Contributions of Imprecision in PET-MRI Rigid Registration to Imprecision in Amyloid PET SUVR Measurements

Christopher G. Schwarz [1]* David T. Jones,[2] Jeffrey L. Gunter,[1,3] Val J. Lowe,[1] Prashanthi Vemuri,[1] Matthew L. Senjem,[1,3] Ronald C. Petersen,[2] David S. Knopman,[2] and Clifford R. Jack Jr.[1] The Alzheimer's Disease Neuroimaging Initiative

[1]*Department of Radiology, Mayo Clinic and Foundation, Rochester, Minnesota*
[2]*Department of Neurology, Mayo Clinic and Foundation, Rochester, Minnesota*
[3]*Department of Information Technology, Mayo Clinic and Foundation, Rochester, Minnesota*

Abstract: Quantitative measurement of β-amyloid from amyloid PET scans typically relies on localizing target and reference regions by image registration to MRI. In this work, we present a series of simulations where 50 small random perturbations of starting location and orientation were applied to each subject's PET scan, and rigid registration using *spm_coreg* was performed between each perturbed PET scan and its corresponding MRI. We then measured variation in the output PET-MRI registrations and how this variation affected the resulting SUVR measurements. We performed these experiments using scans of 1196 participants, half using 18F florbetapir and half using 11C PiB. From these experiments, we measured the magnitude of the imprecision in the rigid registration steps used to localize measurement regions, and how this contributes to the overall imprecision in SUVR measurements. Unexpectedly, we found for both tracers that the imprecision in these measurements depends on the degree of amyloid tracer uptake, and thus also indirectly on Alzheimer's disease clinical status. We then examined common choices of reference regions, and we show that SUVR measurements using supratentorial white matter references are relatively resistant to this source of error. We also show that the use of partial volume correction further magnifies the effects of registration imprecision on SUVR

Additional Supporting Information may be found in the online version of this article.

The Alzheimer's Disease Neuroimaging Initiative: A portion of data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Contract grant sponsor: NIH grants; Contract grant number: R01 AG011378, R00 AG37573, R01 AG041851, U01 AG24904, U01 AG06786, P50 AG16574, R01 AG034676; Contract grant sponsor: The Alexander Family Professorship of Alzheimer's Disease Research, Mayo Clinic; the GHR Foundation; Contract grant sponsor: Elsie and Marvin Dekelboum Family Foundation

*Correspondence to: Christopher G. Schwarz, Ph.D., Mayo Clinic, Diagnostic Radiology, 200 First Street SW, Rochester, MN, 55905. E-mail: schwarz.christopher@mayo.edu

Received for publication 7 December 2016; Revised 6 April 2017; Accepted 9 April 2017.

DOI: 10.1002/hbm.23622

Published online 22 April 2017 in Wiley Online Library (wileyonlinelibrary.com).

measurements. Together, these results suggest that this rigid registration step is an attractive target for future work in improving measurement techniques. *Hum Brain Mapp 38:3323–3336, 2017.*

Key words: reproducibility of results; Alzheimer disease; positron-emission tomography; image processing; computer-assisted; amyloid; florbetapir; Pittsburgh compound B

INTRODUCTION

Quantitative measurements of β-amyloid from Positron Emission Tomography (PET) scans for clinical trials and longitudinal observational studies of Alzheimer's disease (AD) are typically calculated using the Standardized Uptake Value Ratio (SUVR) [Jagust et al., 2010; Salloway et al., 2014; Siemers et al., 2016]. SUVR is a ratio of the measured uptake in a target tissue region of interest (ROI) divided by uptake in a reference tissue ROI that is assumed to be free of pathology [Zasadny and Wahl, 1993]. For amyloid PET scans in studies of AD, the target ROI typically consists of a large region of cortical gray matter (GM). The reference ROI is typically all or part of the cerebellum [Klunk et al., 2004], especially in cross-sectional studies. The reference and target ROIs are typically localized using a series of registrations: 1) rigid registration between the PET scan and a corresponding T1-weighted (T1-w) Magnetic Resonance Imaging (MRI) scan, and 2) nonlinear registration between the corresponding T1-w MRI and a common template space. ROIs defined in the common template space are then backward-propagated to the subject scans, and SUVR is calculated from the mean uptake values in the propagated target and reference ROIs. Many amyloid PET studies also use partial volume correction (PVC), which uses MRI segmentations to estimate and correct for varying tissue fraction per PET voxel [Meltzer et al., 1990]. This step also requires accurate PET-MRI alignment.

Because amyloid PET SUVR measurements play a critical role in studies of Alzheimer's disease, their degree of accuracy (absence of bias or systematic error) and precision (limited by random, non-systematic error) directly impacts the power of these studies. In longitudinal studies measuring change over time, precision of serial measurements is particularly important [Schmidt et al., 2015]. Prior works have primarily focused upon other relevant parameters such as use of PVC, choice of target voxels, and choice of reference region [Brendel et al., 2015; Carbonell et al., 2015; Chen et al., 2015; Landau et al., 2015; Schwarz et al., 2016b]. In this work, we focus specifically on a relatively under-studied measurement step: cross-modality rigid registration. Serial measurements of amyloid PET can show noisy or implausible within-subject trajectories such as those that decrease over time [Landau et al., 2015; Scheinin et al., 2009; Schwarz et al., 2016b], but imprecise rigid registration has not been studied as a quantifiable contribution to the imprecision in these measurements. Unlike other sources of imprecision in PET image measurements such as limited resolution, variations in cerebral blood flow (i.e., wash-in and wash-out effects), variations in subject positioning, and variations in scan time relative to the injection [Schmidt et al., 2015], imprecision in registration occurs in post-processing software. Therefore, it is a source of imprecision that could be improved, and these improvements could be applied retrospectively to existing data.

In this work, we quantify the amount of imprecision in rigid registrations between amyloid PET and MRI, and we measure the effects of this imprecision on the SUVR measurements that depend on them. We then demonstrate that this imprecision is affected by subject amyloid burden and therefore indirectly by clinical diagnosis. Finally, we explore how this variance is affected by different variations in SUVR calculations, such as different reference regions and PVC.

METHODS

Subject Characteristics

Scans were drawn from two primary data sources, each using different amyloid PET ligands: (1) Mayo Clinic studies, and (2) the Alzheimer's Disease Neuroimaging Initiative (ADNI). Mayo Clinic amyloid PET scans used [11C] Pittsburgh Compound B (PiB) [Klunk et al., 2004], and ADNI amyloid PET scans used [18F] AV45 (florbetapir) [Wong et al., 2010]. The two ligands have different binding properties, and we wanted to test our hypotheses using both types of data. Ideally, one would use a large dataset of subjects varying across the entire range of disease burden that were imaged on the same day with both ligands, but such a dataset does not exist. Because the Mayo and ADNI studies use different recruitment strategies, we wanted to ensure that the demographics and distribution of subjects across the clinical spectrum were similar for each dataset, that is, each ligand. Therefore, we selected subjects (with PET scans and corresponding T1-weighted anatomical MRI scans of acceptable quality) from each study by matching them on age, sex, and clinical diagnosis (clinically normal [CN], mild cognitive impairment [MCI], and Alzheimer's Disease dementia [AD]).

