Flortaucipir F 18 Quantitation using a Parametric Estimate of Reference Signal Intensity

(PERSI)

Sudeepti Southekal, Michael D. Devous, Sr., Ian Kennedy, Michael Navitsky, Ming Lu, Abhinay D. Joshi, Michael J. Pontecorvo, Mark A. Mintun

Avid Radiopharmaceuticals, Inc. 3711 Market St Suite 710, Philadelphia PA 19104, USA.

Corresponding Author:

Sudeepti Southekal

Avid Radiopharmaceuticals

3711 Market St, Philadelphia, PA 19104

E-mail: southekal@avidrp.com

215-298-0700

Running Title: Flortaucipir quantitation using PERSI

Keywords: Flortaucipir, AV1451, T807, SUVr, reference region.

Abstract

Introduction: PET imaging of tau pathology in Alzheimer's disease may benefit from the use of white matter reference regions. These regions have shown reduced variability compared to conventional cerebellar regions in amyloid imaging. However, they are susceptible to contamination from partial-volume blurring of tracer uptake in cortex. We present a new technique (PERSI) for Flortaucipir F 18 count normalization that leverages the advantages of white matter reference regions while mitigating potential partial-volume effects.

Methods: Subjects with clinical diagnoses of Alzheimer's Disease (AD), mild cognitive impairment (MCI) or normal cognition (CN) underwent T1 MRI and florbetapir imaging (to determine amyloid (Aβ) status) at screening, and flortaucipir imaging at single or multiple time points. Flortaucipir images acquired as 4x5 minute frames, 80 minutes after a 370 MBq injection, were motion corrected, averaged and transformed to MNI space. The PERSI reference region was calculated for each scan by fitting a bimodal Gaussian distribution to the voxel-intensity histogram within an atlas-based white matter region, and using the center and width of the lower-intensity peak to identify the voxel intensities to be included. Four conventional reference regions were also evaluated: 1) whole cerebellum, 2) cerebellar gray matter, 3) atlas-based white matter, and 4) subject-specific white matter. SUVr was calculated for a statistically-defined neocortical volume-of-interest (MUBADA). Performance was evaluated with respect to test-retest variability in a Phase 2 study of 21 subjects (5-34 days between scans). Baseline variability in controls (SD of SUVr and ΔSUVr values) and effect sizes for group differences (Cohen's d; Aβ+ impaired vs. Aβ-normal) were evaluated in another Phase 2 study with cross-sectional (n=215) and longitudinal (n=142/215; 18±2 months between scans) data.

Results: PERSI showed superior test-retest reproducibility (1.84%) and group separation ability (cross-sectional Cohen's d=9.45; longitudinal Cohen's d=2.34) compared to other reference regions. Baseline SUVr variability and ΔSUVr were minimal in Aβ- control subjects with no specific flortaucipir uptake (SUVr 1.0±0.04, ΔSUVr 0.0±0.02).

Conclusions: PERSI reduced variability while enhancing discrimination between diagnostic cohorts. Such improvements could lead to more accurate disease staging and robust measurements of changes in tau burden over time for the evaluation of putative therapeutic treatments.

Introduction

The growing acceptance of molecular imaging biomarkers in Alzheimer's Disease (AD) research and clinical trials emphasizes the need for reliable and reproducible image quantification methods. Used independently or as an adjunct to visual interpretation by trained physicians, quantification in brain Positron Emission Tomography (PET) involves the determination of target-specific biomarker retention in the presence of confounds from physiological factors (e.g., blood flow and delivery, radiopharmaceutical dose, route of administration, body weight, etc.), technical factors (e.g., scanner model and acquisition parameters) and diffuse or focal non-specific binding to brain structures that may differ between subjects and over time. In clinical trials enrolling large numbers of subjects, ideal quantification via tracer kinetic analysis (*1,2*), including noninvasive methods which obviate arterial blood sampling (*3,4*) can be prohibitively time consuming and complicated. Therefore, an alternative approach is widely used that involves calculating the ratio of the average activity concentration in the target volume-of-interest (VOI) to that in a "reference region", i.e. a tissue region devoid of specific binding of the radiotracer. The reference region serves to achieve intra- and inter-subject count normalization, generating an intercomparable measure of Standardized Uptake Values (SUV; (*5*)), i.e. SUVr. A number of studies (*6–8*) have validated the results from semi-quantitative methods against kinetic models for AD biomarkers.

Reference Regions for Amyloid and Tau Imaging

Reliable reference regions can be critically important to PET investigations. Errors in the mean activity concentration of the reference region directly translate to variability in the SUVr, and could affect the power of a trial to detect meaningful signals, or subtle longitudinal changes in such signals, over the noise in the measurement.

Conventionally, tissue regions that are empirically known to contain a negligible amount of specific binding sites are used as reference regions. The cerebellar cortex, which is expected to remain free of fibrillar amyloid over time, was selected for amyloid imaging with ^{11}C-PiB PET (*6*) since PiB clearance from cerebellar and cerebral gray matter (GM) were observed to be similar, and because of the stability of this region over time. Cerebellar GM regions are susceptible to noise due to their small size, low signal level, and proximity to the axial periphery of the PET scanner bore (where the scanner has a lower sensitivity to detect signal). With the goal of ameliorating some of these concerns, the whole cerebellum (including cerebellar white matter) was validated for Florbetapir F 18 in an autopsy study (*9*) and used to determine an amyloid (florbetapir) positivity threshold (*10*). Whole cerebellum is now routinely used for florbetapir cross-sectional imaging (*11*), and is also currently recommended by the "Centiloid" project (*12*) for the quantitation of amyloid burden with PiB. It has been noticed, however, that the location of the cerebellum within the ~15-cm long axial PET field-of-view (FOV) can make it vulnerable to truncation. For tau imaging, focal areas of uptake in the cerebellum (e.g., dentate nuclei) may also present a challenge.

