 22 pg/mL annually, and this could not be explained by changes in cohort composition. Our approach removed time dependencies [β (SE) = 0.07 (0.59), P = 0.91]. Statistically correcting for drift improved the sensitivity to detect AD dementia to 0.90 (95% CI, 0.89–0.92) from at least 0.66 (95% CI, 0.64–0.69) based on uncorrected data. Specificity became lower (0.69; 95% CI, 0.67–0.70) vs at most 0.80 (95% CI, 0.79–0.82) for uncorrected data.
This approach may also be useful to standardize Aβ 1–42 CSF concentrations across different centers and/or platforms, and to optimize use of CSF biomarker data collected over a long period.
Introduction

Amyloid plaque deposition in the brain is a pathological hallmark in Alzheimer disease (AD). Amyloid-β 1–42 (Aβ 1–42) concentrations in cerebrospinal fluid (CSF) can be used as a biomarker for the presence of abnormal amyloid aggregation in vivo. Reduced CSF Aβ 1–42 concentrations have been robustly associated with the clinical dementia syndrome \(^{1,2}\), are related with postmortem-determined amyloid plaque burden \(^{1,2}\), and are associated in nondemented persons with a substantially increased risk to develop dementia in the future \(^{3,4}\). Consequently, CSF markers for amyloid abnormality are now part of established and widely used research criteria for AD across the cognitive spectrum \(^{5–9}\). Clinical use of CSF Aβ 1–42 concentrations, however, seems hampered by preanalytical and analytical variability in CSF, which complicates the development of a universal cutpoint to determine amyloid abnormality \(^{10,11}\). A further problematic observation is that Aβ 1–42 values as measured with Innotest ELISA, a broadly used assay, seem to have increased gradually over the past 2 decades \(^{10–14}\), which is also reflected by increasing cutpoints over the years reported in several cohorts \(^{12,13–17}\). One approach to remove drift effects is to reanalyze CSF samples in a single ELISA batch \(^{13,14}\), suggesting that the observed drift is not related to changes in patient characteristic over time but rather to changes in ELISA kits. However, this approach is infeasible with growing cohorts, as well as costly, thus impeding high-quality research on CSF markers over and across cohorts.

We propose an alternative approach to statistically correct for drifts in Aβ 1–42 values over time by taking as a starting point the observation that amyloid measurements show a bimodal distribution, which has been robustly observed in CSF and positron emission tomography (PET) data \(^{12,16–20}\). Bimodal distributions reflect the presence of a normal and a pathological population, and a cutpoint for amyloid abnormality can be determined based on the intersection of these 2 distributions \(^{12,14}\). We hypothesized that bimodal distributions can be identified independently of absolute values of amyloid, and that drift effects can be removed in an unbiased way by aligning cutpoints across time.

Our aim was to develop a statistical approach to correct for upward drift in Aβ 1–42 concentrations over time by determining year-specific cutpoints based on Gaussian mixture modeling and aligning these over time. We validated our approach in 3 ways: (a) by comparing drift-corrected values with remeasured Aβ 1–42 concentrations that were available for a subset of individuals with AD-type dementia, (b) by comparing drift-corrected values with visual reads of PET for amyloid abnormality that were available for a subset of individuals, and (c) by evaluating whether our approach improved detection of clinical AD-type dementia compared with uncorrected data.
Material and Methods

Study participant selection

Study participants who visited our memory clinic between 2000 and 2016 were selected from the Amsterdam Dementia Cohort when they had a CSF sample available. Most participants received standardized dementia screening during their first visit as described before. Diagnosis for each participant was made in a multidisciplinary meeting according to established criteria as described previously, and we labeled individuals as having AD-type dementia, mild cognitive impairment (MCI), or subjective cognitive decline (SCD) accordingly. Other participants were labeled as having a non-AD type of dementia when they were diagnosed with frontotemporal dementia ($n = 305$), dementia with Lewy bodies ($n = 146$), or vascular dementia ($n = 76$). Individuals who received a diagnosis unrelated to neurodegenerative disease ($n = 594$) or when their diagnosis could not be established at the first visit ($n = 335$) were labeled as "other." The VU University Medical Center (VUmc) ethical review board approved the study. All participants provided written informed consent to use their data for research purposes.

CSF analysis

CSF samples were obtained by lumbar puncture using a 25-gauge needle and syringe between the L3/L4, L4/L5, or L5/S1 intervertebral space, collected in polypropylene tubes, and processed as previously described. Aβ 1–42 concentrations were determined with sandwich ELISAs (Innotest, β-AMYLOID (1–42), Fujirebio (formerly Innogenetics)) at the neurochemistry laboratory of the Department of Clinical Chemistry of VUmc.