Mayo Clinic data was pooled from two sources: the Mayo Clinic Study of Aging (MCSA) and Mayo Clinic Alzheimer's Disease Research Center (ADRC). MCSA is an epidemiological cognitive aging study in Rochester,

TABLE I. Characteristics of ADNI and MCSA individuals by diagnosis

	CN		MCI		AD dementia	
Characteristic	ADNI ($n = 303$)	Mayo ($n = 303$)	ADNI ($n = 200$)	Mayo ($n = 200$)	ADNI ($n = 95$)	Mayo ($n = 95$)
Age, years						
Median (IQR)	72 (68, 78)	72 (68, 78)	76 (71, 82)	76 (71, 81)	74 (68, 80)	74 (67, 80)
Min, Max	56, 95	56, 95	55, 89	55, 89	55, 88	55, 88
Male gender	138 (46%)	138 (46%)	134 (67%)	134 (67%)	55 (58%)	55 (58%)
Education, years						
Median (IQR)	16 (15, 18)	15 (12, 17)	16 (14, 18)	14 (12, 16)	16 (14, 18)	16 (12, 16)
Min, Max	6, 20	8, 20	9, 20	0, 20	9, 20	7, 20
APOE ϵ4 positive	86 (28%)	89 (29%)	92 (46%)	97 (49%)	64 (67%)	64 (71%)
MMSE						
Median (IQR)	29 (28, 30)	29 (28, 29)	28 (27, 29)	26 (24, 27)	23 (21, 25)	21 (17, 23)
Min, Max	24, 30	24, 30	24, 30	17, 30	19, 26	5, 29
CDR-SB						
Median (IQR)	0 (0, 0)	0 (0, 0)	2 (1, 2)	1 (0, 2)	4 (4, 6)	4 (3, 7)
Min, Max	0, 1	0, 2	0, 6	0, 8	1, 10	0, 13

Olmsted County, MN [Petersen et al., 2010; Roberts et al., 2008]. The ADRC study recruits and follows subjects initially seen as patients in the behavioral neurology practice at Mayo Clinic.

ADNI data used in this study were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). For up-to-date information, see the ADNI homepage [Anon, 2013].

All studies were approved by their respective institutional review boards and all subjects or their surrogates provided informed consent compliant with HIPAA regulations. In total, we used single-timepoint PET-MRI scan-pairs of $n = 1196$ subjects, with $n = 598$ from each study. Within each study, $n = 303$ were clinically normal (CN), $n = 200$ had mild cognitive impairment (MCI), and $n = 95$ had AD dementia (AD). Detailed characteristics of these individuals are given in Table I.

Scan Acquisition Parameters

Mayo Clinic T1-weighted MRI were performed using three General Electric (GE) 3T scanners (models: Discovery MR750, Signa Excite, Signa HDx) with a sagittal 3D magnetization-prepared rapid acquisition gradient-recalled echo (MP-RAGE) sequence. Repetition time (TR) was $\approx$ 2300 ms, echo time (TE) $\approx$ 3 ms, inversion time (TI) = 900 ms, and voxel dimensions were $\approx$ 1.20 $\times$ 1.015 $\times$ 1.015 mm. [11C] Pittsburgh Compound B (PiB) PET/CT studies were acquired using GE scanners (models Discovery 690, Discovery HR, Discovery RX, GE Healthcare, Waukesha, WI). Subjects were injected with PiB and a low-dose computed tomography (CT) scan was acquired. Beginning at 40 min post-injection, subjects underwent a 20-min PET scan sequence with four five-minute dynamic frames. PET images were reconstructed using an iterative reconstruction algorithm (256 matrix, 300 mm field of view, 1.17 mm $\times$ 1.17 mm $\times$ 3.27 mm voxel size). Standard corrections for attenuation, scatter, randoms and decay were applied and a 5 mm Gaussian post filter. The four dynamic frame images were averaged to create a single static PiB image. Further details have been previously published [Jack et al., 2013]. The effective resolution of these images has been experimentally estimated as $\approx$ 8mm [Joshi et al., 2009].

Details of ADNI-standard acquisitions have been previously published for T1-weighted MRI [Jack et al., 2010] and for [18F] AV45 PET [Jagust et al., 2010]. Our study used ADNI's post-processed PET scans, which were previously smoothed to an approximate resolution of $\approx$8mm for uniformity across sites and scanners [Joshi et al., 2009].

Common Processing

All T1-weighted MRI scans were preprocessed for B0 intensity inhomogeneity correction using Unified Segmentation [Ashburner and Friston, 2005] as implemented in SPM12 (v6470) with a population-specific in-house template named *MCSA202*, and several population-specific parameter alterations previously described [Schwarz et al., 2016a]. Nonlinear registration parameters were calculated between each B0-corrected MRI scan and the population-specific template using the Advanced Normalization Tools (ANTs) Symmetric Normalization (SyN) algorithm [Avants et al., 2008]. These parameters were used to transform regions of interest (ROIs) from the template space onto the space of the MRI scan of each subject using Nearest-Neighbor interpolation using the *antsApplyTransforms* tool from ANTs. The above steps were performed only once for each T1-weighted scan and did not vary for each perturbation in the simulation. For each rigid transformation between each subject's corresponding perturbed PET and MRI (see below), PET images were interpolated to the

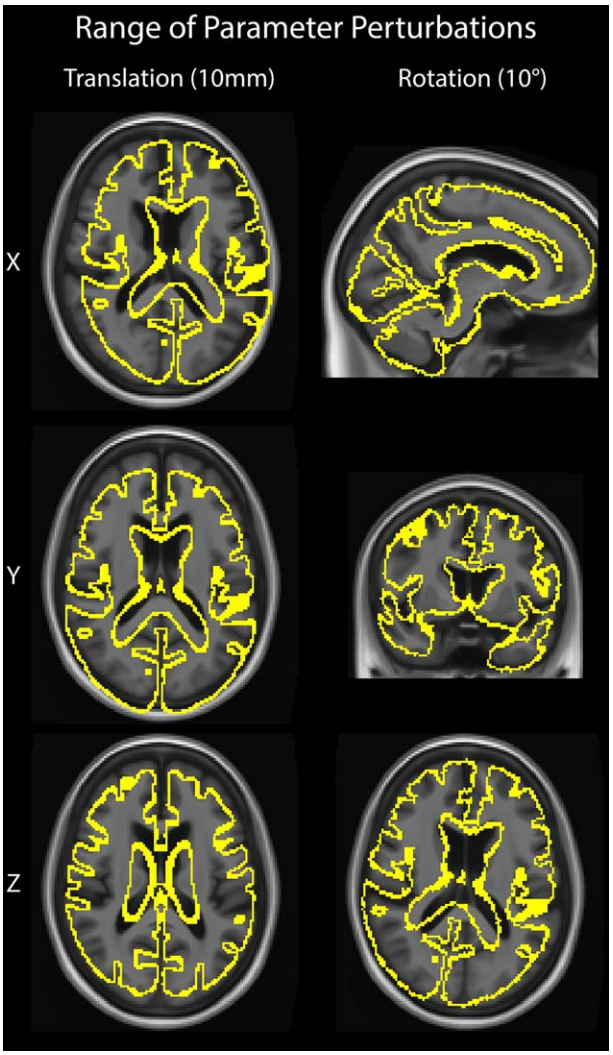

Figure 1.

Range of Parameter Perturbations: Our analysis perturbed the geometry of each PET image's header prior to performing registration with its corresponding MRI scan, and then measured the instability of these registrations and their resultant SUVR calculations. Each translation parameter was perturbed by up to 10mm, and each rotation parameter was perturbed by up to 10° in each direction. In this figure, we illustrate the six rigid registration parameters perturbed by their maximum amounts. Each subfigure shows a brain contour overlaid on an MRI of the same brain, after rigid perturbation by the designated amount. [Color figure can be viewed at wileyonlinelibrary.com]

space of the MRI image using Trilinear interpolation and the *antsApplyTransforms* tool.

In experiments using partial volume correction, we used an in-house implementation of two-class PVC [Meltzer et al., 1990] using tissue-class segmentations produced by the above application of SPM12 and a point spread function of 8 mm.

