Journal of Nuclear Medicine, published on February 9, 2017 as doi:10.2967/jnumed.116.187351

Optimal reference region to measure longitudinal amyloid-beta change with ^{18}F-florbetaben PET

Santiago Bullich[1], Victor L Villemagne[2], Ana M. Catafau[1], Aleksandar Jovalekic[1], Norman Koglin[1], Christopher C Rowe[2], Susan De Santi[3]

(1) Piramal Imaging GmbH, Berlin, Germany

(2) Departments of Medicine and Molecular Imaging, University of Melbourne, Austin Health, Melbourne, VIC, Australia.

(3) Piramal Pharma Inc, Boston, MA, USA

Corresponding author

Santiago Bullich
Tegeler Straße 7
13353 Berlin
Germany
Email: santi.bullich@piramal.com

First author

Santiago Bullich
Tegeler Straße 7
13353 Berlin
Germany
Email: santi.bullich@piramal.com

Word count: 4794/5000 (includes *all* text [title page, figure legends, tables, references, etc.])

Running title: Longitudinal Aβ change with FBB PET [35/40 characters]

Funding: The trial was funded by Bayer Pharma AG, Berlin (Germany), and Piramal Imaging S.A., Matran (Switzerland).

ABSTRACT

Accurate measurement of changes in amyloid-beta (Aβ) deposition over time is important in longitudinal studies, particularly in anti-Aβ therapeutic trials. To achieve this, the optimal reference region (RR) must be selected to reduce variance of Aβ positron emission tomography (PET) measurements, allowing early detection of treatment efficacy. The aim of this study was to determine the RR that allows earlier detection of subtle Aβ changes using ^{18}F-florbetaben (FBB) PET.

Methods. Forty-five patients with mild cognitive impairment (MCI) (72.69 ± 6.54 years; 29 men/16 women) who underwent up to three FBB scans were included. Baseline scans were visually classified as high (Aβ+) or low (Aβ-) amyloid. Six cortical regions were quantified using a standardized region-of-interest atlas applied to the spatially normalized gray-matter image obtained from segmentation of the baseline T1-weighted volumetric magnetic resonance imaging (MRI). Four RRs (cerebellar gray matter (CGM), whole cerebellum (WCER), pons (PONS), and subcortical white matter (SWM)) were studied. Standardized uptake value ratio (SUVR) for each RR was calculated by dividing cortex activity by RR activity, with a composite SUVR averaged over six cortical regions. SUVR increase from baseline to 1 and 2 years, and percentage Aβ deposition per year, were assessed across Aβ+ and Aβ- groups.

Results. SUVs for any RR were not significantly different over time. Percentage Aβ accumulation per year derived from composite SUVR was 0.10±1.72 (Aβ-) and 1.36±1.98 (Aβ+) (p=0.02) for CGM and 0.13±1.47 and 1.32±1.75 (p=0.01), respectively, for WCER. Compared with baseline, composite SUVR increase in Aβ+ scans was significantly larger than in Aβ- scans at 1 year (p_{CGM}=0.04; p_{WCER}=0.03) and 2 years (p_{CGM}=0.02, p_{WCER}=0.01) using these two RR. Significant SUVR changes using PONS as RR were detected only at 2 years ($p_{1\text{-}yr}$=0.46, $p_{2\text{-}yrs}$=0.001). SUVR using SWM as RR showed no significant differences at either follow-up ($p_{1\text{-}yr}$=0.39, $p_{2\text{-}yrs}$=0.09).

Conclusion. RR selection influences reliable early measurement of Aβ changes over time. Compared with SWM and PONS which do not fulfil the RR requirements and have limited sensitivity to detect Aβ

changes, cerebellar RRs are recommended for FBB PET as they allow earlier detection of Aβ accumulation.

[343/350 words]

Key words (3–5): amyloid beta; florbetaben PET; longitudinal; reference region;

NTRODUCTIONThe approved method for the classification of Aβ PET scans in clinical practice is based on visual assessment. However, accurate measurement of changes in Aβ deposition over time is important in longitudinal observational studies and interventional therapeutic trials of Aβ-modifying treatments for Alzheimer's disease (AD). Positron emission tomography (PET) is currently being used to monitor response to therapy in such trials (*1–3*), but longitudinal measurement of amyloid load is challenging because Aβ changes over time can be very subtle (*4,5*). For this reason, optimization of image analysis is crucial to reduce variability and allow early detection of treatment efficacy.

One means of assessing amyloid load is semi-quantitative measurement using the SUVR. SUVR measurement does, however, require normalization of PET activity in the target region to a RR to account for non-displaceable radiotracer binding. A suitable RR must have the same non-displaceable activity (free plus nonspecific binding) as the target region, to ensure it has similar blood flow characteristics as the target region and is free of Aβ (*6*). These requirements are fulfilled by CGM, except in patients with advanced-stage AD (*7*) and in some types of familial AD (*8*), in whom some Aβ may be present in the CGM. In the intended clinical population for brain Aβ imaging, however, CGM is likely to be devoid of Aβ. Furthermore, the effect of cerebellar plaques in cortical FBB SUVRs appears to be negligible, even in advanced stages of AD with high cortical Aβ load (*9*). Limitations of this RR, however, include the small size of the region and its proximity to the edge of the scanner field of view, where truncation or scatter influence can decrease the signal to noise ratio (*6,10*).

The potential issues with using the cerebellar cortex as the RR provide the rationale for testing other RRs, such as WCER, white matter or PONS, for which nonspecific tracer retention means they do not fulfill the requirements for a RR. Whole cerebellum has been used successfully to track amyloid changes in therapeutic clinical trials using ^{18}F-florbetapir (*3*). However, recent research suggests that SWM may improve the power to track longitudinal Aβ changes using ^{18}F-florbetapir (*11–15*), while PONS seems superior to cerebellar cortex for detecting change over time with ^{18}F-flutemetamol PET (*12*). These findings are supported by a recent study in which SUV stability in different RRs was compared for

different amyloid tracers (^{18}F-flutemetamol, ^{18}F-florbetapir, and FBB), across time, across clinical conditions and across cerebral Aβ status (*16*). The study concluded that the RR with largest stability of SUV can be tracer-specific, not allowing the generalization of the findings from one Aβ radiotracer to the others.