PET imaging

A subset of individuals had visual amyloid PET scans available within a year of CSF sampling (obtained, on average, within 1.7 (2.4) months). All PET scans were labeled as amyloid positive or negative based on visual reading by an experienced nuclear medicine physician (BvB). Amyloid PET imaging was conducted using $^{11}$C-PiB ($n = 177$), $^{18}$F-flutemetamol ($n = 112$), $^{18}$F-florbetaben ($n = 300$), and $^{18}$F-florbetapir ($n = 13$). PET methodology has been previously described in detail, except for $^{18}$F-florbetapir. $^{18}$F-florbetapir images were made on a Gemini TF-64 PET/CT scanner (Philips Medical Systems) or a Philips Ingenity TF PET/CT scanner and consisted of 2 X 10-min frames acquired 50 to 70 min after bolus net injection of approximately 370 MBq $^{18}$F-florbetapir.

Statistical analyses

Changes in cohort composition over time were assessed with linear regression for continuous variables of age and Mini-Mental State Examination (MMSE) score, with Pearson trend tests for

1. Chapter 4

Material and Methods

Study participant selection

Study participants who visited our memory clinic between 2000 and 2016 were selected from the Amsterdam Dementia Cohort when they had a CSF sample available. Most participants received standardized dementia screening during their first visit as described before. Diagnosis for each participant was made in a multidisciplinary meeting according to established criteria as described previously, and we labeled individuals as having AD-type dementia, mild cognitive impairment (MCI), or subjective cognitive decline (SCD) accordingly. Other participants were labeled as having a non-AD type of dementia when they were diagnosed with frontotemporal dementia ($n = 305$), dementia with Lewy bodies ($n = 146$), or vascular dementia ($n = 76$). Individuals who received a diagnosis unrelated to neurodegenerative disease ($n = 594$) or when their diagnosis could not be established at the first visit ($n = 335$) were labeled as "other." The VU University Medical Center (VUmc) ethical review board approved the study. All participants provided written informed consent to use their data for research purposes.

CSF analysis

CSF samples were obtained by lumbar puncture using a 25-gauge needle and syringe between the L3/L4, L4/L5, or L5/S1 intervertebral space, collected in polypropylene tubes, and processed as previously described. Aβ 1–42 concentrations were determined with sandwich ELISAs (Innotest, β-AMYLOID (1–42), Fujirebio (formerly Innogenetics)) at the neurochemistry laboratory of the Department of Clinical Chemistry of VUmc.

PET imaging

A subset of individuals had visual amyloid PET scans available within a year of CSF sampling (obtained, on average, within 1.7 (2.4) months). All PET scans were labeled as amyloid positive or negative based on visual reading by an experienced nuclear medicine physician (BvB). Amyloid PET imaging was conducted using $^{11}$C-PiB ($n = 177$), $^{18}$F-flutemetamol ($n = 112$), $^{18}$F-florbetaben ($n = 300$), and $^{18}$F-florbetapir ($n = 13$). PET methodology has been previously described in detail, except for $^{18}$F-florbetapir. $^{18}$F-florbetapir images were made on a Gemini TF-64 PET/CT scanner (Philips Medical Systems) or a Philips Ingenity TF PET/CT scanner and consisted of 2 X 10-min frames acquired 50 to 70 min after bolus net injection of approximately 370 MBq $^{18}$F-florbetapir.