The use of a subcortical white matter (WM) reference region (centrum semiovale) for longitudinal studies with florbetapir has been investigated by several groups (*13–15*) . This region has the advantage of shared axial location with the cortex, and is typically larger (hence less sensitive to noise) than cerebellar regions. WM regions showed lower variability compared to cerebellar cortex, especially for longitudinal data. To take advantage of this reduced variability, yet retain the merits of autopsy-validated SUVR values referenced to cerebellum, a 2-step normalization process

was proposed for longitudinal studies (*16*). Data are first normalized to whole cerebellum, then scaled by the ratio of subcortical WM for follow up studies to the subcortical WM of the first visit.

A phenomenon known as Partial Volume Effect (PVE) presents an additional challenge for fixed anatomical regions. PVE refers to blurring caused by the limited spatial resolution of PET scanners (on the order of 4 – 6 mm (*17*)) which leads to apparent "spillover" or cross-contamination of counts between adjacent structures. If counts from areas with specific tracer binding spill in to the reference region, SUVr will be underestimated. This is of particular concern for WM reference regions, due to their proximity to cortical GM (containing specific tracer uptake in positive scans). The fraction of counts spilling in to the reference region is dependent on the degree of tracer binding (cortical SUVr) and the size of the structure with specific uptake (*18*) (*19*). Since both cortical atrophy as well as uptake levels could change over the course of time, a WM reference region could be differentially impacted by PVE at different time points, thus obfuscating longitudinal measures.

In this paper, we introduce and evaluate a novel technique called PERSI (Parametric Estimation of Reference Signal Intensity) to identify a subject-specific WM reference region for tau biomarker flortaucipir that addresses the above concerns associated with traditional reference regions.

PERSI

PERSI is a subject-specific, data-driven technique that identifies voxels with apparent non-specific flortaucipir uptake within a white matter region, based on the signal intensity histogram of the region. The underlying assumption of the approach is that the white matter in a flortaucipir PET image is composed of the following categories of voxels in differing proportions: 1) voxels with

negligible specific binding; 2) voxels with contamination (or spillover) of counts from adjacent cortical tissues with specific uptake or regions with off-target binding; and 3) rare occurrences of focal WM areas with elevated binding. Of these, the voxels in category #1 are considered to be appropriate for count normalization across subjects and scans. The goal of PERSI is to identify voxels in category 1, which will generally be the majority of voxels within the white matter intensity histogram and are expected to have a lower intensity than voxels from other categories within that histogram. This is achieved by fitting the histogram to a bimodal Gaussian distribution (i.e. the sum of two Gaussian distributions). The higher-intensity peak captures spill-in from tissues with elevated flortaucipir uptake, while the lower-intensity peak represents non-specific (reference) signal intensity (Figure 1). Since this technique is not constrained by anatomical boundaries, it maximizes the number of voxels, and hence minimizes the variance, of the reference region. Further, PERSI is designed to be relatively robust to issues related to positioning, registration and PVE.

Methods

Design Summary

We evaluated the performance of PERSI for flortaucipir, relative to traditional reference regions using data from two previously-reported studies which were performed in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments, or comparable ethical standards (NCT01992380 (*20*), NCT02016560 (*21*); no new subjects were administered flortaucipir for this work). The first was a test-retest study, wherein subjects underwent flortaucipir scanning on two occasions within approximately 4 weeks. These data were used to determine the stability of the reference region

without changes in the underlying pathology over time. The second study was a Phase 2 clinical trial with cross-sectional (3 cohorts) as well as longitudinal (18±2 months between scans) components. Knowledge of the amyloid (Aβ) status of the subjects in this study allowed us to evaluate the baseline variability of PERSI vis-à-vis traditional reference regions, since Aβ- older controls are presumed not to accumulate tau pathology outside the mesial temporal lobe. The second study also allowed us to assess the cross-sectional and longitudinal performance of the technique at varying levels of impairment. Details of these analyses are described below.

Flortaucipir Image Acquisition and Preprocessing

Flortaucipir images were acquired as 4 x 5-min dynamic frames, beginning 80 minutes after an approximately 370 MBq injection. Frames 2-4 were rigidly registered to the first frame using the MCFLIRT tool provided by FSL software for inter-frame motion correction. For images acquired at a start time offset from 80 minutes post-injection, time correction factors were calculated for each voxel as the slope of the linear regression line through the four time points, multiplied by the time offset (*22*). Motion- and time-corrected frames were then averaged and co-registered to the subject's MRI using the FSL FLIRT linear registration tool. Next, the unified segmentation and normalization algorithm (*23*) in SPM8 was run to spatially normalize the T1 MRI images acquired at screening to the Montreal Neurological Institute (MNI) brain template (*24*), while simultaneously generating probabilistic segmentations for the GM, WM and CSF.

PERSI implementation

The MNI atlas was segmented using FSL to generate a binary WM image (MNI-WM). PET images spatially normalized to MNI space were then masked with the MNI-WM, and the resulting voxel

8

intensities were plotted as a histogram. The histogram for each subject was fit to a bimodal Gaussian distribution, as shown in Figure 1. The peak location (μ_1) and the sigma (σ_1) of the first Gaussian peak were extracted for analyses. Voxels with values within the full width at half maximum {FWHM = σ_1*2$\sqrt{(2\ln(2))}$} of μ_1, i.e. $\mu_1 \pm$ FWHM, were included in the reference region for that subject.

SUVr Analysis

For comparison with PERSI, SUVr values were generated using four additional reference regions: whole cerebellum (wholeCere), cerebellar GM (cereCrus), atlas-based WM (atlasWM) and subject-specific WM (ssWM). wholeCere was delineated manually on a template (*10*). cereCrus was derived from the AAL VOI "cere-crus-1", and modified to avoid potential overlap with other structures by translating it inferiorly by 6mm. For the atlasWM region, the MNI atlas was segmented to extract a binary WM mask. Then, two steps were undertaken to eliminate voxels susceptible to PVE spillover of counts from the cortex: 1) a mask made up of a combination all volumes from the AAL template, a cerebellum volume and a brainstem volume was created, and applied to the WM segmentation to remove the associated voxels; 2) the resulting volume was eroded using a 5-voxel box erosion in FSL. The same two steps were applied to generate ssWM regions, after a threshold was applied to subject-specific, probabilistic MRI segmentations to preserve voxels with >50% probability of being WM. SUVr values were calculated for a target volume-of-interest (VOI; Figure 2) relative to the five candidate reference regions. This VOI was derived by employing a discriminant analysis technique (MUBADA; (*25*)) to identify voxels most relevant to separation by diagnostic group and amyloid status. More details are provided in the companion manuscript (*20*).