Although CGM has been extensively validated in FBB cross-sectional studies (*17–19*), little is known about its robustness in tracking longitudinal changes over time and how it compares with other RRs. The objective of the present work was to identify a RR that allows early detection of subtle Aβ changes using FBB PET.

MATERIALS AND METHODS

Participants

The study population consisted of 45 patients with MCI (mean age, 72.69 ± 6.54 years; 29 male/16 female) and has previously been described in detail *(20–22)*. This study was conducted in accordance with the Declaration of Helsinki and was approved by the Austin Health Human Research Ethics Committee. All participants (or their legal representatives) provided written informed consent to undergo brain MRI and PET scanning with FBB.

Image acquisition

Participants underwent magnetic resonance imaging scans and FBB PET scans at baseline (n=45), 1 year (n=41), and 2 years (n=36). Imaging was performed with a Philips Allegro PET camera as previously described (*22*). A 2-minute transmission scan using a rotating ^{137}Cs source was performed for attenuation correction immediately before scanning. Each participant then received intravenous FBB (average, 286 ± 19 MBq), and images were acquired between 90 and 110 minutes post injection. Images were reconstructed using a three-dimensional row-action maximum-likelihood algorithm (Philips, Cleveland, USA). A three-dimensional T1-weighted MRI was performed before the PET scan.

Image analysis

Image analysis was performed using SPM12 software (http://www.fil.ion.ucl.ac.uk/spm/doc/). Motion correction was conducted for each PET scan and an average PET image was generated. Each patient's three PET scans were then realigned to the average. The average PET image was co-registered to the baseline MRI image, and the same transformation was applied to each individual PET scan. Baseline MRI images were segmented into gray matter, white matter, and cerebrospinal fluid. Baseline MRI was normalized on the T1 template provided with SPM and the same transformation was applied to the PET images and MRI segmentation.

The standardized Automated Anatomical Labeling volume of interest template (*23*) was applied to the spatially normalized gray segmentation of the MRI to generate regions-of-interest (ROIs) in the CGM, frontal, lateral temporal, occipital, parietal, anterior cingulate and posterior cingulate. SWM and PONS ROIs were generated by applying a manually defined mask around the centrum semiovale to the spatially smoothed normalized white segmentation of the MRI with a Gaussian kernel (Full width at half maximum=10 mm). WCER ROI was generated by using a manually defined mask to the sum of the normalized gray and white segmentation of the MRI (Fig. 1).

Mean radioactivity values were obtained from each ROI without correction for partial volume effects applied to the PET data. The SUV, defined as the decay-corrected brain radioactivity concentration normalized for injected dose and body weight, was calculated for all regions. These values were then used to derive the SUVR, as the ratio of the activity in the cerebral cortical regions to the activity of the RR. Four RRs (CGM, WCER, PONS and SWM) were studied, and a composite SUVR was calculated for each RR by averaging the SUVR of six cortical regions (frontal, occipital, parietal, lateral temporal and posterior and anterior cingulate regions) (*24*).

Visual assessment

Baseline FBB PET images underwent visual assessment by five independent nuclear medicine physicians blinded to clinical data following the reading methodology previously described by Seibyl et

al. (*25*). A scan was read as positive if increased tracer uptake was visible in any of the frontal, parietal, temporal, or posterior cingulate/precuneus cortices compared to white matter. The final result of the visual assessment was based on the majority read (i.e. agreement of the majority of the five independent blinded readers).

Study hypothesis

Longitudinal measurement of Aβ is challenging, given the lack of a standard of truth. Aβ change cannot be inferred from post-mortem histopathological determination of Aβ accumulation in the brain, as performed in cross-sectional studies. The study hypothesis was that the accumulation of Aβ over the 2 year follow-up will progress in those MCI patients visually assessed as positive while no or little Aβ accumulation will be observed in MCI patients visually assessed as negative (*26*). The RR that yielded the lowest SUVR variance allowing the earliest detection of Aβ changes between Aβ positive and negative MCI patients was determined.

Statistical analysis

Statistical analysis was performed using R (http://www.R-project.org/). A p-value lower than 0.05 was considered significant.

SUV analysis. SUVs at baseline, 1 year, and 2 years were assessed using a repeated measures analysis of variance to analyze possible change over time. Relative changes in SUVs between two RRs (REF1 and REF2) were assessed by calculating the ratio of the SUVs of these two RRs ($SUVR_{RR}$= $SUVR_{REF2}$/ $SUVR_{REF1}$) at baseline, 1 year and 2 years. A linear regression model was fitted to each participant's data ($SUVR_{RR} = \alpha_0 + \alpha_1 \cdot t$), where $SUVR_{RR}$ is the standardized uptake value ratio, t is the scan time, and α_0 and α_1 are the coefficients of the model. The slope of the regression line (α_1) in the groups of participants visually assessed as positive and negative was compared to zero using t-tests ($H_0: \alpha_1 = 0$ versus $H_1: \alpha_1 \neq 0$). Rejection of the null hypothesis in these tests indicates that the two RRs

compared show different behavior over time. Due to the explorative nature of this analysis, no corrections for multiple comparisons were made.

SUVR analysis. The change in SUVR from baseline ($SUVR_{baseline}$) to 1 year ($SUVR_{1\ year}$) and 2 years ($SUVR_{2\ years}$) was calculated as follows:

$$\Delta SUVR_1 = SUVR_{1\ year} - SUVR_{baseline}$$

$$\Delta SUVR_2 = SUVR_{2\ year} - SUVR_{baseline}$$

The SUVR change between the groups of participants visually assessed as positive and negative was compared using unpaired t-tests ($H_0: \Delta SUVR_{negative} = \Delta SUVR_{positive}$ versus $H_1: \Delta SUVR_{negative} < \Delta SUVR_{positive}$). A linear regression model was fitted to each participant's data ($SUVR = \beta_0 + \beta_1 \cdot t$), where SUVR is the standardized uptake value, t is the scan time, and β_0 and β_1 are the coefficients of the model. The percentage of Aβ deposition per year ($A\beta_{dep}$) was defined as $A\beta_{dep}=100\cdot\beta_1/SUVR_{baseline}$. The percentage of Aβ deposition per year was compared in the groups of participants visually assessed as positive and negative using an unpaired t-test ($H_0: A\beta_{dep,negative} = A\beta_{dep,positive}$ versus $H_1: A\beta_{dep,negative} < A\beta_{dep,positive}$). Cohen's d was used to measure the effect size of the percentage Aβ accumulation per year between scans visually assessed as negative and positive.