Statistical analyses

Changes in cohort composition over time were assessed with linear regression for continuous variables of age and Mini-Mental State Examination (MMSE) score, with Pearson trend tests for
nominal variables in diagnostic group, sex, or apolipoprotein E (APOE) ε4 genotype, and time as an ordered categorical variable. Changes in CSF Aβ 1–42 concentrations over time were first determined for the total cohort with linear regression. If cohort composition (i.e., diagnostic group, age, sex, or APOE ε4 genotype) showed time dependencies, we performed additional analyses to test whether drift effects could be explained by such variables. We calculated year-specific cutpoints to determine amyloid status with Gaussian mixture modeling. First, we determined the number of distributions providing the best fit on the data with the R function “boot.comp.” Second, for each year we determined a cutpoint as the value where 2 fitted normal distributions intersected. Only a limited number of CSF samples were available for the period 2000 to 2002, and so we grouped these together to ensure reliable cutpoint estimation for these early years. We removed drift effects by aligning year-specific cutpoints, an unbiased method that keeps intact all other information present in the data. Per year, a scaling factor was determined based on the difference between the reference cutpoint obtained from the most recent period when cutpoints remained stable (2015–2016), the upper limit of abnormal values, and the lower limit of normal values. Scaling factors were added to the raw normal and abnormal Aβ 1–42 values to align year-specific cutpoints to the reference. For a subset of AD-type dementia individuals, Aβ 1–42 estimates were available that were remeasured using the same batch of ELISA kits, which removed drift effects. We used these new measurements to validate whether our approach accurately removed time dependencies from CSF data by studying whether the difference between corrected and remeasured values was related to the time that CSF was obtained. Furthermore, we compared drift-corrected values with uncorrected data in terms of concordance with visual amyloid PET read for the total cohort and per year. Concordance was defined as the percentage of individuals who were classified identically using PET and CSF methods. Finally, we compared uncorrected and corrected data in terms of sensitivity—specificity for clinical AD-type dementia vs all other individuals for the total cohort and per year. For uncorrected data, we used cutpoints we previously determined over the years using different methods: 550 pg/mL (based on clinical discrimination in 2010), 640 pg/mL (based on optimization of concordance with amyloid PET in 2014), and 680 pg/mL (based on Gaussian mixture modeling in 2017). Concordance, sensitivity, and specificity estimates were considered statistically different from each other when an estimate was not contained in the 95% CI of another estimate.

nominal variables in diagnostic group, sex, or apolipoprotein E (APOE) ε4 genotype, and time as an ordered categorical variable. Changes in CSF Aβ 1–42 concentrations over time were first determined for the total cohort with linear regression. If cohort composition (i.e., diagnostic group, age, sex, or APOE ε4 genotype) showed time dependencies, we performed additional analyses to test whether drift effects could be explained by such variables. We calculated year-specific cutpoints to determine amyloid status with Gaussian mixture modeling. First, we determined the number of distributions providing the best fit on the data with the R function “boot.comp.” Second, for each year we determined a cutpoint as the value where 2 fitted normal distributions intersected. Only a limited number of CSF samples were available for the period 2000 to 2002, and so we grouped these together to ensure reliable cutpoint estimation for these early years. We removed drift effects by aligning year-specific cutpoints, an unbiased method that keeps intact all other information present in the data. Per year, a scaling factor was determined based on the difference between the reference cutpoint obtained from the most recent period when cutpoints remained stable (2015–2016), the upper limit of abnormal values, and the lower limit of normal values. Scaling factors were added to the raw normal and abnormal Aβ 1–42 values to align year-specific cutpoints to the reference. For a subset of AD-type dementia individuals, Aβ 1–42 estimates were available that were remeasured using the same batch of ELISA kits, which removed drift effects. We used these new measurements to validate whether our approach accurately removed time dependencies from CSF data by studying whether the difference between corrected and remeasured values was related to the time that CSF was obtained. Furthermore, we compared drift-corrected values with uncorrected data in terms of concordance with visual amyloid PET read for the total cohort and per year. Concordance was defined as the percentage of individuals who were classified identically using PET and CSF methods. Finally, we compared uncorrected and corrected data in terms of sensitivity—specificity for clinical AD-type dementia vs all other individuals for the total cohort and per year. For uncorrected data, we used cutpoints we previously determined over the years using different methods: 550 pg/mL (based on clinical discrimination in 2010), 640 pg/mL (based on optimization of concordance with amyloid PET in 2014), and 680 pg/mL (based on Gaussian mixture modeling in 2017). Concordance, sensitivity, and specificity estimates were considered statistically different from each other when an estimate was not contained in the 95% CI of another estimate.
Results