Validation of PERSI Implementation for Longitudinal Data

For longitudinal datasets with more than one scan per subject, the PERSI algorithm was run separately for each scan, and a common reference region was generated from the voxels that were shared by the individual regions. Since this approach may not be practical for all studies, a comparison of SUVr values was performed using both, the combined reference region as well individual reference regions for each scan.

Evaluation of Test-Retest Reproducibility

Twenty-four subjects (10 AD, 8 MCI and 6 CN) were imaged at baseline, followed by a repeat scan between 48 hours and 4 weeks later (*20*). Twenty-one of these subjects had technically adequate scans for inclusion in these analyses. The standard deviation of SUVr percent difference was calculated and compared across reference regions as a measure of test-retest reproducibility. A paired t-test was also performed to evaluate differences between test and retest scans.

Cross-sectional and Longitudinal Analyses

215 subjects from a Phase 2 study (*21*) were imaged at baseline, of which 142 returned for a follow-up scan 18±2 months after the screening visit (shown in parenthesis in Table 1). Subjects also received florbetapir scans at the screening visit that were interpreted as amyloid positive (Aβ+) or amyloid negative (Aβ-) based on visual reads augmented by quantitative information as described in (*26*).

To assess reference region performance in cognitively normal subjects, SUVrs were calculated for 68 CN subjects (see Table 1) determined to be Aβ- (Florbetapir). Of these, 16 were between 20 and 40 years of age (YCN) and the remaining were over 50 years of age (OCN). The accuracy and

precision across reference regions were assessed, given an expected SUVr of 1 for Aβ- CN subjects with no specific uptake. For each reference region, a one-sample t-test was performed to test the hypothesis that the mean SUVr for this cohort was 1.

Cross-sectional SUVr, longitudinal SUVr change and group separation were evaluated in subjects classified into 6 groups based on clinical diagnosis (CN, MCI, AD) and amyloid status (Aβ+, Aβ-; Table 1). SUVr mean and standard deviation values for each group were computed using each of the 5 reference regions. Baseline variability in ΔSUVr was evaluated for the Aβ-, OCN cohort (YCN were not evaluated longitudinally), given expected ΔSUVr=0. The impact of reference regions on Cohen's d effect sizes for differences between Aβ+ impaired (AD and MCI) and Aβ-CN subjects were assessed for cross-sectional SUVr as well as longitudinal SUVr change (ΔSUVr = $SUVr_{follow-up}$ - $SUVr_{baseline}$).

Results

Figure 3 shows representative signal intensity histograms (within the WM mask) used to derive the PERSI reference region for representative subjects clinically diagnosed as YCN, OCN, or Aβ+ AD. For the two normal subjects, only a single peak was observed. For the subject with elevated flortaucipir uptake in the GM, counts that spilled into the WM formed a second peak. PERSI reference regions are comprised of a larger number of voxels than traditional reference regions, even though the volume of the PERSI region reduces with increasing count spillover from the GM (Table 2). The average PERSI region for Aβ+ subjects (though smaller than the average Aβ-PERSI region due to spillover) is a factor of 2.5 larger than the whole cerebellum VOI.

Evaluation of Test-Retest Reproducibility

Test and retest datasets were significantly correlated (p<0.01) for all reference regions (Table 3, Figure 4). Reproducibility (SD) varied depending on reference region, with WM regions (1.8-2.7%) outperforming cerebellar regions (3.7%-4.2%). PERSI SUVr values were the most reproducible of all the SUVr evaluated. The absolute range of SUVr values for WM regions was observed to be smaller than cerebellar regions.

Validation of PERSI Implementation for Longitudinal Data

The two methods employed to identify PERSI volumes in longitudinal datasets generated consistent results, as shown in Figure 5. Although the use of individual masks for each scan produced slightly lower SUVr values overall, the SUVr values were highly correlated (R=1) across the disease spectrum.

Cross-sectional and Longitudinal Analyses

Table 4 (row 1) shows the mean ± SD of SUVr values for Aβ- CN subjects (YCN and OCN). For this cohort, presumed to have no tau pathology, all reference regions generated SUVr values consistent with expectations (expected SUVr = 1), indicating low baseline variability in flortaucipir SUVr overall. However, there appeared to be higher variability with cerebellar regions (7-11%) compared to WM regions (4-7%). Again, PERSI outperformed other regions in both accuracy and reproducibility. One-sample t-tests indicated that PERSI SUVr values for OCN and YCN were not significantly different from 1, whereas statistically significant differences were noted for the other regions.

Similarly, when measuring changes in SUVr over 18 months for Aβ- OCN subjects (Table 4, row 2), PERSI registered the lowest mean and standard deviation of change. This is consistent with the understanding that this demographic was not expected to accumulate additional tau over 18 months. PERSI also generated the tightest range of ΔSUVr values (Table 4, row 3) for this demographic, again suggesting low baseline variability in the measurement of longitudinal changes using PERSI.

Box plots of cross-sectional SUVr grouped by clinical diagnosis and amyloid status for each of the reference regions being evaluated are presented in Figure 6. Overall, Flortaucipir SUVrs were effective at distinguishing Aβ+ impaired subjects from Aβ- subjects (*21*). Cerebellar reference regions again generated a larger range of SUVr values compared to WM. Unlike atlasWM and ssWM, however, the smaller range of PERSI SUVrs did not translate to a lowered ability to distinguish between groups, as revealed by Cohen's d effect sizes in Table 4, row 4. PERSI had the highest effect size for differentiating impaired from control cohorts among all reference regions tested. Although PERSI, atlasWM and ssWM had comparable test-retest reproducibility (see Table 3) and baseline variability (see Table 4; row 1), PERSI had an almost twofold increase in effect size.