RESULTS

SUV analysis

No statistically significant differences over time were detected in SUV for any of the RRs studied (p=0.53 (CGM); p=0.57 (WCER); p=0.68 (PONS);p=0.89 (SWM)) (Fig. 2). Further analysis showed that the ratio of SUV in WCER and PONS to the SUV in CGM ($SUVR_{RR}$) was stable over time indicating that WCER and PONS follow similar time courses in comparison with CGM ($p=0.78$ (SUV_{WCER} to SUV_{CGM} ratio), $p_{PONS\text{-}CGM}=0.37$ (SUV_{PONS} to SUV_{CGM} ratio)). SUV of the SWM showed, however, a statistically significant tendency to decrease with respect to the SUV of the CGM in the negatives group ($p_{SWM\text{-}}$

$_{CGM}$=0.05 (SUV_{SWM} to SUV_{CGM} ratio)) indicating that SWM did not follow similar time course as CGM for all the subjects.

SUVR change and percentage of Aβ deposition per year (composite SUVR)

Using CGM and WCER as RRs, composite SUVR increase in positive scans (n=23) was significantly larger than those in negative scans (n=18) between baseline and 1 year ((p=0.04 (CGM); p=0.03 (WCER)) and between baseline and 2 years (n=17 (Aβ-), n=19 (Aβ+); p=0.02 (CGM); p=0.01 (WCER))(n=) (Table 1; Fig. 3). Using PONS as a RR detected significant changes between positive and negative scans only at 2 years (p=0.46 (1 year); p=0.001 (2 years)), while using white matter showed no significant differences at either follow-up scan (p=0.15 (1 year); p=0.16 (2 years)). Both CGM and WCER RRs enabled early detection of cortical SUVR changes between positive and negative scans that were consistent with the hypothesized change for patients with MCI. Average percentage Aβ accumulation per year (mean ± SD) derived from composite SUVR in negative and positive scans was 0.10 ± 1.72 and 1.36 ± 1.98, respectively, using CGM as RR, and 0.13 ± 1.47 and 1.32 ± 1.75, respectively, using WCER. The effect size (Cohen's d) of the percentage Aβ accumulation per year was higher for WCER (d_{WCER}=0.73) and CGM (d_{CGM}=0.67) than for PONS (d_{PONS}=0.48) and SWM (d_{SWM}=-0.04). All RRs showed significant $A\beta_{dep}$ in the group of subjects visually assessed as positive (Table 1; Figure 3). The $A\beta_{dep}$ in the subjects visually assessed as negative was not significantly different from zero for CGM, WCER and PONS but reached statistical significance for SWM (Table 1; Fig. 3).

SUVR change and percentage of Aβ deposition per year (regional SUVR)

Regional percentage Aβ accumulation per year in positive scans was significantly larger than those in negative scans in frontal cortex, lateral temporal cortex, anterior cingulate, posterior cingulate, and parietal cortex (all p<0.05) when using WCER or CGM as the RR (Table 2; Fig. 4). Regional SUVR using WCER or CGM as RR increased significantly (p<0.05) between baseline and 1 year for the frontal cortex (WCER, p=0.043; CGM, p=0.049), anterior cingulate (p=0.043, p=0.048), and posterior cingulate (p=0.003, p=0.004) (Table 3; Fig. 5).

DISCUSSION

The results of this study suggest that the accurate selection of an appropriate RR is crucial for detection of subtle changes in longitudinal FBB PET analysis. The study also confirms the robustness of FBB quantification using CGM and WCER as RRs, as they allow earlier detection of Aβ change. These results are consistent with a previous study reporting that the highest SUV stability for FBB is achieved in the CGM (*16*), in contrast with other Aβ radioligands, for which SWM and PONS have been reported as the most accurate RRs (*11–15*) . This emphasizes the probable ligand-specific nature of the optimal RR.

Although the rate of Aβ deposition depends on the disease stage (*26*), the accumulation of Aβ per year measured in the present study using FBB (CGM, 1.36 ± 1.98%; WCER, 1.32 ± 1.75%) is similar to previously reported results for other Aβ radioligands. In a study with ^{18}F-florbetapir, reported deposition was 1.1 ± 4.9% for CGM and 1.1 ± 4.0 for WCER in a sample of early MCI (*13*), with an overall range for ^{18}F-florbetapir of 1–2% reported in a second study (*10*). Deposition in a study with MCI subjects with ^{18}F-flutemetamol was found to be 1.6 ± 3.3% (*12*).

Each of the RRs identified in our analysis has several advantages and disadvantages. CGM fulfils all the requirements for a RR, in that it is free of Aβ and has similar non-displaceable activity (free + nonspecific binding) and blood flow as the target region (*6*). CGM can, however, be affected by amyloid accumulation in late disease stages, but this effect is unlikely to happen in patients with MCI, such as those enrolled in the present study, or the clinically intended population for brain Aβ imaging. Indeed, the effect of cerebellar plaques in cortical FBB SUVRs appears to be negligible, even in advanced stages of AD with high cortical Aβ load (*9*). Another potential limitation of CGM, which is shared by the PONS, is that it seems to be vulnerable to scatter and truncation as a result of its position close to the edge of the scanner field of view. CGM and PONS are also small regions that tend to have a low number of counts and higher variability. Unlike CGM, the PONS, together with WCER and SWM, are mainly white matter