Study participants characteristics

Over the past 2 decades, CSF Aβ 1–42 measurements were available for 4397 participants (Table 1 and Figure 1A). The majority of these individuals had AD-type dementia (31%), followed by SCD (18%), non-AD dementia (16%), and MCI (14%), with an additional 21% labeled as “other.” Cohort composition showed changes over time in the relative proportions of diagnostic groups (χ²(4) = 12.82; P = 0.01; Figure 1B), with the proportion of individuals with SCD increasing over time (Z = 3.23; P = 0.001), and the group “other” showing a trend to decrease (Z = -1.69; P = 0.09). Proportions of AD-type dementia, non-AD dementia, and MCI remained stable over time (all P > 0.05). Furthermore, the age of the total cohort decreased over time (β(SE) = -0.12 (0.04); P < 0.001; Figure 1C). MMSE scores, sex distributions, and the proportion of APOE ε4 carriers remained stable over the years (all P > 0.05; Figure 1, D–F).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total group (n = 4397)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years [mean (SD)]</td>
<td>63.5 (9.1)</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>1881 (43)</td>
</tr>
<tr>
<td>APOE ε4 carrier, n (%)</td>
<td>1940 (44)</td>
</tr>
<tr>
<td>MMSE, median (IQR)</td>
<td>27 (26–29)</td>
</tr>
<tr>
<td>Subject group, n (%)</td>
<td>1362 (31)</td>
</tr>
<tr>
<td>AD dementia</td>
<td></td>
</tr>
<tr>
<td>Non-AD dementia710 (14)</td>
<td></td>
</tr>
<tr>
<td>MCI</td>
<td>624 (14)</td>
</tr>
<tr>
<td>SCD</td>
<td>771 (17.5)</td>
</tr>
<tr>
<td>Other</td>
<td>929 (21)</td>
</tr>
</tbody>
</table>

Table 1. Clinical and biological description of subjects seen between 2000 and 2016 with available CSF Aβ 1–42 measurements.

IQR, interquartile range.

Data missing for 383 subjects.

Drift over time in CSF Aβ 1–42 values depends on diagnosis

The total cohort showed increases in Aβ 1–42 values of 22.07 pg/mL per year (SE = 1.13; P < 0.001). Repeating analyses additionally correcting for age did not change these results [β (SE) = 21.07 (1.09); P < 0.001], suggesting that higher Aβ 1–42 values in more recent years were not explained by an increasingly younger age of the cohort. The slope of the drift showed a dependency on diagnosis (P<0.001), with Aβ 1–42 values increasing 50% less strongly for individuals with AD-type dementia, 13 pg/mL per year, compared with increases of 25 pg/mL per year observed in individuals with SCD, 27 pg/mL in those with MCI, 23 pg/mL in participants with non-AD dementia, and 21 pg/mL in those with MCI and non-AD dementia. The results remained stable over the years (all P > 0.05; Figure 1, D–F).

Drift over time in CSF Aβ 1–42 values depends on diagnosis

The total cohort showed increases in Aβ 1–42 values of 22.07 pg/mL per year (SE = 1.13; P < 0.001). Repeating analyses additionally correcting for age did not change these results [β (SE) = 21.07 (1.09); P < 0.001], suggesting that higher Aβ 1–42 values in more recent years were not explained by an increasingly younger age of the cohort. The slope of the drift showed a dependency on diagnosis (P<0.001), with Aβ 1–42 values increasing 50% less strongly for individuals with AD-type dementia, 13 pg/mL per year, compared with increases of 25 pg/mL per year observed in individuals with SCD, 27 pg/mL in those with MCI, 23 pg/mL in participants with non-AD dementia, and 21 pg/mL in those with MCI and non-AD dementia. The results remained stable over the years (all P > 0.05; Figure 1, D–F).
Correcting the upward drift in CSF Aβ 1–42 results

with non-AD dementia, and 25 pg/mL in individuals with an "other" diagnosis (Figure 2A). The specificity of slope differences for AD-type dementia suggests that the rate of annual increases depended on amyloid concentrations, and this explanation was supported by exploratory analyses on data stratified based on the median, which showed a similar difference in slope [low concentrations: β (SE) = 9.30 (0.57), P < 0.001; high concentrations: β (SE) = 14.13 (1.14), P < 0.001].

Bimodal distribution of CSF Aβ 1–42 concentrations can be identified independently of absolute concentrations

Bimodal distributions provided the most optimal fit for each year of CSF Aβ 1–42 concentrations (all P < 0.05). Year-specific cutpoints increased from 527 pg/mL for samples analyzed in 2002 to 807 pg/mL for samples analyzed in 2016, corresponding to annual increases of 19 pg/mL (P < 0.001; Figure 2B). Using these year-specific cutpoints, we labeled 2156 individuals as having abnormal amyloid [1226 (57%) with AD-type dementia, 343 (16%) with MCI, 248 (11%) with non-AD dementia, 215 (10%) "other," and 124 (6%) with SCD]. Using these year-specific cutpoints as a reference, we compared the proportion of mislabeling of amyloid status when uncorrected data with a single cutpoint are used. At most, 680 (15%) individuals were mislabeled when using a cutpoint of 550 pg/mL, and at least 282 (6%) individuals were mislabeled when using a cutpoint of 680 pg/mL compared with year-specific cutpoints (Table 2 and Figure 2C).