Similar advantages were observed with PERSI in the measurement of ΔSUVr 18 months after baseline. The low variability in ΔSUVr for Aβ- groups led to an improved signal-to-noise ratio for the measurement of ΔSUVr for impaired subjects, resulting in a greater effect size (Table 4, row 5). Longitudinal Cohen's d for was 2.34 with PERSI, 1.3 times higher than the next highest value (atlasWM). The occurrence of ΔSUVr values substantially below 0 (considered biologically unlikely) for Aβ+ subjects relative to Aβ- subjects was also reduced with PERSI compared to the

other candidate reference regions (Figure 6). There was one Aβ- AD subject who had ΔSUVr < 0 regardless of reference region, including PERSI; however, the quantification of this subject may have been affected by subject motion.

Discussion

We have evaluated PERSI, a new inter- and intra-subject count normalization technique for Flortaucipir. PERSI estimates non-specific (reference) signal intensity within WM by modeling the spill-in of counts from GM. Our analyses showed that PERSI was associated with the highest effect sizes for diagnostic group separation and lowest variability among reference regions evaluated, for cross-sectional as well as longitudinal measures. In the absence of a truth standard, PERSI generated results consistent with expectations based on the pathophysiology of AD. PERSI should, in theory, also be applicable to other tauopathies.

Unique individual patterns of Flortaucipir uptake (*21*) (and associated spillover) make it challenging to identify anatomical reference tissues for subjects across the disease spectrum. Issues with subject positioning and image truncation, errors in spatial normalization to atlas, and unexpected areas of truly increased tracer uptake in selected reference tissues (e.g., dentate nuclei in cerebellum for flortaucipir) may further compound the problem. Strategies to overcome PVE contamination in the reference region typically involve erosion or shrinking of WM segmentations, similar to our atlasWM and ssWM regions, to minimize voxels with expected contamination. Our cross-sectional and longitudinal effect size results show that this does not completely eliminate PVE. Further, such regions require a number of pre-processing steps that might be sensitive to noise. PERSI is designed to leverage the advantages of WM reference regions (*27*) while mitigating the potential negative impact of PVE. With PERSI, the optimal reference signal is

identified based on voxel intensity rather than anatomical location. Therefore, it is less likely to be affected by differences in uptake patterns or image processing errors. While partial volume correction (PVC) algorithms (*19,28–31*) also correct VOI counts from the effects of spillover, PERSI reduces potential contamination of the reference region when PVC is not applied.

The WM mask described in this manuscript was derived from the MNI template and is common to all subjects. Alternatively, individual WM segmentations can be used as subject-specific masks, if volumetric MRI scans are acquired for each subject. In our studies, data were analyzed using both, subject-specific as well as template-based masks and the results were in agreement. For ease of implementation and broader applicability in studies without MRI data acquisition, our implementation of PERSI employed the template-based WM mask.

For longitudinal datasets, a common PERSI region was obtained for all scans of a given subject from the intersection of the individual PERSI masks. While this additional step was executed to improve the consistency of our longitudinal results, its omission did not alter our conclusions.

PERSI reference regions had several advantages over other candidates. Our results showed excellent test-retest reproducibility (<2%) and equivalently low variability in ΔSUVr over 18±2 months using PERSI. However, baseline variability cannot be used in isolation to select a reference region, since the true variability of the data could be underestimated. PERSI surpassed other candidate regions in differentiating between groups stratified based on clinical diagnoses and amyloid status. AtlasWM and ssWM had similar variability to PERSI, but the group separation ability for these regions was inferior.

With all reference regions, some subjects appeared to decrease flortaucipir binding over time. Although it is not impossible for tau to decrease in later disease stages due to loss of tau producing neurons and atrophy, in these mostly MCI and early AD subjects the observed decreases are most likely artifacts of image processing methods. The occurrence of these negative ΔSUVr values was reduced with PERSI.

The volume of the PERSI region (Table 2) varies depending on the number of voxels deemed to be contaminated by spillover; yet the smallest PERSI region seen in our datasets is substantially larger than traditional reference regions. Large regions are another contributor to low variability, further substantiating the suitability of PERSI for longitudinal measurements.

White matter reference regions may suffer quantitative biases due to count variations brought about by ischemic changes in subcortical white matter. Limited literature (*32*) suggests that calcification after ischemic changes could result in elevated flortaucipir uptake. In theory, PERSI should be unaffected by focal irregularities in counts because the method specifically excludes areas with a detectably different signal intensity from the overall WM signal. However, this remains to be proven in future work.

A drawback to PERSI is that it involves additional processing compared to standard reference regions. However, population-based SUV normalization methods have been proposed in the past. Turkheimer et al (*33*) developed a supervised clustering method for the neuroinflammatory marker ^{11}C-PK11195 which assumed a weighted linear combination of 6 kinetic classes. A novel method (*34*) was recently introduced for tau imaging using ^{11}C-PBB3. This method involved first calculating SUVr images using a standard cerebellar cortex reference tissue method, then generating a new reference region for each subject using only the voxels with SUVr between the

mean and 2 SD below the mean SUVr of a database of healthy controls. These methods could provide advantages over traditional methods, and with rigorous evaluation may prove to be critical to future trials.

In summary, PERSI reduced variability while enhancing discrimination between diagnostic cohorts. These improvements could lead to more accurate disease staging and robust measurements of changes in tau burden over time for evaluation of the efficacy of putative therapeutic interventions.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Financial disclosure

SS, MDD, IK, MN, ML, ADJ, MJP, MAM, are employees of Avid Radiopharmaceuticals, a wholly owned subsidiary of Eli Lilly and company. Avid Radiopharmaceuticals sponsored this study.

References

1. Huang S, Phelps M. Principles of tracer kinetic modeling in positron emission tomography and autoradiography. *Positron Emiss Tomogr Autoradiogr Princ Appl Brain Heart*. 1986:287-346.