Evaluation Criteria and Statistical Methods

To measure imprecision in rigid registration between PET and MRI, we applied 50 different rigid (6 degrees of freedom) perturbations to each subject's PET image's header geometry (i.e., its starting position in world coordinate space) and examined the resulting variation in registration parameters and their resulting SUVR measurements. We further describe these steps below, and we provide a flowchart in Supporting Information Figure S1.

If a hypothetical registration algorithm performed perfectly, it could be given a pair of input images and return for all starting positions an identical, registered moving image that forms a global-minimum registration for that image pair. Our simulations were designed to examine the variation in output registrations when given a variety of starting locations, that is, the degree to which this ideal occurs in practice. Because it is known that the algorithm in practice requires some degree of overlap between initial image positions, caution was necessary to ensure that our perturbations met this condition. To this end, we first performed a single rigid registration for each pair of PET and MRI scan and altered the header information of the PET image files such that the two were pre-aligned. These pre-aligned scan pairs were assessed visually to ensure that none were grossly misaligned.

We performed 50 random perturbations around each pre-registered starting point by selecting each of the six free parameters of the rigid perturbing transformation independently from a uniform distribution. Translation parameters (x,y,z) were each independently selected from a range [−10 mm,10 mm]. Rotation parameters (around each axis x,y,z)[1] were each independently selected from a range [−10°,10°]. To select these ranges, we examined the distribution of each parameter from the initial PET-MRI rigid registrations. Among these, rotation parameters up to 20° and translation parameters up to 40 mm were frequently observed. Because we wished to avoid gross registration failures of the perturbed input images, we selected 10° and 10 mm as conservative upper limits for each parameter. This degree of perturbation was small enough not to cause total registration failures and did not overestimate the true variation in subject head placement during scans. We illustrate the upper limits of our chosen parameter ranges in Figure 1.

A consistent random seed was used to select perturbations for each pair of scans from each subject, such that all scans underwent the same set of 50 perturbations. For each random perturbation, a file copy of the PET image was made, and this copy was perturbed by modifying the image header geometry (nifti format s-form/q-form parameters) without resampling or otherwise modifying the image voxels. These 50 geometry-perturbed files were used to provide 50 different starting points for PET-MRI registration. PET-MRI registrations were then performed for

[1]In this work, (x,y,z) directions map directly to a Right, Anterior, Superior (RAS) coordinate system. Rotations were about the origin as defined in the PET image after initial co-registration to the T1-w image.

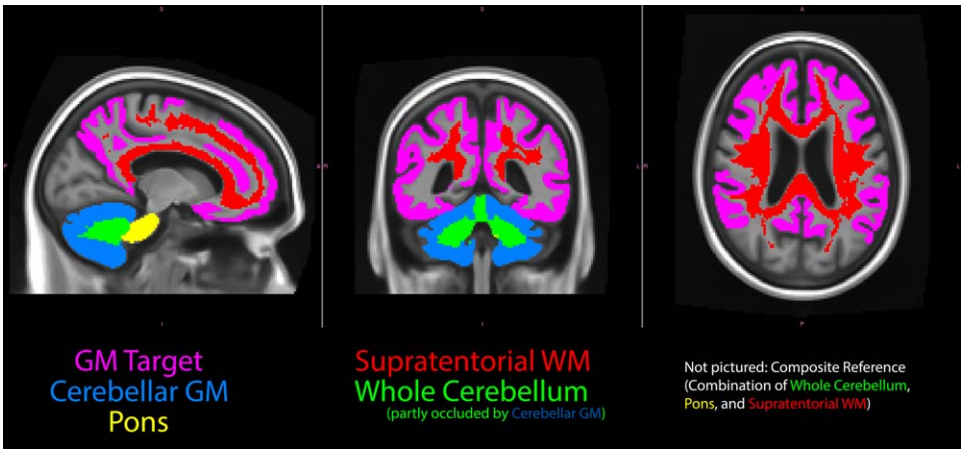

Figure 2.
ROIs used during SUVR computations in this analysis. The GM Target region (pink) was used as the target ROI (SUVR numerator) for all variants. The other ROIs pictured are the multiple choices of reference region (SUVR denominator) that were compared. [Color figure can be viewed at wileyonlinelibrary.com]

each pair of original MRI scan and a header-perturbed copy of the PET image. We then examined these outputs: (1) computed registration parameters and their deviation from the expected inverse of each perturbation, (2) coefficient of variation (CV = mean/standard deviation) in computed SUVR values from these registrations using a variety of potential reference regions, both with and without PVC.

All PET-MRI registrations were performed using SPM's coregister function (*spm_coreg*) as implemented in SPM12. The spm_coreg method is so commonly employed in SUVR analysis of amyloid PET [Edison et al., 2013; Klunk et al., 2015; Landau and Jagust, 2015; Lowe et al., 2009; Ng et al., 2007; Schain et al., 2014; Schwarz et al., 2016a; Zhou et al., 2007] that we consider it the standard of reference, and therefore choose it as the primary method that we examine in this work. We used algorithm-default registration/convergence parameters with six degrees of freedom (rigid) registration and the default, normalized mutual information cost function. We also replicated our experiments with all per-parameter tolerances (designed to directly set precision limits for each parameter) set to 1/4th of their normal amounts, and this had no qualitative difference upon any results or conclusions (data not presented). Experiments were also replicated using an alternate implementation of the *spm_coreg* algorithm included with FreeSurfer version 6 (see Strengths and Limitations of Current Study section). No resampling was performed during the pre-alignment, perturbation, or post-perturbation alignment steps; these only modified image header geometry information to alter each scan's location in world coordinate space.

SUVR computations all used a cerebral GM target ROI consisting of those regions primarily affected by β-amyloid deposition: parietal, cingulate precuneus, prefrontal, orbitofrontal, temporal, and anterior cingulate [Jack et al., 2013], measured only in those voxels segmented as GM (with ≥50% confidence) in the corresponding T1-weighted image. We used this target with five different choices of reference region: cerebellar GM, supratentorial white matter (WM), whole cerebellum, pons, and a composite reference including all voxels in the whole cerebellum, pons, and supratentorial WM. Our supratentorial white matter ROI was defined by a combination of the centrum semiovale, corona radiata, and corpus callosum ROIs from the Johns Hopkins University single-subject WM atlas [Oishi et al., 2009], and measured only in those voxels segmented as WM (with ≥50% confidence). The cerebellar GM ROI was defined by an in-house modification of the AAL atlas, and measured only in those voxels segmented as GM (with ≥50% confidence). The pons ROI was defined by the same in-house AAL atlas, and measured only in those voxels segmented as WM (with ≥50% confidence). The whole-cerebellum ROI was defined by an in-house lobar atlas and measured only in those voxels segmented as either GM or WM. We illustrate these regions in Figure 2.

For all box plots, whiskers extend from the hinge to the highest/lowest value within 1.5 × the inter-quartile range (IQR). Notches are given at values 1.5 × (IQR/sqrt(n)), representing a 95% confidence interval around the median. Therefore, non-overlapping notched areas indicate significant differences in median values, with 95% confidence [McGill et al., 1977].