Removing drift from CSF Aβ 1–42 data and determining a uniform cutpoint

Our approach of adding year-specific constant values to the data such that cutpoints based on bimodal distributions of amyloid concentrations aligned over time removed the drift effect [β (SE) = 0.07 (0.59); P = 0.91; Figure 2, D and E]. Gaussian mixture modeling on drift-corrected data of the total cohort resulted in a cutpoint of 813 (95% CI, 761–859).

Validation 1: Correspondence with remeasured CSF Aβ 1–42 concentrations

We compared drift-corrected Aβ 1–42 concentrations with remeasured CSF Aβ 1–42 concentrations that were available for a subset of 148 individuals with AD-type dementia (initially measured between 2000 and 2014 and then remeasured in 2015) 11. As previously reported, the absolute differences between uncorrected and remeasured Aβ 1–42 values increased over time [β (SE) = 16.25 (2.27); P < 0.001]. In contrast, absolute differences between drift-corrected and remeasured Aβ 1–42 concentrations were unrelated with time (β (SE) = 0.33 (2.11); P = 0.88), which supports our approach successfully removing the drift.

4

Bimodal distribution of CSF Aβ 1–42 concentrations can be identified independently of absolute concentrations

Bimodal distributions provided the most optimal fit for each year of CSF Aβ 1–42 concentrations (all P < 0.05). Year-specific cutpoints increased from 527 pg/mL for samples analyzed in 2002 to 807 pg/mL for samples analyzed in 2016, corresponding to annual increases of 19 pg/mL (P < 0.001; Figure 2B). Using these year-specific cutpoints, we labeled 2156 individuals as having abnormal amyloid [1226 (57%) with AD-type dementia, 343 (16%) with MCI, 248 (11%) with non-AD dementia, 215 (10%) "other," and 124 (6%) with SCD]. Using these year-specific cutpoints as a reference, we compared the proportion of mislabeling of amyloid status when uncorrected data with a single cutpoint are used. At most, 680 (15%) individuals were mislabeled when using a cutpoint of 550 pg/mL, and at least 282 (6%) individuals were mislabeled when using a cutpoint of 680 pg/mL compared with year-specific cutpoints (Table 2 and Figure 2C).

Removing drift from CSF Aβ 1–42 data and determining a uniform cutpoint

Our approach of adding year-specific constant values to the data such that cutpoints based on bimodal distributions of amyloid concentrations aligned over time removed the drift effect [β (SE) = 0.07 (0.59); P = 0.91; Figure 2, D and E]. Gaussian mixture modeling on drift-corrected data of the total cohort resulted in a cutpoint of 813 (95% CI, 761–859).

Validation 1: Correspondence with remeasured Aβ 1–42 concentrations

We compared drift-corrected Aβ 1–42 concentrations with remeasured CSF Aβ 1–42 concentrations that were available for a subset of 148 individuals with AD-type dementia (initially measured between 2000 and 2014 and then remeasured in 2015) 11. As previously reported, the absolute differences between uncorrected and remeasured Aβ 1–42 values increased over time [β (SE) = 16.25 (2.27); P < 0.001]. In contrast, absolute differences between drift-corrected and remeasured Aβ 1–42 concentrations were unrelated with time (β (SE) = 0.33 (2.11); P = 0.88), which supports our approach successfully removing the drift.
Validation 2: Concordance with amyloid PET

A subset of 602 (14%) participants had a visual reading of amyloid PET available. Compared with the total cohort, this subset had 10% more individuals with AD-type dementia and 10% fewer individuals classified as “other” (see Table 1 in the Data Supplement). Across this subset, drift-corrected CSF Aβ 1–42 concentrations showed the highest concordance of 88% (95% CI, 85–90) with amyloid PET readings, which was higher than concordance estimates of uncorrected CSF Aβ 1–42 values with amyloid PET readings (Table 2). When comparing concordance of amyloid PET readings with uncorrected and drift-corrected data per year, drift-corrected data showed consistently high concordance with amyloid PET compared with uncorrected data for the period between 2006 and 2016 (Figure 3).