regions that do not fulfill the conditions for an RR. In the WCER, however, this seems to have limited impact, and WCER is a sensitive region for detection of Aβ changes. Although the use of SWM as RR showed significant Aβ accumulation in the positive group, it did not allow detection of differences between negative and positive subjects because of the SUVR increase over time in the negative group. The reason for this low sensitivity is likely to be the different time course of the SUV in SWM with respect to the SUV in CGM, WCER and PONS, possibly caused by the different pathologies involved in the positive group (i.e. MCI subjects that converted to AD) and negative group (i.e. MCI subjects that progressed to progressive supranuclear palsy, temporal lobar degeneration, amyotrophic lateral sclerosis, or dementia with Lewy bodies). These results are in concordance with those of Villemagne et al. showing that the SWM region is not stable across diagnoses and Aβ status (*16*). Additionally, SWM may be affected by white matter atrophy, and vascular lesions, which are less likely to be found in cerebellum and PONS. SWM is also likely to be affected by spillover in regards to cortical activity (*13*).

One general limitation of all longitudinal measurements of Aβ accumulation over time is the lack of a standard of truth, such as post-mortem histopathological data, which is used in cross-sectional studies. In the present study, Aβ accumulation measured with FBB PET was compared between patients with MCI expected to remain stable over time (FBB negative at baseline) and those expected to increase Aβ accumulation (FBB positive at baseline). While some FBB-negative patients may also accumulate Aβ and some positive patients may have reached a plateau and may not increase Aβ deposition further, it is unlikely that this biases our results in favor of a particular RR (*27*). Another potential issue is that ROIs were based on each participant's baseline structural MRI scan. This may introduce the possibility that atrophy may have occurred, which could in turn result in over- or underestimation of change. The analysis described in this manuscript was replicated in a subsample (n=36) in which at least one follow-up MRI was available. The SUVR obtained using follow-up MRIs showed an excellent correlation with SUVRs using only the baseline MRI (R^2>0.988; composite region across RRs) and the conclusions did not differ, indicating the small impact of atrophy in this sample. Impact of atrophy and partial-volume correction is

an important area of future investigation that may be important in some populations but may increase measurement variability (*13,28*).

As noted above, the findings of this analysis are specific to FBB SUVR, and should therefore not be generalized to other amyloid PET tracers, or to other analysis methods (e.g. dynamic acquisitions and tracer kinetics). Dynamic acquisition may further reduce the variability, as shown for ^{11}C-Pittsburgh compound B (*29*), but it is difficult to apply in routine clinical practice and beyond the scope of this study. Other sources of variability such as physiologic changes (e.g. blood flow), consistency of acquisition conditions (e.g. scan start time), reconstruction and tracer quantification methods (e.g. distribution volume ratio, SUVR) that are also beyond the scope of this work, do warrant further investigation.

CONCLUSION

RR selection influences the reliable and early measurement of Aβ changes over time. Compared with SWM or PONS which do not fulfil the RR requirements and have limited sensitivity to detect Aβ changes, cerebellar RRs (CGM and WCER) are recommended for FBB PET as they were more sensitive to detect subtle Aβ accumulation.

DISCLOSURE

Santiago Bullich, Aleksandar Jovalekic and Norman Koglin are employees of Piramal Imaging GmbH.

Ana M Catafau was employee of Piramal Imaging GmbH until January 2016. Currently works as a freelance consultant on Clinical Molecular Imaging.

Victor L Villemagne is supported in part by NHMRC Research Fellowship 1046471 and has received speaker's honoraria from Piramal Imaging, GE Healthcare, Avid Pharmaceuticals, AstraZeneca, Hoffmann-La Roche, and consulting fees for Novartis, Lundbeck, Abbvie and Hoffmann-La Roche.

Christopher C Rowe has received research grants from Bayer Schering Pharma, Piramal Imaging, Avid Radiopharmaceuticals, Navidea, GE Healthcare, Astra Zeneca, Biogen.

Susan de Santi is an employee of Piramal Pharma Inc.

ACKNOWLEDGEMENTS

Medical writing support was provided by Dan Booth (Bioscript Medical, London, UK) and funded by Piramal Imaging S.A.

REFERENCES

1. Rinne JO, Brooks DJ, Rossor MN, *et al*. 11C-PiB PET assessment of change in fibrillar amyloid-beta load in patients with Alzheimer's disease treated with bapineuzumab: a phase 2, double-blind, placebo-controlled, ascending-dose study. *Lancet Neurol.* 2010;9:363-372.

2. Sevigny J, Suhy J, Chiao P, *et al*. Amyloid PET screening for enrichment of early-stage Alzheimer disease clinical trials: experience in a phase 1b clinical trial. *Alzheimer Dis Assoc Disord.* 2016;30:1-7.

3. Sevigny J, Chiao P, Bussiere T, *et al*. The antibody aducanumab reduces Abeta plaques in Alzheimer's disease. *Nature.* 2016;537:50-56.

4. Lopresti BJ, Klunk WE, Mathis CA, *et al*. Simplified quantification of Pittsburgh Compound B amyloid imaging PET studies: a comparative analysis. *J Nucl Med.* 2005;46:1959-1972.

5. Tolboom N, Yaqub M, Boellaard R, *et al*. Test-retest variability of quantitative [11C]PIB studies in Alzheimer's disease. *Eur J Nucl Med Mol Imaging.* 2009;36:1629-1638.

6. Schmidt ME, Chiao P, Klein G, *et al*. The influence of biological and technical factors on quantitative analysis of amyloid PET: Points to consider and recommendations for controlling variability in longitudinal data. *Alzheimers Dement.* 2015;11:1050-1068.

7. Thal DR, Rub U, Orantes M, Braak H. Phases of A beta-deposition in the human brain and its relevance for the development of AD. *Neurology.* 2002;58:1791-1800.

8. Knight WD, Kim LG, Douiri A, Frost C, Rossor MN, Fox NC. Acceleration of cortical thinning in familial Alzheimer's disease. *Neurobiol Aging.* 2011;32:1765-1773.