RESULTS AND DISCUSSION

Imprecision in PET-MRI Rigid Registration Is Larger in Rotation Parameters and Increases with Clinical Disease Severity

For each perturbation, we multiplied the perturbation matrix with the transformation matrix output by the registration

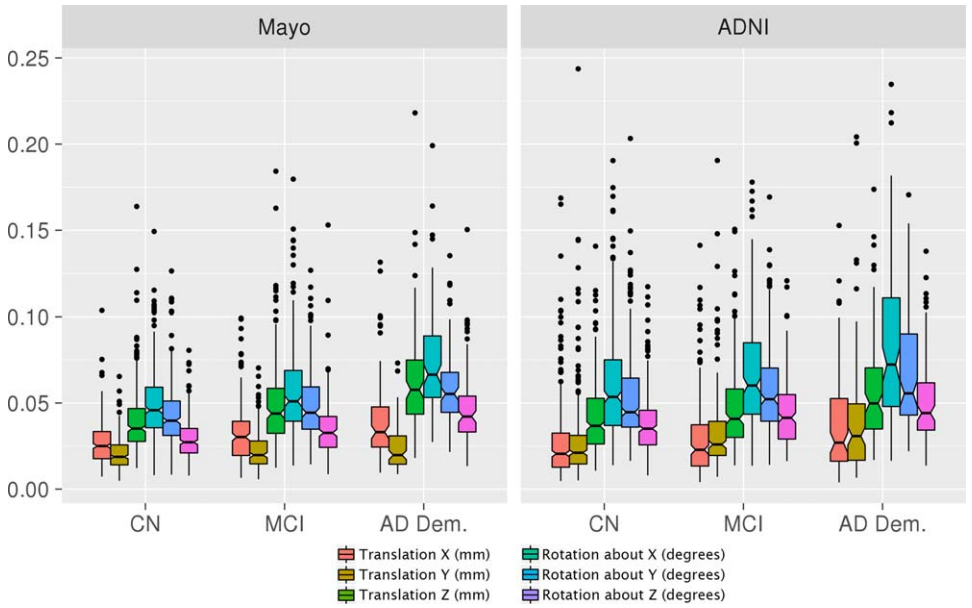

Figure 3.

Standard deviation of registration parameters across 50 random perturbations of each subject's input PET scan. Translation parameters are given in millimeters. Rotation parameters are given in degrees. Non-overlap of notched boxplot areas indicates significant differences in median values, with 95% confidence. In both datasets/tracers, imprecision in all parameters increased with clinical disease severity, i.e. tracer uptake. [Color figure can be viewed at wileyonlinelibrary.com]

algorithm. This matrix product is the result of applying the transform from the random perturbation followed by the transform from the registration that attempted to undo this perturbation. Ideally, this matrix would be identity, but in practice it gives the positional deviation caused by the perturbation and re-registration steps. We then extracted the 6 translation and rotation parameters from this matrix using *spm_imatrix()*. The standard deviation of each parameter across all 50 iterations was defined as its jitter or imprecision. This value measures the variation in ending locations (in world coordinate space) returned by the registration algorithm across the set of perturbed starting locations, after multiplying-out the inverse of the perturbation transforms themselves. We plot these jitter values for each parameter in Figure 3.

In Figure 3, there was a trend where the jitter in all parameters increased with clinical disease severity. In the Mayo dataset, differences between the CN and AD dementia groups were significant for all parameters except for *Y* translation. In the ADNI dataset, differences between these two groups were significant for all parameters except for *X* translation. Therefore, registrations for subjects with AD dementia were significantly less precise than for control subjects. Among the translation parameters, the *Z* (superior–inferior) direction had the least precision in registration. Among the rotation parameters, rotation about the *X* axis (pitch, the direction of the subject's nose moving up/down) had the least precision. These results are unfortunate because they are the directions in which subject positioning typically has the least physical restriction, and thus the most variance. Rigid registration corrects for these variations in subject positioning, but with less precision than we expected. Overall, registration was more precise in estimating translation parameters than rotation parameters.

It is important to note that each of the 50 registrations produced per scan pair was qualitatively valid. The jitter for each parameter (Fig. 3) was typically smaller than 0.1 mm (translation) or 0.1° (rotation). This level of variation was extremely small such that a human observer could barely see any variation (movement) when observing animations of the 50 post-perturbation registered PET scans resampled in the common space of the corresponding MRI. Because all registrations were visually similar, observers would not select any as comparatively better or worse, or as failures, in an image registration QC paradigm. We provide in Figure 4, a typical example of the variation in PET image voxels included when automatically localizing ROIs used for SUVR computation. From this heat map of a typical Mayo subject with AD Dementia, it is apparent that although no perturbed registration was a gross failure, this variation in parameters is sufficient to cause small inconsistencies in which voxels were included in the ROIs that form SUV ratios. In following sections, we examine the effects of this imprecision on computed SUVR values.

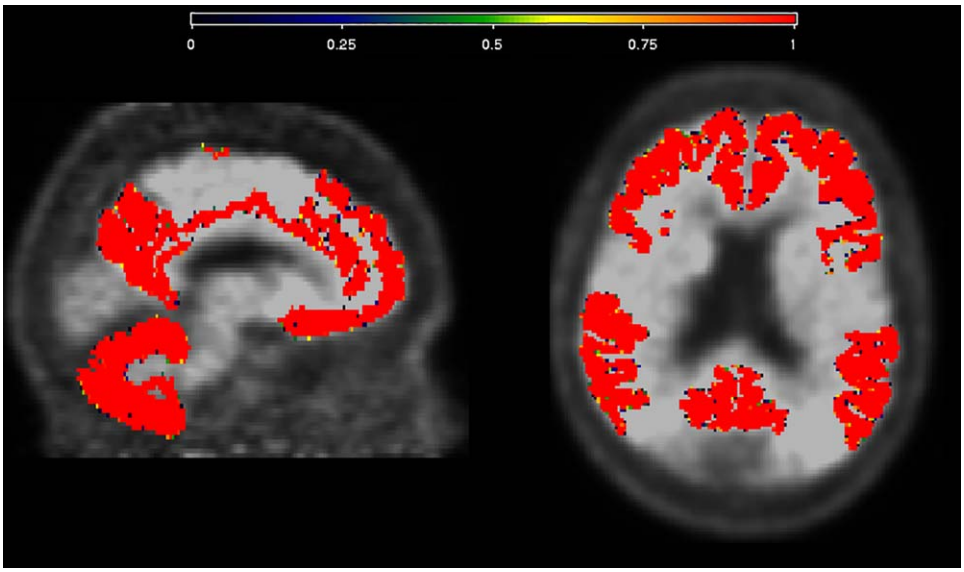

Figure 4.

Heat map of voxels included in cortical target and cerebellar GM reference ROIs across 50 perturbed registrations for a typical example subject. Red voxels (those in the color corresponding to a ratio of 1) were included in the SUVR computation ROIs located using all 50 perturbed PET-MRI registrations. Voxels in cooler colors were included in a smaller fraction of registrations, i.e. were less frequently located within these quantification ROIs. Variations in SUVR values across perturbations are due to the varying presence/absence of these voxels within the quantification ROIs as a result of PET-MRI rigid registration imprecision. [Color figure can be viewed at wileyonlinelibrary.com]

Imprecision in PET-MRI Rigid Registration is a Source of Imprecision in SUVR Measurements that Increases with Clinical Disease Severity

In Figure 5, we plot the coefficient of variation (CV) of measured SUVR values across the 50 random rigid perturbations applied to each PET scan in corresponding PET-MRI scan pairs prior to their co-registration. Because CV = (standard deviation/mean), CV of SUVR values are comparable even if the SUVR values have different ranges (such as for different ligands, reference regions, etc.). For results in this section, we used the standard SUVR method: a cerebellar GM reference region, without PVC. We present analyses with different combinations in later sections. In both datasets, imprecision in SUVR increased with clinical disease severity. Differences between all groups were significant in the Mayo dataset, while differences between the CN and MCI groups, and between the CN and AD Dementia groups, were significant in the ADNI dataset. This was expected from the previous section: imprecision in rigid registration parameters increased with clinical severity, so it naturally follows that SUVR measurements using these registrations would also become less precise. We plot the same data on a continuous axis against mean SUVR values in Figure S2 (Supporting Information).