Figure 1. Cohort characteristics plotted according to year of first visit. Number of individuals (A); relative proportions of diagnostic groups (B); age (C); MMSE scores [data were missing for 179 (4%) individuals] (D); sex distribution (E); and proportion of APOE ε4 carriers [data missing for 383 (9%) individuals] (F).
Correcting the upward drift in CSF Aβ 1-42 results

Figure 2. Dashed horizontal line throughout the figure shows the uncorrected cutpoint 680 pg/mL unless specified otherwise. Aβ 1–42 concentrations increase over time (A). Uncorrected Aβ 1–42 concentrations with individuals colored blue when having abnormal amyloid or orange when having normal amyloid (B). Year-specific cutpoints determined with mixture modeling increase over time (C). The dotted lines reflect the 95% CIs of the cutpoints. Drift-corrected Aβ 1–42 concentrations with cutpoint of 813 pg/mL (D). Drift-corrected year-specific cutpoints (E). The dotted lines reflect the 95% CIs of the cutpoints.
Table 2. Comparison of uncorrected and drift-corrected values for different cutpoints in terms of labeling individuals according to amyloid abnormality, and sensitivity–specificity analyses to detect AD dementia.

<table>
<thead>
<tr>
<th>Cutpoint</th>
<th>False positive, n (%)</th>
<th>False negative, n (%)</th>
<th>Concordance %</th>
<th>Sensitivity (95% CI)</th>
<th>Sensitivity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>550 pg/mL</td>
<td>4 (0.1)</td>
<td>676 (15)</td>
<td>70 (66–74)</td>
<td>0.66 (0.64–0.69)</td>
<td>0.81 (0.79–0.82)</td>
</tr>
<tr>
<td>640 pg/mL</td>
<td>64 (1.5)</td>
<td>225 (5)</td>
<td>81 (78–84)</td>
<td>0.85 (0.83–0.86)</td>
<td>0.72 (0.71–0.74)</td>
</tr>
<tr>
<td>680 pg/mL</td>
<td>149 (3)</td>
<td>133 (3)</td>
<td>84 (81–87)</td>
<td>0.89 (0.87–0.91)</td>
<td>0.68 (0.67–0.70)</td>
</tr>
<tr>
<td>813 pg/mL</td>
<td>Reference</td>
<td>Reference</td>
<td>88 (85–90)</td>
<td>0.90 (0.89–0.92)</td>
<td>0.69 (0.67–0.70)</td>
</tr>
</tbody>
</table>

PET data were available for the period 2006 to 2016.

*From Mulder et al. 
*From Zwan et al. 
*Significantly different from uncorrected values with cutpoints 550 pg/mL and 640 pg/mL.
*From Bertens et al. 
*Significantly different from 550 pg/mL and 640 pg/mL.

**Significantly different from 550 pg/mL, 640 pg/mL, and 680 pg/mL.

**Significantly different from 550 pg/mL, 640 pg/mL, and 680 pg/mL.

Validation 3: Clinical outcome

Across the total group, sensitivity and specificity estimates for classifying individuals as having AD-type dementia vs all other participants were largely comparable between corrected Aβ 1–42 values with a cutpoint of 813 pg/mL and uncorrected Aβ 1–42 values with a cutpoint of 680 pg/mL (Table 2). Compared with uncorrected data with other cutpoints (550 and 640 pg/mL), corrected Aβ 1–42 values showed better sensitivity to detect clinical AD dementia at the cost of lower specificity. When comparing sensitivity and specificity estimates between uncorrected and drift-corrected data per year, the largest differences in sensitivity and specificity were observed in the most recent years when using a cutpoint of 550 pg/mL (Figure 3). Drift-corrected values also showed improved sensitivity to detect AD-type dementia for the period 2014 to 2016 compared with uncorrected data with 640 and 680 pg/mL as cutpoints. In this period, specificity for drift-corrected values was worse compared with uncorrected data for all cutpoints. We further explored whether lower specificity observed in this period for drift-corrected values reflected an improved detection of predementia AD. To this end, we compared the percentage of MCI and SCD participants labeled as having abnormal amyloid determined based on corrected and uncorrected CSF Aβ 1–42 concentrations with previously reported prevalence estimates. Using uncorrected CSF Aβ 1–42 concentrations resulted in a lower than expected percentage of individuals with abnormal amyloid, regardless of the cutpoint used.
Correcting the upward drift in CSF Aβ 1–42 results (Table 3), whereas amyloid status determined with corrected CSF values resulted in percentages of MCI and SCD participants with abnormal amyloid that were in line with expected prevalences. Figure 3. Annual concordance (top), sensitivity (middle), and specificity (bottom) estimates for corrected and uncorrected Aβ 1–42 values with different cutpoints (c1 = drift corrected; c2 = uncorrected 550 pg/mL; c3 = uncorrected 640 pg/mL; c4 = uncorrected 680 pg/mL). PET data were available from 2006 onward; owing to small sample sizes, years 2006 to 2008 were taken together to obtain reliable 95% CI. *P < 0.05 c2 and all other cutpoints; aP < 0.05 c2 and c4; bP < 0.05 c3 and c1; cP < 0.05 c3 and c4; dP < 0.05 c3 and c1, c4; eP < 0.05 c4 and c1.
Table 3. Percentage of SCD and MCI participants with abnormal amyloid as determined with corrected and uncorrected CSF amyloid values for different cutpoints and for the period 2014 to 2016.