9. Catafau AM, Bullich S, Seibyl JP, *et al*. Cerebellar amyloid-beta plaques: How frequent are they, and do they influence 18F-Florbetaben SUVR? *J Nucl Med.* 2016;57(11):1740-1745.

10. Matthews D, Marendic B, Andrews A, *et al*. Longitudinal amyloid measurement for clinical trials: a new approach to overcome variability. Presented at 8th Human Amyloid Imaging, Miami, FL, USA (Abstract #P47), 2014.

11. Chen K, Roontiva A, Thiyyagura P, *et al*. Improved power for characterizing longitudinal amyloid-beta PET changes and evaluating amyloid-modifying treatments with a cerebral white matter reference region. *J Nucl Med.* 2015;56:560-566.

12. Rowe C, Doré V, Bourgeat P, *et al*. Longitudinal assessment of Aβ accumulation in non-demented individuals: an 18F-flutemetamol study. *J Nucl Med.* 2015;56 (Suppl 3):Abstract #193.

13. Landau SM, Fero A, Baker SL, *et al*. Measurement of longitudinal beta-amyloid change with 18F-florbetapir PET and standardized uptake value ratios. *J Nucl Med.* 2015;56:567-574.

14. Brendel M, Hogenauer M, Delker A, *et al*. Improved longitudinal [(18)F]-AV45 amyloid PET by white matter reference and VOI-based partial volume effect correction. *Neuroimage.* 2015;108:450-459.

15. Joshi AD, Pontecorvo MJ, Navitsky MA, Kennedy IA, Mintun MA, Devous Sr MD. Measuring change in beta amyloid burden over time using florbetapir-PET and a subcortical white matter reference region. *Alzheimers Dement.* 2014;10 (4 Suppl):P902 (Abstract P4-316).

16. Villemagne VL, Bourgeat P, Doré V, *et al*. Amyloid imaging in therapeutic trials: the quest for the optimal reference region. *Alzheimers Dement.* 2015;11 (7 Suppl):P21-P22 (Abstract #IC-P-016).

17. Barthel H, Bullich S, Sabri O, *et al*. 18F-Florbetaben (FBB) PET SUVR quantification: which reference region? *J Nucl Med.* 2015;56 (Suppl 3):Abstract #1563.

18. Bullich S, Catafau AM, Seibyl J, De Santi S. Classification of positive and negative 18F-Florbetaben scans: comparison of SUVR cutoff quantification and visual assessment performance. *J Nucl Med.* 2016;57:516 (Abstract).

19. De Santi S, Catafau AM, Seibyl J, Bullich S. Robustness of 18F-Florbetaben SUVR cutoff quantification across reference regions and standards of truth. *J Nucl Med.* 2016;57:458 (Abstract).

20. Ong K, Villemagne VL, Bahar-Fuchs A, *et al*. (18)F-florbetaben Abeta imaging in mild cognitive impairment. *Alzheimers Res Ther.* 2013;5:4.

21. Bahar-Fuchs A, Villemagne V, Ong K, *et al*. Prediction of amyloid-beta pathology in amnestic mild cognitive impairment with neuropsychological tests. *J Alzheimers Dis.* 2013;33:451-462.

22. Ong KT, Villemagne VL, Bahar-Fuchs A, *et al*. Abeta imaging with 18F-florbetaben in prodromal Alzheimer's disease: a prospective outcome study. *J Neurol Neurosurg Psychiatry.* 2015;86:431-436.

23. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, *et al*. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage.* 2002;15:273-289.

24. Rowe CC, Ackerman U, Browne W, *et al*. Imaging of amyloid beta in Alzheimer's disease with 18F-BAY94-9172, a novel PET tracer: proof of mechanism. *Lancet Neurol.* 2008;7:129-135.

25. Seibyl J, Catafau AM, Barthel H, *et al*. Impact of training method on the robustness of the visual assessment of 18F-florbetaben PET scans: results from a Phase 3 trial. *J Nucl Med.* 2016;57:900-906.

26. Villemagne VL, Burnham S, Bourgeat P, *et al*. Amyloid beta deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol.* 2013;12:357-367.

27. Villain N, Chetelat G, Grassiot B, *et al*. Regional dynamics of amyloid-beta deposition in healthy elderly, mild cognitive impairment and Alzheimer's disease: a voxelwise PiB-PET longitudinal study. *Brain.* 2012;135:2126-2139.

28. Rullmann M, Dukart J, Hoffmann KT, *et al*. Partial-Volume Effect Correction Improves Quantitative Analysis of 18F-Florbetaben beta-Amyloid PET Scans. *J Nucl Med.* 2016;57:198-203.

29. van Berckel BN, Ossenkoppele R, Tolboom N, *et al*. Longitudinal amyloid imaging using 11C-PiB: methodologic considerations. *J Nucl Med.* 2013;54:1570-1576.