The mean CV across all Mayo CN subjects was ≈0.00042, that is, the standard deviation across all registration perturbations was ≈0.042% of their current amyloid burden. For Mayo AD dementia subjects, this value was ≈0.11%. Previous reproducibility studies have estimated test-retest variability for similar SUVR measurements from PiB scans as $4.4 \pm 4.2\%$ for CN subjects and $8.0 \pm 7.0\%$ for AD dementia subjects [Tolboom et al., 2009]; our data suggests that a fraction of this (≈1%) may be due to registration imprecision, and that differences in registration imprecision between groups may contribute to the increased measurement variability in AD dementia.

When calculating change over time by subtracting independent serial PiB measurements, these errors become larger. For example, the average standard deviation (SD) of SUVR values across Mayo AD dementia subjects was ≈0.0022. When measuring SUVR change by subtracting independently-measured values from two time points, these errors propagate quadratically, yielding a SD of ≈0.0032. The median annual rate of change in PiB PET SUVR in MCI/AD-dementia subjects is approximately 0.048 SUVR units/year [Jack et al., 2013]. In a hypothetical but reasonable study with a 1-year follow-up period, the error due to imprecision in rigid registration for PiB SUVR change-over-time measurements of AD dementia subjects would be ≈6.58% of the expected annual rate of change. Therefore, even though rigid registration imprecision

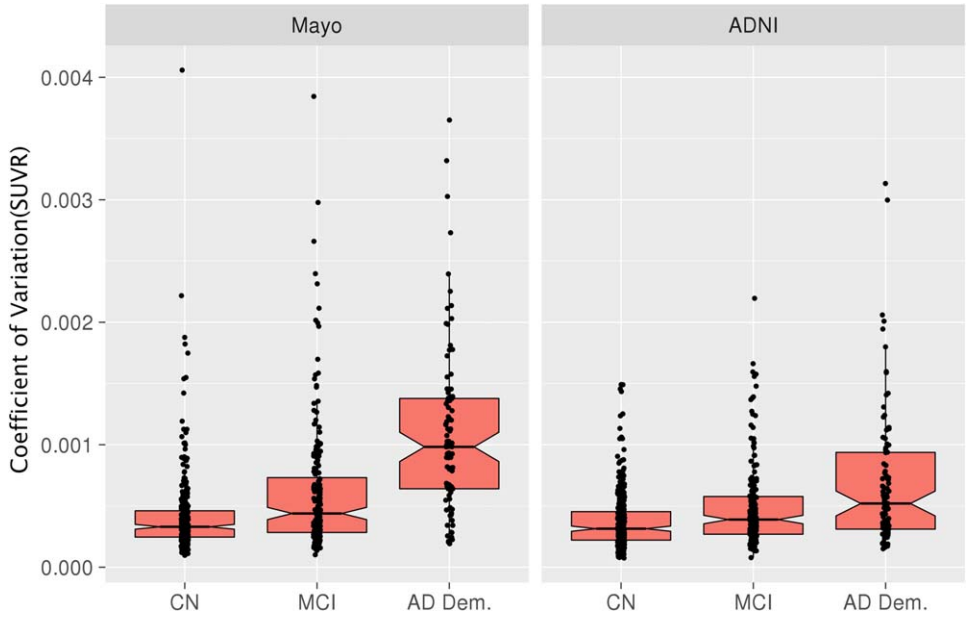

Figure 5.

Coefficient of variation (CV) of measured SUVR values (using the cerebellar GM reference region) across 50 random rigid perturbations of PET scans applied prior to rigid registration, separated by clinical disease severity. Non-overlap of notched boxplot areas indicates significant differences in median values, with 95% confidence. In both datasets/tracers, imprecision in SUVR due to imprecision in PET-MRI rigid registration increased with clinical disease severity, i.e. tracer uptake. [Color figure can be viewed at wileyonlinelibrary.com]

contributes only ≈1% of the total PiB SUVR test-retest imprecision, this fraction is not negligible when compared to the small annual change magnitudes that are of interest in amyloid PET studies.

These estimates are similar for ADNI florbetapir scans. The average standard deviation across ADNI AD dementia subjects was ≈0.0011. When subtracting values, this propagates quadratically for a SD of ≈0.0016. The annual rate of change for ADNI florbetapir scans of MCI/AD-dementia subjects is ≈1.4% or ≈0.022 SUVR units/year [Landau et al., 2015]. Therefore, the error due to imprecision in rigid registration for ADNI florbetapir change-over-time measurements of AD dementia subjects is ≈7.08% of the expected annual rate of change. Previous works have shown implausible within-subject trajectories that decrease over time in 20-25% of subjects [Landau et al., 2015; Schwarz et al., 2016b] when using similar cerebellar GM reference regions. Our data suggest that reducing this registration imprecision could potentially reduce the incidence of these implausible measurements.

Imprecision in SUVR Measurements Due to PET-MRI Rigid Registration is Affected by Choice of Reference Region

We present boxplots of CV's of SUVR values with different typical choices of reference regions in Figure 6. For all reference regions, the same set of perturbations was applied prior to registration, so differences in CV across them are only attributable only to stability of the ROI's mean values across variation in imprecise registrations.

In both datasets, SUVRs using whole-cerebellum references had a lower CV, that is, superior stability under the same varying registrations, than those using cerebellar GM references. This difference was significant in the Mayo AD Dementia group, and in the ADNI CN and MCI groups. This greater stability of whole-cerebellum measurements offers a potential, partial explanation for previous comparisons favoring whole cerebellar over cerebellar GM references [Joshi et al., 2015; Schwarz et al., 2016b]. Pontine references had the worst resistance to registration imprecision among all reference ROIs in all groups, and this difference was significant in all groups except for the Mayo AD Dementia group. Supratentorial white matter (SWM) reference ROIs had the best stability in all groups, and this difference vs. cerebellar ROIs was significant for all groups except for the ADNI AD Dementia group.

This relative resistance to variance in registration may partly explain previous comparisons favoring SWM references for longitudinal amyloid PET [Brendel et al., 2015; Landau et al., 2015; Schwarz et al., 2016b].

The composite reference, including all voxels from the supratentorial WM, pons, and the whole cerebellum, had performance in between that of the SWM (better) and the

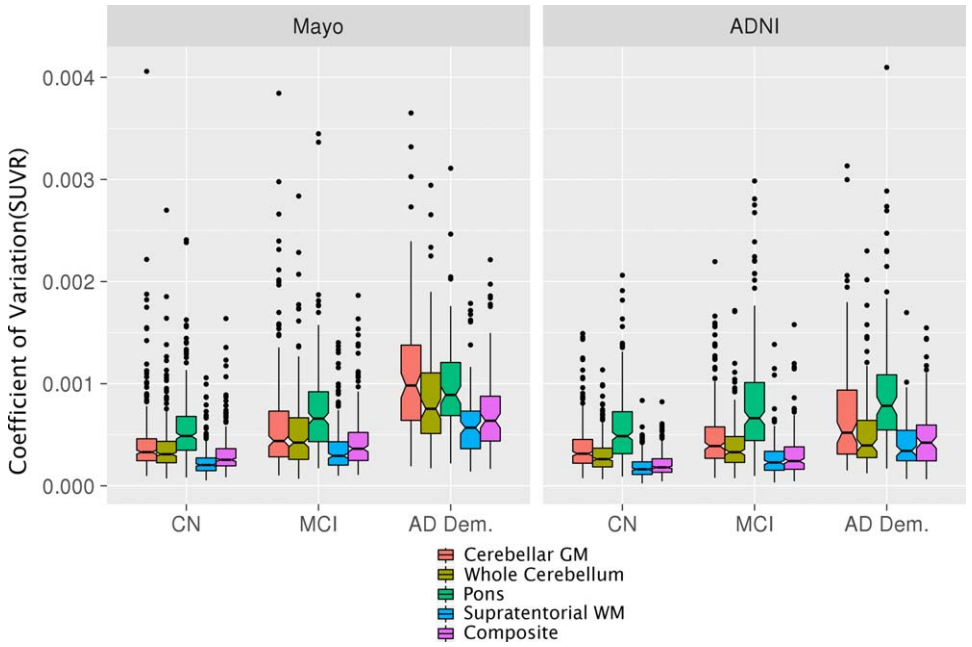

Figure 6.