2. Mintun MA, Raichle ME, Kilbourn MR, Wooten GF, Welch MJ. A quantitative model for the in vivo assessment of drug binding sites with positron emission tomography. *Ann Neurol*. 1984;15:217-227.

3. Logan J. Graphical analysis of PET data applied to reversible and irreversible tracers. *Nucl Med Biol*. 2000;27:661-670.

4. Gunn RN, Lammertsma AA, Hume SP, Cunningham VJ. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *NeuroImage*. 1997;6:279-287.

5. Thie JA. Understanding the Standardized Uptake Value, Its Methods, and Implications for Usage. *J Nucl Med*. 2004;45:1431-1434.

6. Price JC, Klunk WE, Lopresti BJ, et al. Kinetic Modeling of Amyloid Binding in Humans using PET Imaging and Pittsburgh Compound-B. *J Cereb Blood Flow Metab*. 2005;25:1528-1547.

7. Shcherbinin S, Devous MD, Schwarz AJ, Joshi AD, Navitsky M, Mintun MA. Region-dependent kinetics of the Tau PET tracer [18 F]-AV-1451 (T807). *Alzheimers Dement J Alzheimers Assoc*. 2015;11:P113.

8. Wong DF, Rosenberg PB, Zhou Y, et al. In Vivo Imaging of Amyloid Deposition in Alzheimer Disease Using the Radioligand 18F-AV-45 (Flobetapir F 18). *J Nucl Med*. 2010;51:913-920.

9. Clark CM, Pontecorvo MJ, Beach TG, et al. Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-β plaques: a prospective cohort study. *Lancet Neurol*. 2012;11:669-678.

10. Joshi AD, Pontecorvo MJ, Clark CM, et al. Performance characteristics of amyloid PET with florbetapir F 18 in patients with alzheimer's disease and cognitively normal subjects. *J Nucl Med Off Publ Soc Nucl Med*. 2012;53:378-384.

11. Landau SM, Mintun MA, Joshi AD, et al. Amyloid deposition, hypometabolism, and longitudinal cognitive decline. *Ann Neurol*. 2012;72:578-586.

12. Klunk WE, Koeppe RA, Price JC, et al. The Centiloid Project: Standardizing quantitative amyloid plaque estimation by PET. *Alzheimers Dement J Alzheimers Assoc*. 2015;11:1-15.e4.

13. Wong K-P, Wardak M, Shao W, et al. Quantitative analysis of [18F]FDDNP PET using subcortical white matter as reference region. *Eur J Nucl Med Mol Imaging*. 2009;37:575-588.

14. Joshi A, Pontecorvo M, Navitsky MA, Kennedy IA, Mintun M, Devous MD. Measuring change in beta-amyloid burden over time using florbetapir-pet and a subcortical white matter reference region. *Alzheimers Dement J Alzheimers Assoc*. 2014;10:P902.

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

16. Joshi AD, Pontecorvo MJ, Lu M, Skovronsky DM, Mintun MA, Devous MD. A Semiautomated Method for Quantification of F 18 Florbetapir PET Images. *J Nucl Med*. 2015;56:1736-1741.

17. Khalil MM. *Basic Science of PET Imaging*. Springer; 2016.

18. Erlandsson K, Buvat I, Pretorius PH, Thomas BA, Hutton BF. A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Phys Med Biol*. 2012;57:R119.

19. Thomas BA, Erlandsson K, Modat M, et al. The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease. *Eur J Nucl Med Mol Imaging*. 2011;38:1104-1119.

20. Devous MD Sr, Joshi AD, Navitsky M, et al. Test–Retest Reproducibility for the Tau PET Imaging Agent 18F-AV-1451. *J Nucl Med*. 2017 *Manuscr Submitt Publ*.

21. Pontecorvo MJ, Devous MD, Navitsky M, et al. Relationships between flortaucipir PET tau binding and amyloid burden, clinical diagnosis, age and cognition. *Brain*. 2017;140:748-763.

22. Southekal S, Devous MDS, Navitsky M, Kennedy I, Joshi AD, Mintun MA. Correction for acquisition time discrepancies in SUVr analyses of 18F-AV1451 tau images. In: *Human Amyloid Imaging Conference*. Miami, USA; 2016.

23. Ashburner J, Friston KJ. Unified segmentation. *NeuroImage*. 2005;26:839-851.

24. Fonov VS, Evans AC, McKinstry RC, Almli CR, Collins DL. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *ResearchGate*. 2009;47.

25. Abdi H, Williams LJ, Beaton D, Posamentier MT, Harris TS, Krishnan A, et al. Analysis of Regional Cerebral Blood Flow Data to Discriminate Among Alzheimer's Disease, Frontotemporal Dementia, and Elderly Controls: Multi-Block Barycentric Discriminant (Mubada) Methodology. *J Alzheimers Dis JAD*. 2012;31:S189-S201.

26. Pontecorvo MJ, Arora AK, Devine M, et al. Quantitation of PET signal as an adjunct to visual interpretation of florbetapir imaging. *Eur J Nucl Med Mol Imaging*. 2017:1-13.

27. Landau SM, Fero A, Baker SL, et al. Measurement of Longitudinal β-Amyloid Change with 18F-Florbetapir PET and Standardized Uptake Value Ratios. *J Nucl Med*. 2015;56:567-574.

28. Meltzer CC, Leal JP, Mayberg HS, Wagner HN, Frost JJ. Correction of PET data for partial volume effects in human cerebral cortex by MR imaging. *J Comput Assist Tomogr*. 1990;14:561-570.

29. Müller-Gärtner HW, Links JM, Prince JL, et al. Measurement of Radiotracer Concentration in Brain Gray Matter Using Positron Emission Tomography: MRI-Based Correction for Partial Volume Effects. *J Cereb Blood Flow Metab*. 1992;12:571-583.

30. Rousset O, Ma Y, Evans A. Correction for partial volume effects in PET: principle and validation. *J Nucl Med Off Publ Soc Nucl Med*. 1998;39:904-911.