Figure legends

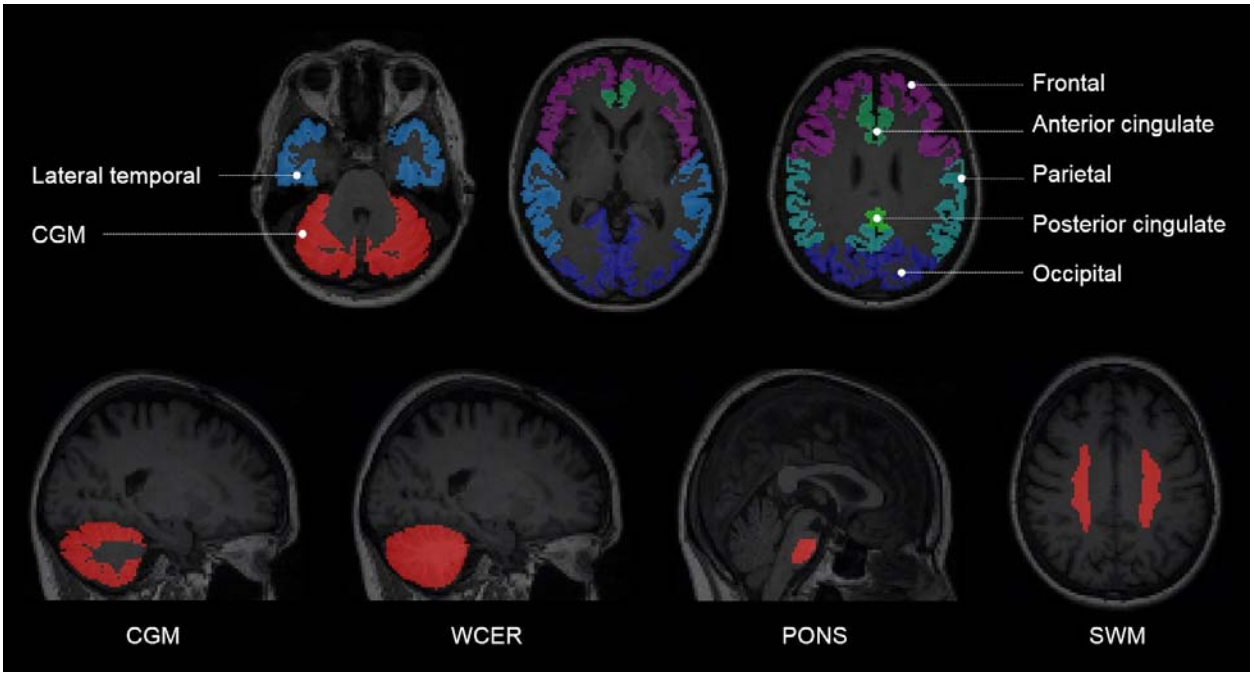

Figure 1. Example of the regions of interest generated from the segmentation of each subject's MRI.

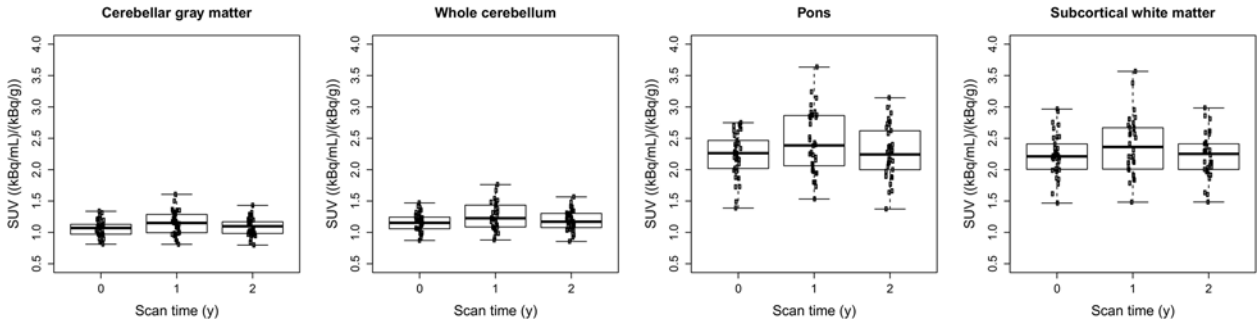

Figure 2. Standardized uptake values over time for different reference regions.

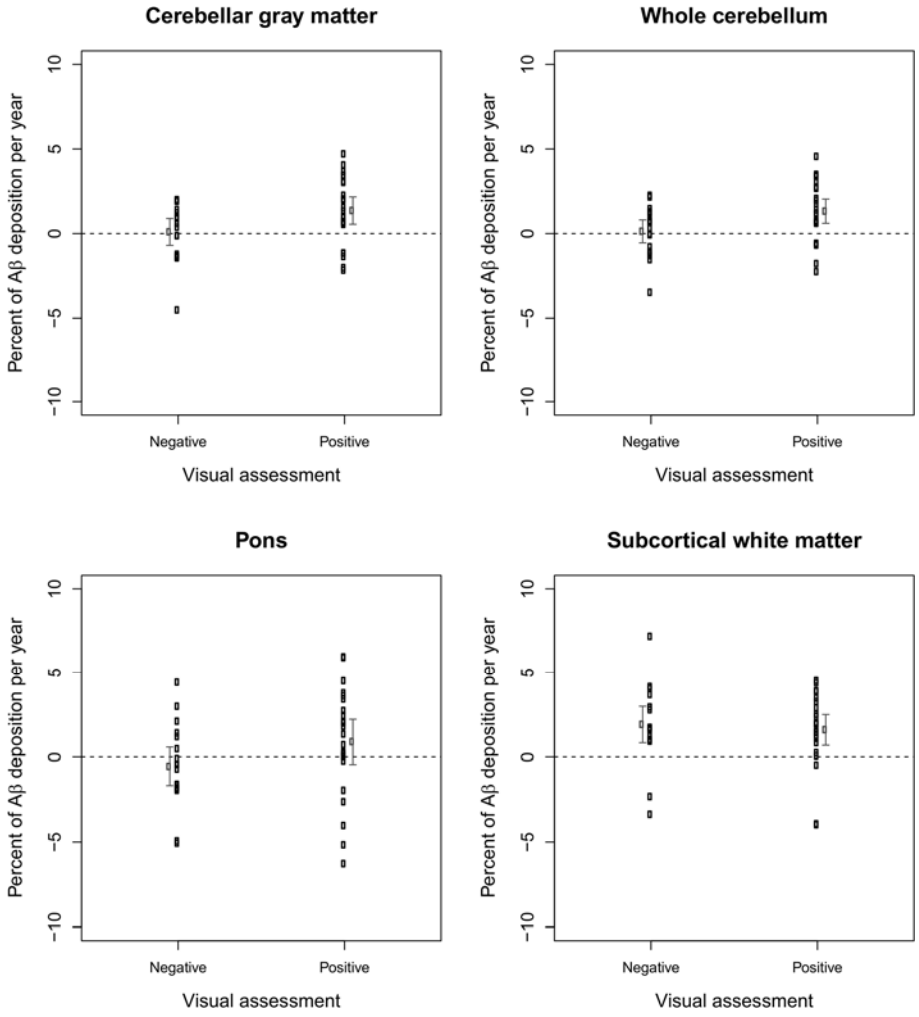

Figure 3. Percentage amyloid-beta deposition (and mean ± 95% confidence interval of the mean) per year for participants visually assessed as negative or positive measured with composite standardized uptake value ratios using four different reference regions.