CV of measured SUVR values when using five different classes of reference region across 50 random rigid perturbations of PET scans applied prior to rigid registration, plotted for each clinical diagnostic severity group. Non-overlap of notched boxplot areas indicates significant differences in median values, with 95% confidence. Supratentorial white matter reference ROIs had the best resistance to PET-MRI rigid registration imprecision, in all groups. [Color figure can be viewed at wileyonlinelibrary.com]

whole cerebellum (worse). This result is not surprising: the superior performance of the SWM was reduced by inclusion of the cerebellar and pontine ROIs. Still, the composite reference's resistance to registration imprecision, relative to cerebellar and pontine references, could partly explain its superior performance in some prior comparisons of longitudinal measurements [Schwarz et al., 2016b].

We emphasize that we are not advocating for the use of a particular reference region based on these results. In particular, we are not advocating for the use of differing reference regions for different subjects according to clinical severity or amount of tracer uptake. These results are only intended to quantify the amount of imprecision in SUVR measurements under a variety of conditions. Choice of reference region for amyloid PET SUVR measurements should depend on multiple factors; however, these results provide a potential mechanism to explain some results from earlier comparisons [Brendel et al., 2015; Chen et al., 2015; Landau et al., 2015; Schwarz et al., 2016a].

Cerebellar Reference Regions that are Eroded or Located Further from Target Tissue are Not More Robust to PET-MRI Rigid Registration Imprecision

We also hypothesized that ROIs located further from anatomic boundaries would be more resistant to registration imprecision. Specifically, we hypothesized that eroding cerebellar ROIs, that is, using only those voxels located in the center of homogeneous regions, would make ROI mean values more stable when their boundaries were imprecisely localized. We further hypothesized that SUVRs using cerebellar reference ROIs would be especially imprecise because registration imprecision could cause voxels in the occipital cortex to be included in the reference. To test these hypotheses, we introduced the "bottom" cerebellar variants using only cerebellar GM voxels located inferior to the cerebellar WM, that is, only those furthest from the tentorium. We also introduced "eroded" cerebellar variants, where morphological erosion was performed on the standard cerebellar GM reference, and on the "bottom" cerebellar GM reference, with a radius of 5 voxels. We then repeated the experiment from the previous section, comparing these with the standard cerebellar variants.

In both datasets, we found that these smaller, alternative cerebellar references had significantly higher CVs (i.e., significantly more imprecision in SUVR values under the same amount of antecedent rigid registration imprecision) than the standard whole cerebellum and cerebellar GM references (Fig. 7). This data suggest that smaller reference ROIs located further from anatomic boundaries and/or further from the tentorium do not yield more stable SUVR values. This may be because these regions are smaller, making their mean values less precise. The poor performance of cerebellar ROIs located further from the tentorium may be because

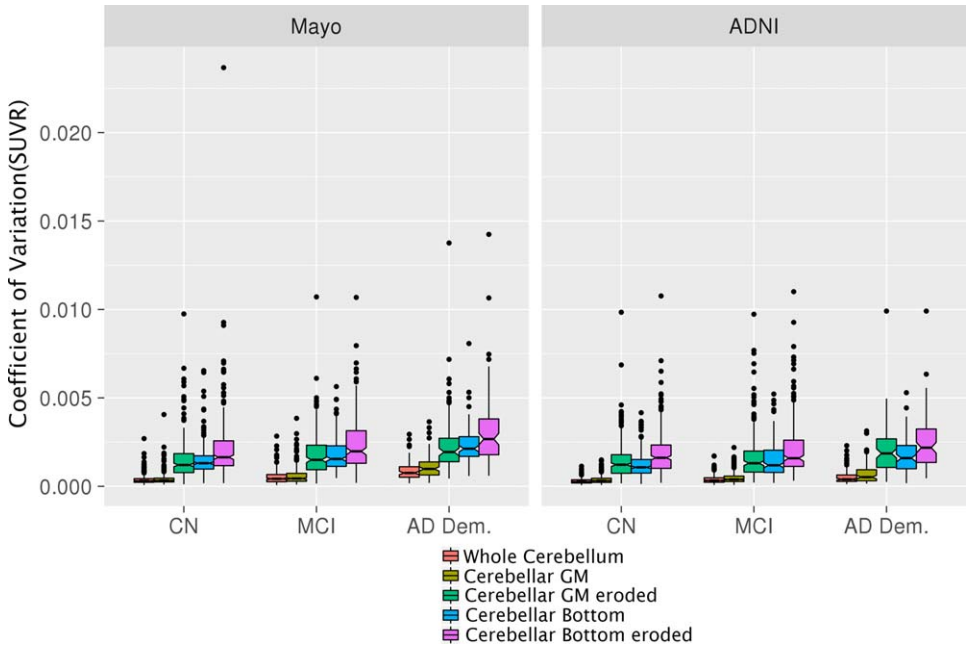

Figure 7.

CV of measured SUVR values when using cerebellar variant reference regions across 50 random rigid perturbations of PET scans applied prior to rigid registration, plotted by clinical diagnostic severity group. Non-overlap of notched boxplot areas indicates significant differences in median values, with 95% confidence. The smaller cerebellar variant reference ROIs located further from anatomic boundaries and/or further from the tentorium were relatively less resistant to imprecision in PET-MRI rigid registration. [Color figure can be viewed at wileyonlinelibrary.com]

these regions are further from the center of the PET camera, thus having a lower image signal-to-noise ratio and less-stable mean values. Previous works have proposed that the smaller size and peripheral location of cerebellar ROIs may explain their worse performance vs. SWM ROIs [Chen et al., 2015; Landau et al., 2015]. It is also plausible that more-peripheral ROIs would be more severely displaced by identical rotations than those nearer to the center of rotation.

Imprecision in SUVR Measurements Due to PET-MRI Rigid Registration is Larger When Using PVC

We show in Figure 8, the same plot as in Figure 6, but with additional corresponding variants using two-class voxel-based PVC [Meltzer et al., 1990]. Variants using PVC consistently had higher average CV values than those corresponding variants not using PVC, for all clinical groups, all reference regions, and both datasets. These differences were significant for almost all pairs in the Mayo data (exceptions: MCI whole-cerebellum, AD Dementia Supra-WM/Composite). In the ADNI group, these differences were mostly not significant.

These results are not surprising because PVC inherently relies on accurate PET-MRI registration, using each voxel's MRI segmentation to estimate the tissue fraction in PET voxels. Sensitivity of PVC methods to errors in PET-MRI registration has been reported since their inception [Meltzer et al., 1990]. We emphasize that it would be improper to conclude from this data that PVC should not be used; however, SUVR pipelines using PVC should especially prioritize precision when choosing a registration method.

Conclusions

In this work, we have presented a series of experiments to measure imprecision in PET-MRI rigid registration of amyloid PET scans and its effects on SUVR calculations. From these experiments, we have presented four major conclusions: (1) Imprecision in PET-MRI rigid registration contributes ≈1% of the total test-retest imprecision in amyloid PET SUVR measurements, and this fraction causes subtractive change-over-time measurements to be imprecise by ≈7% of the expected reference values for annual change in subjects with AD dementia (≈6.58% for PiB, ≈7.08% for florbetapir). (2) Imprecision in PET-MRI rigid registration increases with amyloid load and thus indirectly with clinical disease severity. (3) Imprecision in SUVR measurements due to PET-MRI rigid registration is affected by choice of reference region, and supratentorial WM references have relatively superior resistance to registration imprecision. (4) Imprecision in SUVR measurements due to PET-MRI rigid registration is larger when using PVC.