<table>
<thead>
<tr>
<th>Cognitive status</th>
<th>Age, years [mean (SD)]</th>
<th>Reference</th>
<th>Observed</th>
<th>Corrected</th>
<th>Uncorrected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>813 pg/mL</td>
<td>550 pg/mL</td>
<td>640 pg/mL</td>
</tr>
<tr>
<td>MCI (n = 142)</td>
<td>65 (8)</td>
<td>42.8 (38.7–47.1)</td>
<td>44.4</td>
<td>13.4*</td>
<td>22.5*</td>
</tr>
<tr>
<td>SCD (n = 207)</td>
<td>59 (8)</td>
<td>17.4 (11.6–25.2)</td>
<td>17.9</td>
<td>3.9*</td>
<td>8.7*</td>
</tr>
</tbody>
</table>

*From Jansen et al. 15.
*From Mulder et al. 16.
*From Zwan et al. 17.
*Based on Bertens et al. 18.
*Significantly different from reference prevalence.

Discussion

Over a period of almost 2 decades, we observed a drift in CSF Aβ 1–42 concentrations, with an average annual increase of roughly 20 pg/mL across the total cohort. Cutpoints to determine amyloid abnormality increased with a similar rate from 527 pg/mL in 2002 up to 807 pg/mL in 2016. During this period, the composition of the cohort showed some changes in the relative proportions of diagnosis and younger age over time. These changes in cohort composition, however, did not explain the increase in Aβ 1–42 concentrations, suggesting that the drift is mainly caused by variability in ELISA kits over the years. Drift effects were less strong for individuals with AD-type dementia compared with other clinical diagnoses, indicating that the increase in Aβ 1–42 values depended on Aβ 1–42 concentrations. We demonstrated and validated that drift effects can be removed from the data by a simple scaling factor derived from year-dependent cutpoints for amyloid abnormality. This resulted in a uniform cutpoint value and avoided potential removal of patient information that might be relevant for disease and/or related to amyloid levels.

ELISAs detect soluble Aβ 1–42, both in monomers and as complex with other proteins or oligomers. Many studies have demonstrated that CSF Aβ 1–42 concentrations are reduced in individuals with AD, and this has resulted in the incorporation of CSF Aβ 1–42 in research criteria for AD. 1–8 Yet, precise causes for reduced Aβ 1–42 concentrations continue to be uncertain. Postmortem studies comparing amyloid plaque load in the brain support the concept that reduced concentrations in Aβ 1–42 reflect aggregation into plaques. 2 However, it is unclear what explains the drift over time in Aβ 1–42 concentrations as measured with ELISA, and why the magnitude of the drift depends on Aβ 1–42 concentrations. If Aβ 1–42 concentrations
Correcting the upward drift in CSF Aβ 1-42 results

reflect only a sink of soluble Aβ 1–42 into plaques, this leaves the question unexplained as to why drift rates would depend on whether amyloid levels are within normal or pathological ranges. The difference in drift rates we observed suggests that CSF samples contain additional differences in their composition that are associated with the presence of abnormal low or normal high Aβ 1–42 concentrations. Previous studies have demonstrated that CSF of AD-type dementia patients contains large Aβ 1–42 aggregates 26 and soluble oligomers 27–29 (although the latter are more difficult to detect) 30–32. A recent study showed that monomeric Aβ 1–42 concentrations can be recovered when CSF samples containing Aβ 1–42 aggregates are treated with a chemical to dissolve the aggregates 33. Possibly, during the past 17 years, changes in ELISA kits and/or calibration data have occurred that might be influenced by the presence of Aβ 1–42 aggregates. Although the precise causes of increasing Aβ 1–42 concentrations over time remain to be determined, our results implicate that when correcting this drift from historic data, Aβ 1–42 concentrations need to be considered because these may influence the strength of the drift.

Compared with year-specific cutpoints for amyloid abnormality, the use of a single threshold on uncorrected data had minimal impact on the total cohort, with misclassification of amyloid abnormality for 6% to 15% of the individuals, depending on the cutpoint used. Impact in terms of classification for the clinical diagnosis of AD-type dementia was minimal for uncorrected data when a cutpoint of 680 pg/mL was used, showing similar sensitivity and specificity estimates in the total group of 4397 patients compared with corrected data with a cutpoint of 813 pg/mL.