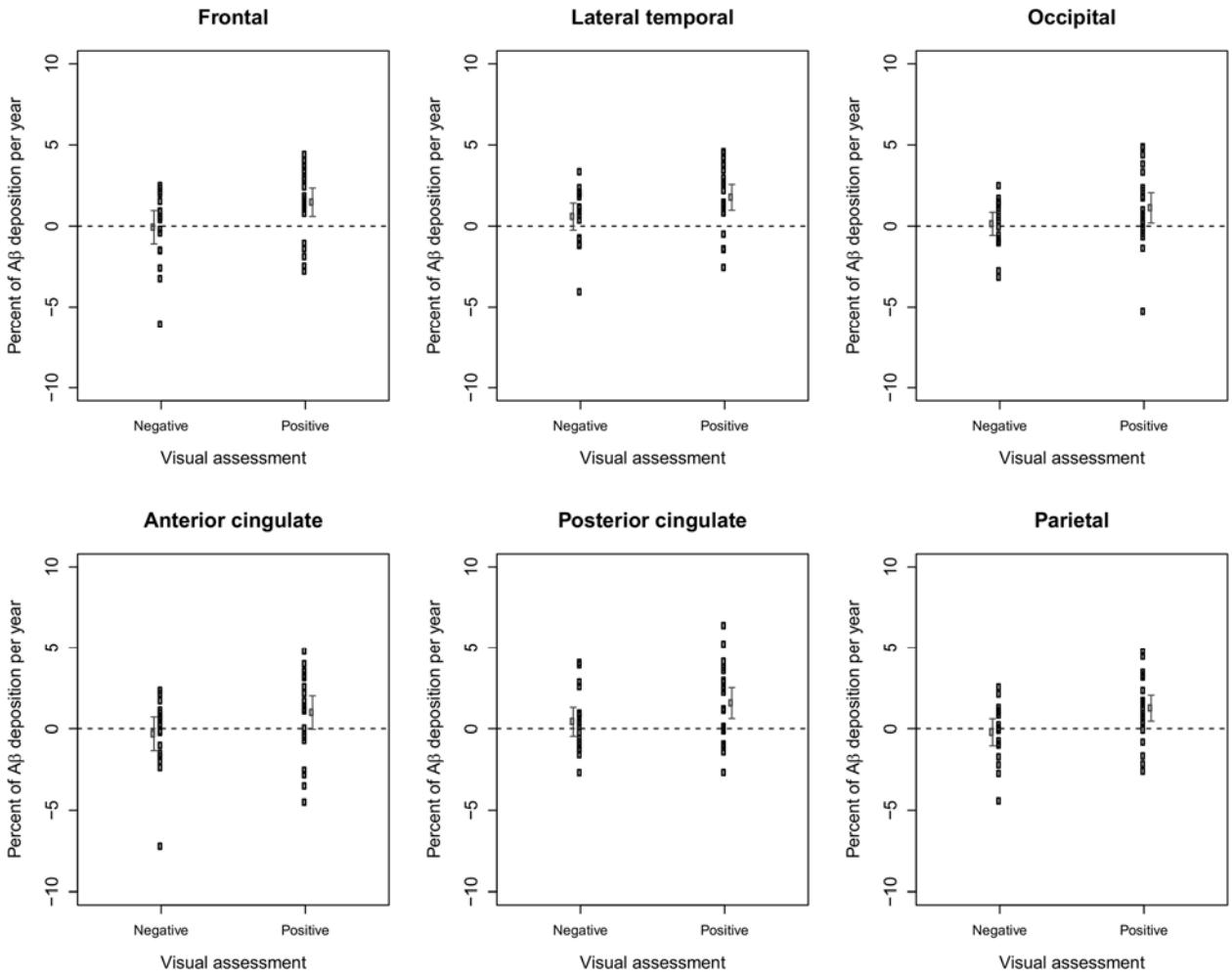

Figure 4. Percent of amyloid-beta deposition per year (and mean ± 95% confidence interval of the mean) for subjects visually assessed as negative and positive (reference region: cerebellar gray matter).