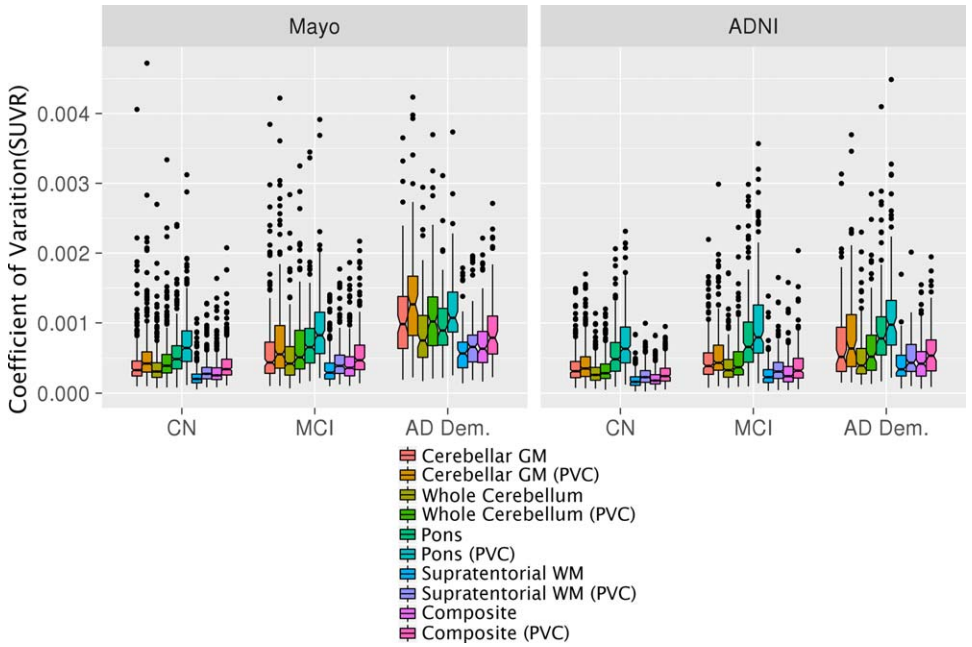

Figure 8.

CV of measured SUVR values across 50 random rigid perturbations of PET scans applied prior to rigid registration, plotted by clinical diagnostic severity group, using five different reference regions, with and without two-class partial volume correction. Non-overlap of notched boxplot areas indicates significant differences in median values, with 95% confidence. Variants using PVC consistently had worse resistance to PET-MRI rigid registration imprecision than those corresponding variants not using PVC. [Color figure can be viewed at wileyonlinelibrary.com]

We emphasize that our conclusions are not intended to suggest that particular processing choices of reference region, PVC, etc. be chosen. Instability of the measurements as a result of imprecision in the antecedent rigid registration is only one factor in such methodological choices, and these are better left to more thorough comparisons using multiple criteria [Brendel et al., 2015; Chen et al., 2015; Landau et al., 2015; Schwarz et al., 2016b]. Instead, this work is designed to highlight that PET-MRI rigid registration is a previously-underestimated source of amyloid PET SUVR imprecision that could be addressed by software algorithm improvements to improve the statistical power of studies using previously-acquired or future data. It suggests that software pipelines that analyze change in serial amyloid PET scans by simultaneously using data across serial scans to improve per-timepoint precision (potentially at the cost of per-timepoint accuracy) could reduce this source of imprecision in change-over-time measurements. Future works will focus on development and validation of such pipelines.

STRENGTHS AND LIMITATIONS OF THIS STUDY

This study's strength is in quantifying the imprecision in amyloid PET SUVR measurements due to imprecision in PET-MRI registration, using large datasets from two different tracers (PiB and AV45), under a variety of common methodological choices. Relatively little attention has been paid to accurate PET-MRI registration in amyloid PET SUVR quantification. Our study suggests that this step contributes to measurement imprecision, which is particularly important for studies of change over time. These findings also offer a potential explanation for previous reports of decreased test-retest reliability in AD dementia subjects [Tolboom et al., 2009]. It has been reported that intra-scan motion is typically higher for AD dementia subject PET scans than for controls [Ikari et al., 2012], which is another source of variability that would only compound this issue. Our findings show that worse reliability in AD dementia subjects cannot be attributed solely to increased motion.

Although *spm_coreg* is only one of many tools that could be applied for PET-MRI registration (e.g., tools in FSL, FreeSurfer, ANTs), it is popular for amyloid PET studies [Edison et al., 2013; Klunk et al., 2015; Landau and Jagust, 2015; Lowe et al., 2009; Ng et al., 2007; Schain et al., 2014; Schwarz et al., 2016a; Zhou et al., 2007] . PET-MRI registration performance likely differs across these methods, but our goal was to examine this issue as it applies to common measurement methods of SUVR. Our results should not be interpreted to suggest that *spm_coreg* is

unsuitable; comparisons between it and others are needed to assess its relative performance.

The upcoming 6.0 release of the popular FreeSurfer package also includes a PET processing pipeline using *mri_coreg*, based on *spm_coreg* [Greve, 2016a]. We repeated our simulations using *mri_coreg* with its recommended settings [Greve, 2016a] in place of *spm_coreg*. Although there were differences between *mri_coreg* and the original *spm_coreg*, both in algorithmic features and in the values produced for identical input pairs, we found no qualitative differences in imprecision between the methods, and all major findings were replicated with *mri_coreg* (data not presented). The recommended use of FreeSurfer 6 for longitudinal PET data is to use longitudinal FreeSurfer for corresponding serial MRI and independently register each PET scan to the nearest MRI time point [Greve, 2016b]. Therefore, we expect that SUVR values produced with FreeSurfer 6, when released, will have a similar degree of imprecision due to independent PET-MRI registrations, regardless of whether the longitudinal processing stream is used.

Several methods have been proposed for amyloid PET quantification without MRI [Fripp et al., 2008; Joshi et al., 2015; Raniga et al., 2007; Thurfjell et al., 2014; Zhou et al., 2014]. We did not focus on MRI-independent methods in this work because (1) MRI is typically acquired during PET quantification in longitudinal observational and clinical trials of Alzheimer's disease [Jagust et al., 2010; Salloway et al., 2014; Siemers et al., 2016]. (2) Partial volume correction techniques, which are generally recommended by previous comparisons using amyloid PET [Brendel et al., 2014; Rullmann et al., 2016; Schwarz et al., 2016b; Su et al., 2015; Thomas, 2012], are not possible with MRI-independent methods. MRI-independent methods do not have a rigid PET-to-MRI registration step discussed in this work, but they do necessarily have some form of PET-to-PET registration to a common template for ROI localization. It is unknown whether these registrations have a similar degree of imprecision, but we did not examine this in this work.

To measure parameter jitter (Evaluation Criteria and Statistical Methods and Imprecision in PET-MRI Rigid Registration is Larger in Rotation Parameters and Increases with Clinical Disease Severity sections), we assumed that an initial PET-MRI registration produced a reasonable minimum. This initial registration was necessary to produce a valid starting point around which to vary registration parameters. Without it, deviations around poor initial locations resulted in frequent gross registration failures. It is possible that sometimes the initial registration produced a relatively shallow local minimum, that is, some post-perturbation solutions may have been superior. This could have biased the parameter jitter results (Imprecision in PET-MRI Rigid Registration is Larger in Rotation Parameters and Increases with Clinical Disease Severity section). However, the variation-in-SUVR results (Imprecision in PET-MRI rigid registration is a Source of Imprecision in SUVR Measurements that Increases with Clinical Disease Severity–Imprecision in SUVR Measurements due to PET-MRI Rigid Registration is Larger When Using PVC sections) do not depend on this assumption and were consistent with the conclusions of the Imprecision in PET-MRI Rigid Registration is Larger in Rotation Parameters and Increases with Clinical Disease Severity section, suggesting that these biases were negligible. Additionally, our imprecision estimates may be conservative because our ranges of parameter perturbations (Evaluation Criteria and Statistical Methods section) were also designed to be conservative underestimates, in order to avoid gross registration failures.