31. Sattarivand M, Kusano M, Poon I, Caldwell C. Symmetric geometric transfer matrix partial volume correction for PET imaging: principle, validation and robustness. *Phys Med Biol*. 2012;57:7101.

32. Lockhart SN, Ayakta N, Winer JR, Joie RL, Rabinovici GD, Jagust WJ. Elevated 18F-AV-1451 PET tracer uptake detected in incidental imaging findings. *Neurology*. February 2017:10.1212/WNL.0000000000003724.

33. Turkheimer FE, Edison P, Pavese N, et al. Reference and Target Region Modeling of [11C]-(R)-PK11195 Brain Studies. *J Nucl Med*. 2007;48:158-167.

34. Kimura Y, Endo H, Ichise M, et al. A new method to quantify tau pathologies with 11C-PBB3 PET using reference tissue voxels extracted from brain cortical gray matter. *EJNMMI Res*. 2016;6:24.
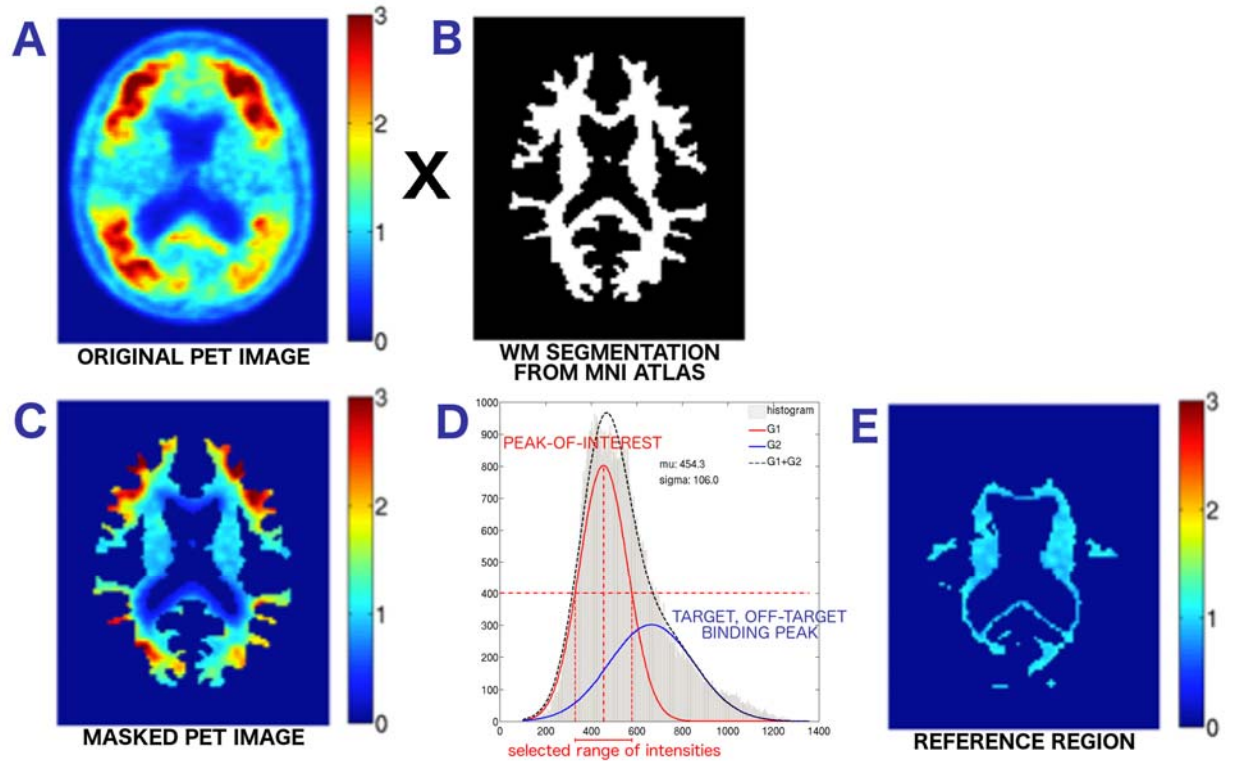
Figure 1. Graphical description of PERSI. The original PET image A is masked using B to extract counts in the WM (C). The resulting non-zero counts are plotted as a histogram (D) and fit to a bimodal Gaussian distribution. The parameters of the lower-intensity peak, shown in red, are used to select the range of values to be used as the reference region E.
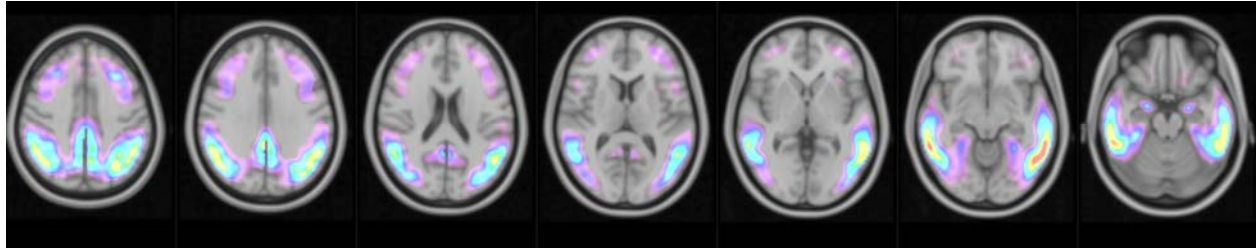