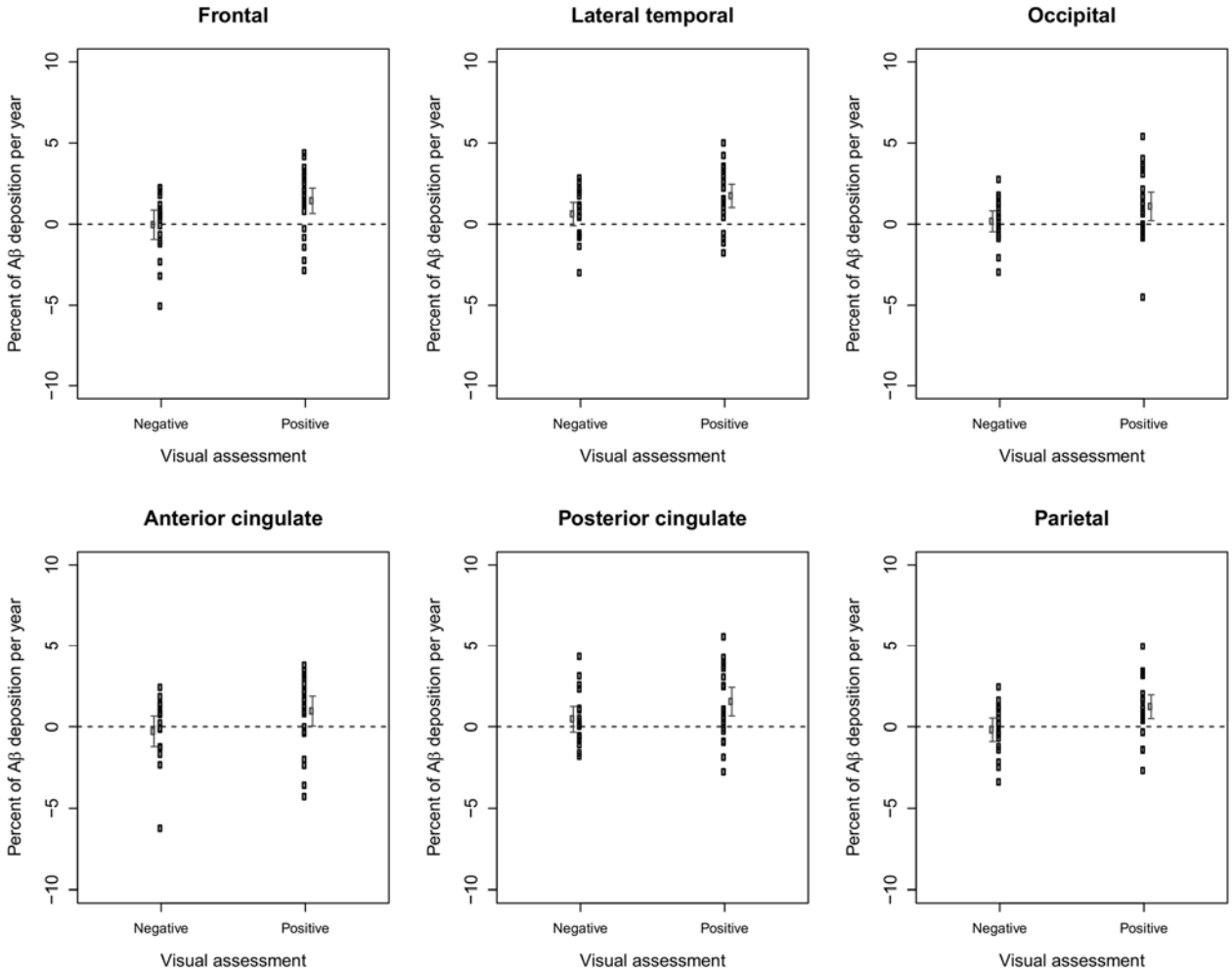

Figure 5. Percent of amyloid-beta deposition per year (and mean ± 95% confidence interval of the mean) for subjects visually assessed as negative and positive (reference region: whole cerebellum).

Table 1. Detection of amyloid-beta changes (composite standardized uptake value ratio)

Reference region	$A\beta_{dep}$ (mean ± standard deviation)		p-values (negative vs. positive) (t-test)[a]		
	Negative	**Positive**	**$A\beta_{dep}$**	**$\Delta SUVR_{1\ year}$**	**$\Delta SUVR_{2\ years}$**
Cerebellar gray matter	0.10 ± 1.72	1.36 ± 1.98[b]	**0.02**	**0.04**	**0.02**
Whole Cerebellum	0.13 ± 1.47	1.32 ± 1.75[b]	**0.01**	**0.03**	**0.01**
Pons	-0.54 ± 2.45	0.87 ± 3.29[b]	0.06	0.46	**0.00**
Subcortical white matter	2.10± 2.40[b]	2.00± 2.31[b]	0.55	0.15	0.16

$A\beta_{dep}$ = percentage Aβ deposition per year; ΔSUVR = change in standardized uptake value ratio
[a]Statistically significant p-values ($p<0.05$) shown in bold; [b]Reference region showed significant $A\beta_{dep}$ (i.e. statistically superior to zero)

Table 2. Regional amyloid-beta change detection using cerebellar gray matter as reference region.

Reference region	$A\beta_{dep}$ (mean ± standard deviation)		p-values (negative vs. positive) (t-test)[a]		
	Negative	**Positive**	**$A\beta_{dep}$**	**$\Delta SUVR_{1\ year}$**	**$\Delta SUVR_{2\ years}$**
Frontal cortex	-0.07 ± 2.22	1.47 ± 2.14	**0.02**	**0.05**	**0.03**
Lateral temporal cortex	0.59 ± 1.81	1.77 ± 1.95	**0.03**	0.07	**0.01**
Occipital cortex	0.14 ± 1.55	1.13 ± 2.29	0.05	0.15	**0.01**
Anterior cingulate cortex	-0.30 ± 2.21	0.99 ± 2.53	**0.04**	**0.05**	0.21
Posterior cingulate cortex	0.43 ± 1.90	1.57 ± 2.33	**0.05**	**0.00**	**0.04**
Parietal cortex	-0.21 ± 1.76	1.26 ± 1.96	**0.01**	0.12	**0.01**
Composite	0.10 ± 1.72	1.36 ± 1.98	**0.02**	**0.04**	**0.02**

$A\beta_{dep}$ = percentage Aβ deposition per year; ΔSUVR = change in standardized uptake value ratio
[a]Statistically significant p-values ($p<0.05$) shown in bold

Table 3. Regional amyloid-beta change detection using whole cerebellum as reference region.

Reference region	Aβdep (mean ± standard deviation)		p-values (negative vs. positive) (t-test)[a]		
	Negative	**Positive**	**$Aβ_{dep}$**	**$\Delta SUVR_{1\ year}$**	**$\Delta SUVR_{2\ years}$**
Frontal cortex	-0.04 ± 1.96	1.44 ± 1.92	**0.01**	**0.04**	**0.02**
Lateral temporal cortex	0.62 ± 1.56	1.74 ± 1.76	**0.02**	0.07	**0.00**
Occipital cortex	0.17 ± 1.39	1.09 ± 2.15	0.05	0.16	**0.01**
Anterior cingulate cortex	-0.27 ± 2.00	0.95 ± 2.25	**0.04**	**0.04**	0.18
Posterior cingulate cortex	0.47 ± 1.68	1.54 ± 2.16	**0.04**	**0.00**	**0.03**
Parietal cortex	-0.18 ± 1.53	1.23 ± 1.79	**0.01**	0.12	**0.01**
Composite	0.13 ± 1.47	1.32 ± 1.75	**0.01**	**0.03**	**0.01**

$Aβ_{dep}$ = percentage Aβ deposition per year; ΔSUVR = change in standardized uptake value ratio
[a]Statistically significant p-values ($p<0.05$) shown in bold