For the most recent period between 2014 and 2016, uncorrected Aβ 1–42 concentrations for all cutpoints showed better specificity; however, this was associated with an underestimation of amyloid abnormality in those with MCI and SCD, whereas estimates based on corrected values were in line with those reported in the literature 34, 35. Still, our results imply that when studying historic data for individuals from populations that less frequently show pathological amyloid, it is important to have a cutpoint that considers the drift to avoid selecting false-positive cases (i.e., participants selected from an earlier period) and false-negative cases (i.e., participants selected from more recent periods). Here we made use of the observation that the bimodal distributions of Aβ 1–42 concentrations are independent of absolute values of Aβ 1–42, and so by realigning these distributions, we could derive in an unbiased way a single cutpoint that can be used for all data.

A potential limitation of the present study is that we do not have pathological information available for all individuals, and so the possibility that our drift correction might still lead to misclassification of amyloid abnormality cannot be excluded. However, the validation analyses we performed support the idea that our approach effectively removed time effects when comparing with remeasured Aβ 1–42 values. Moreover, the latter values demonstrated
improved concordance with amyloid PET compared with uncorrected data. A strong aspect of our approach is that we determined cutpoints for amyloid abnormality using a Gaussian mixture modeling approach that does not rely on clinical outcome and so is unbiased. With this approach, we avoided introducing dependencies in CSF data on clinical diagnosis. Another advantage of our approach is that it does not require regression modeling, which might remove too much information from amyloid concentrations that may potentially be related to time-dependent changes of (possibly yet to be discovered) factors in the cohort. An alternative, more costly approach would be to remeasure Aβ 1–42 for each novel batch. Systematic differences also exist across different assays to assess Aβ 1–42 in CSF; for this reason, ongoing international initiatives are developing novel measurement methods to reduce such variability with the aim to determine a universal cutpoint. Our approach provides the opportunity to standardize Aβ 1–42 concentrations across platforms and cohorts, which may enable identifying a universal cutpoint until better measurement methods exist. Once improved measurement methods have arrived, our approach will enable preserving the use of large, valuable historic data sets by calibrating these to new data.

In conclusion, we observed increasing Aβ 1–42 CSF concentrations over time owing to changes in ELISA kits. These increases were doubled in individuals with high Aβ 1–42 CSF concentrations, suggesting that drift over time depends on whether amyloid metabolism is normal or abnormal. Drift effects can be adjusted for by realigning year-specific cutpoints for amyloid abnormality, resulting in an unbiased and uniform cutpoint for the entire period. Potentially, this approach might also be useful to align CSF data and thereby reduce variability between multiple centers.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:

Employment or Leadership: None declared.
Consultant or Advisory Role: C.E. Teunissen, Fujirebio.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: The VUmc Alzheimer Center is part of the neurodegeneration research program of Neuroscience Amsterdam and is supported by Stichting Alzheimer Nederland and Stichting VUmc Fonds. The clinical database structure was developed with funding from Stichting Diorapthe. B.M. Tijms, the Memorabel grant program of ZonMW (number 733050506).
Expert Testimony: None declared.

Patents: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or final approval of manuscript.


Correcting the upward drift in CSF Aβ 1-42 results

Data supplement

Supplementary table 1. Clinical and biological description of subjects with both CSF and amyloid PET available.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Sample with both CSF and amyloid PET available (n=602)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD)</td>
<td>63.3 (7.7)</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>244 (40%)</td>
</tr>
<tr>
<td>APOE e4 carrier, n (%)</td>
<td>288 (48%)</td>
</tr>
<tr>
<td>MMSE, median (IQR)</td>
<td>26 (22-28)</td>
</tr>
<tr>
<td>Positive amyloid PET read, n (%)</td>
<td>324 (54%)</td>
</tr>
<tr>
<td>Subject group, n (%):</td>
<td></td>
</tr>
<tr>
<td>AD dementia</td>
<td>253 (42%)</td>
</tr>
<tr>
<td>Non-AD dementia</td>
<td>70 (12%)</td>
</tr>
<tr>
<td>MCI</td>
<td>92 (15%)</td>
</tr>
<tr>
<td>SCD</td>
<td>114 (19%)</td>
</tr>
<tr>
<td>Other</td>
<td>73 (12%)</td>
</tr>
</tbody>
</table>

AD is Alzheimer’s disease, MCI is mild cognitive impairment, SCD is subjective cognitive decline. Subjects seen between 2006-2016.