Figure 2. MUBADA VOI obtained using discriminant analysis used to identify voxels predominantly associated with group separation between cohorts, overlaid on MNI template. This VOI is applied to measured PET data in a weighted fashion to generate SUVr values.
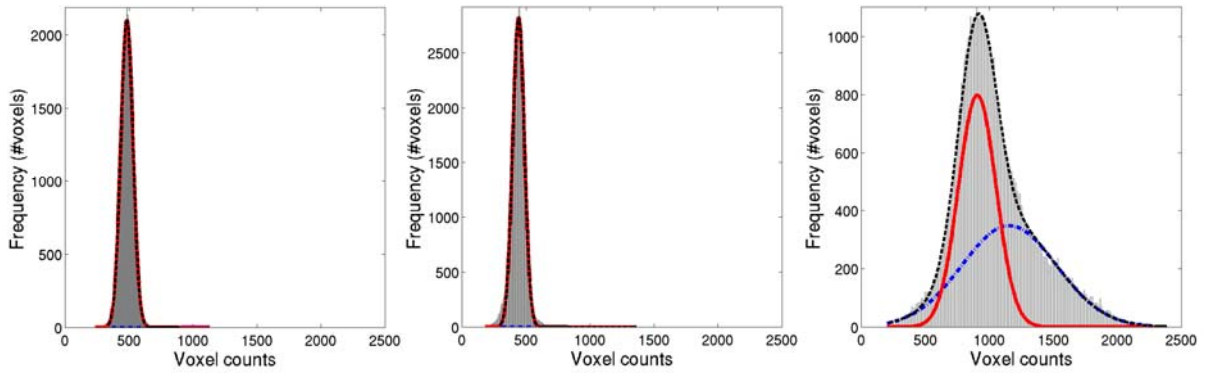
Figure 3. Typical WM histograms for 3 subjects: (left) 36 y.o. cognitively normal Aβ-; (center) 69 y.o. cognitively normal Aβ-; and (right) 75 y.o. AD Aβ+. The red solid line denotes lower-intensity signal and the blue dot-dash line indicates higher-intensity peak. Black dashed line shows sum of two peaks (bimodal Gaussian). For the two normal subjects, only a single peak is observed. For the subject with elevated Flortaucipir uptake in the GM, counts that have spilled into the WM are captured by the blue peak.
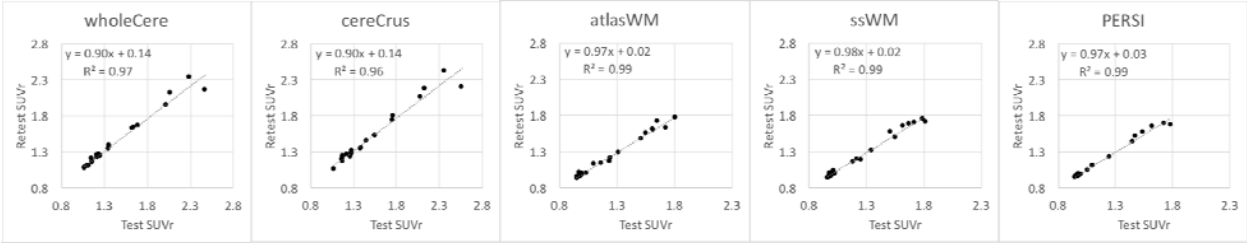
Figure 4. Comparison of test and retest conditions for 21 subjects using 5 reference regions
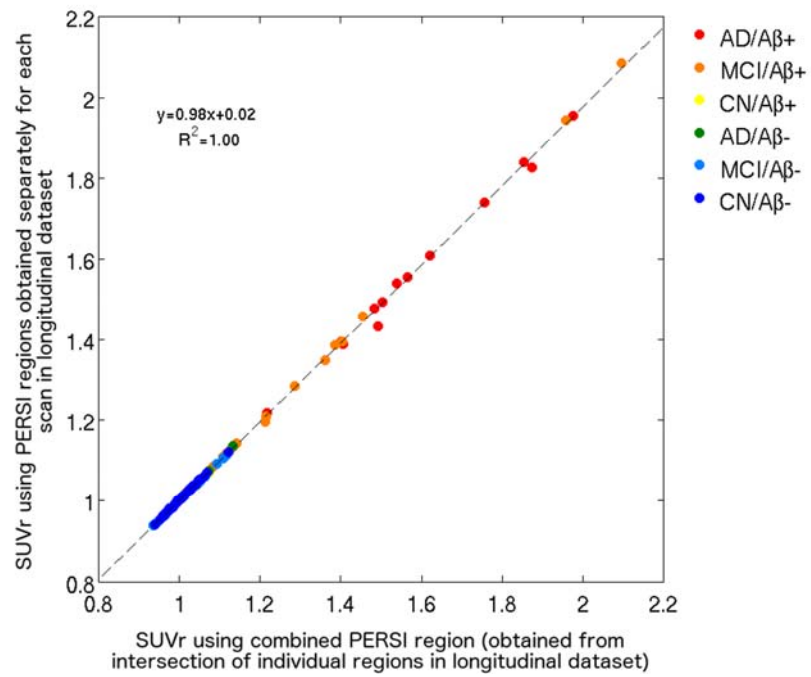
Figure 5. PERSI SUVr comparison for longitudinal data: Individual vs. Combined mask
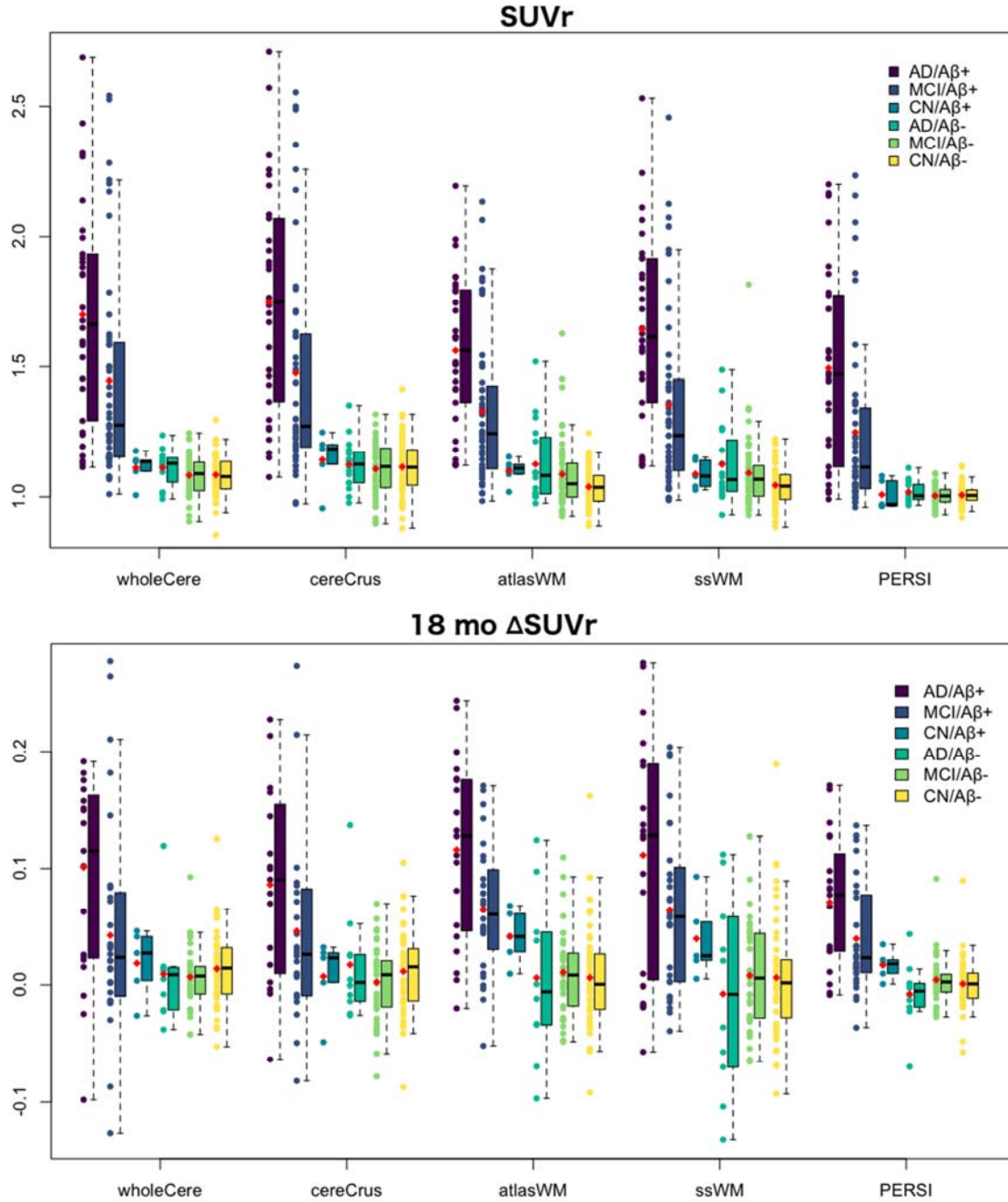
Figure 6. Comparison of SUVr (top) and ΔSUVr (bottom) values using 5 reference regions and MUBADA as the target region. Subjects are grouped by clinical diagnosis and amyloid status. Red dot denotes mean.

TABLES

Table 1. Cross-sectional and longitudinal study participants

Clinical Diagnosis Group	Symbol	N Visit1 (Visit2)		
		Aβ-	Aβ+	Total
Probable AD	AD	16 (10)	30 (19)	46 (29)
Mild Cognitive Impairment	MCI	49 (33)	47 (28)	96 (61)
Older Cognitively Normal (≥ 50y)	OCN	52 (47)	5 (5)	57 (52)
Young Cognitively Normal (< 40y)	YCN	16 (0)	0 (0)	16 (0)
All		**133 (90)**	**82 (52)**	**215 (142)**

Table 2. Reference region sizes

Region	Size (number of 2 x 2 x 2 mm voxels)
Whole Cerebellum (wholeCere)	15,878
Cerebellar gray (cereCrus)	1,815
Pons	1,226
Atlas-based WM (atlasWM)	8,384
Average subject-specific WM region (ssWM)	1,140
Average PERSI region (Aβ-)	44,222
Average PERSI region (Aβ+)	38,954

Table 3. Comparison of reference regions using test-retest data

Reference Region	R^2	Mean %Difference	SD %Difference
Whole cerebellum	0.971	0.066	3.700
Cerebellum crustaneous	0.961	0.042	4.179
Atlas WM	0.986	-0.764	2.616
Subject-specific WM	0.987	-0.869	2.663
PERSI	0.990	-0.165	1.841
(All correlations were statistically significant with p<0.01)			

Table 4. Comparison of 5 reference regions using cross-sectional and longitudinal data. 1. Baseline variability of SUVr values measured in Aβ- CN subjects. 2. Baseline variability of ΔSUVr values measured in Aβ- CN subjects. 3. Range of ΔSUVr values in Aβ- OCN subjects 4. Cross-sectional effect sizes (Cohen's d) 5. Longitudinal effect sizes (Cohen's d)

Measure	Group	wholeCere	cereCrus	atlasWM	ssWM	PERSI
1. SUVr Mean ± SD	YCN (n=16)	1.06 ± 0.08	1.10 ± 0.08	0.98 ± 0.06	0.99 ± 0.07	1.01 ± 0.04
	OCN (n=52)	1.09 ± 0.07	1.12 ± 0.11	1.06 ± 0.07	1.06 ± 0.07	1.01 ± 0.04
2. ΔSUVr Mean ± SD	OCN/Aβ- (n=47)	0.01 ± 0.03	0.01 ± 0.04	0.01 ± 0.04	0.01 ± 0.04	0.00 ± 0.02
3. ΔSUVr Min/ Max	OCN/Aβ- (n=47)	-0.05 / 0.13	-0.09 / 0.11	-0.09 / 0.16	-0.09 / 0.19	-0.06 / 0.09
4. Cohen's d SUVr	Aβ+ AD & MCI vs. Aβ- CN	6.10	4.57	5.12	5.43	9.45
	Aβ+ AD vs. Aβ- CN	8.14	6.20	7.01	7.68	13.61
	Aβ+ MCI vs. Aβ- CN	4.80	3.53	3.91	3.99	6.79
5. Cohen's d ΔSUVr	Aβ+ AD & MCI vs. Aβ- CN	1.65	1.39	1.82	1.47	2.34
	Aβ+ AD vs. Aβ- CN	2.75	2.03	2.51	2.00	3.19
	Aβ+ MCI vs. Aβ- CN	0.99	0.95	1.34	1.11	1.76