Our experiments used two tracers for late-uptake amyloid PET, but did not examine other types of PET in brain imaging such as full-dynamic acquisitions, Fluorodeoxyglucose or recently developed Tau ligands such as AV-1451[Xia et al., 2013]. However, typical processing methods for these tracers also use SUVR computed via registration with MRI, so we hypothesize that SUVR calculations using these PET images would be affected similarly. We also did not examine other methods of PVC. Although all PVC methods using co-registered MRI would be affected by registration imprecision, we hypothesize that region-based methods [Rousset et al., 1998] would be affected less severely than voxel-based, such as that examined here.

ACKNOWLEDGMENTS

The authors would like to thank Heather J. Wiste for her assistance in selecting subjects for inclusion in this analysis. Data collection and sharing for the ADNI dataset used in this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health

(www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

REFERENCES

ADNI Home (2013): Available at: www.adni-info.org. Last accessed: 04/14/2017.

Ashburner J, Friston KJ (2005): Unified segmentation. Neuroimage 26:839–851.

Avants BB, Epstein CL, Grossman M, Gee JC (2008): Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Article. Med Image Anal 12:26–41.

Brendel M, Delker A, Rötzer C, Böning G, Carlsen J, Cyran C, Mille E, Gildehaus FJ, Cumming P, Baumann K, Steiner H, Haass C, Herms J, Bartenstein P, Rominger A (2014): Impact of partial volume effect correction on cerebral β-amyloid imaging in APP-Swe mice using [18F]-florbetaben PET. Neuroimage 84:843–853.

Brendel M, Högenauer M, Delker A, Sauerbeck J, Bartenstein P, Seibyl J, Rominger A (2015): Improved longitudinal [18F]-AV45 amyloid PET by white matter reference and VOI-based partial volume effect correction. Neuroimage 108:450–459.

Carbonell F, Zijdenbos AP, Charil A, Grand'Maison M, Bedell B (2015): Optimal target region for subject classification based on amyloid PET images. J Nucl Med 56:1351–1358.

Chen K, Roontiva A, Thiyyagura P, Lee W, Liu X, Ayutyanont N, Protas H, Luo JL, Bauer R, Reschke C, Bandy D, Koeppe RA, Fleisher AS, Caselli RJ, Landau S, Jagust WJ, Weiner MW, Reiman EM (2015): Improved power for characterizing longitudinal amyloid-β PET changes and evaluating amyloid-modifying treatments with a cerebral white matter reference region. J Nucl Med 56:560–566.

Edison P, Carter SF, Rinne JO, Gelosa G, Herholz K, Nordberg A, Brooks DJ, Hinz R (2013): Comparison of MRI based and PET template based approaches in the quantitative analysis of amyloid imaging with PIB-PET. Neuroimage 70:423–433.

Fripp J, Bourgeat P, Raniga P, Acosta O, Villemagne V, Jones G, O'Keefe G, Rowe C (2008): MR-less high dimensional spatial normalization of 11C PiB PET images on a population of elderly. Mild cognitive impaired and Alzheimer disease patients. Med Image Comput Comput Interv 5241:442–449.

Greve D (2016a): FreeSurfer wiki: PETSurfer. FreeSurfer wiki. Available at: https://surfer.nmr.mgh.harvard.edu/fswiki/PetSurfer Last accessed: 04/14/2017.

Greve DN (2016b): [Freesurfer] Longitudinal surface analysis of PET data. Freesurfer mailing list. Available at: https://mail.nmr.mgh.harvard.edu/pipermail///freesurfer/2016-September/047921.html Last accessed: 04/14/2017.

Ikari Y, Nishio T, Makishi Y, Miya Y, Ito K, Koeppe RA, Senda M (2012): Head motion evaluation and correction for PET scans with 18F-FDG in the Japanese Alzheimer's disease neuroimaging initiative (J-ADNI) multi-center study. Ann Nucl Med 26:535–544.

Jack CRJ, Bernstein MA, Borowski BJ, Gunter JL, Fox NC, Thompson PM, Schuff N, Krueger G, Killiany RJ, DeCarli CS, Dale AM, Weiner MW, Carmichael OW, Tosun D, Weiner MW (2010): Update on the magnetic resonance imaging core of the Alzheimer's disease neuroimaging initiative. Article. Alzheimer's Dement 6:212–220.

Jack CRJ, Wiste HJ, Lesnick TG, Weigand SD, Knopman DS, Vemuri P, Pankratz VS, Senjem ML, Gunter JL, Mielke MM, Lowe VJ, Boeve BF, Petersen RC (2013): Brain β-amyloid load approaches a plateau. Neurology 80:890–896.

Jagust WJ, Bandy D, Chen K, Foster NL, Landau SM, Mathis CA, Price JC, Reiman EM, Skovronsky D, Koeppe RA (2010): The Alzheimer's Disease Neuroimaging Initiative positron emission tomography core. Alzheimer's Dement 6:221–229.

Joshi AD, Pontecorvo MJ, Lu M, Skovronsky DM, Mintun MA, Devous MD (2015): A semiautomated method for quantification of F 18 florbetapir PET images. J Nucl Med 56:1736–1741.

Joshi A, Koeppe RA, Fessler JA (2009): Reducing between scanner differences in multi-center PET studies. Neuroimage 46:154–159.

Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, Bergström M, Savitcheva I, Huang GF, Estrada S, Ausén B, Debnath ML, Barletta J, Price JC, Sandell J, Lopresti BJ, Wall A, Koivisto P, Antoni G, Mathis CA, Långström B (2004): Imaging brain amyloid in Alzheimer's disease with Pittsburgh compound-B. Ann Neurol 55:306–319.

Klunk WE, Koeppe RA, Price JC, Benzinger TL, Devous MD, Jagust WJ, Johnson KA, Mathis CA, Minhas D, Pontecorvo MJ, Rowe CC, Skovronsky DM, Mintun MA (2015): The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. Alzheimer's Dement 11:1–15.

Landau S, Jagust W (2015): ADNI Florbetapir processing methods. ADNI PET Image Analysis. Available at: https://ida.loni.usc.edu/ Last accessed: 04/14/2017.

Landau SM, Fero A, Baker SL, Koeppe R, Mintun M, Chen K, Reiman EM, Jagust WJ (2015): Measurement of longitudinal B-amyloid change with 18F-florbetapir PET and standardized uptake value ratios. J Nucl Med 56:567–574.

Lowe VJ, Kemp BJ, Jack CR, Senjem M, Weigand S, Shiung M, Smith G, Knopman D, Boeve B, Mullan B, Petersen RC (2009): Comparison of 18F-FDG and PiB PET in cognitive impairment. J Nucl Med 50:878–886.

McGill R, Tukey JW, Larsen WA (1977): Variations of box plots. Am Stat 32:12–16.

Meltzer CC, Leal JP, Mayberg HS, Wagner HNJ, Frost JJ (1990): Correction of PET data for partial volume effects in human cerebral cortex by MR imaging. J Comput Assist Tomogr 14: 561–570.

Ng S, Villemagne VL, Berlangieri S, Lee S-T, Cherk M, Gong SJ, Ackermann U, Saunder T, Tochon-Danguy H, Jones G, Smith C, O'Keefe G, Masters CL, Rowe CC (2007): Visual assessment versus quantitative assessment of 11C-PIB PET and 18F-FDG PET for detection of Alzheimer's disease. J Nucl Med 48: 547–552.

Oishi K, Faria A, Jiang H, Li X, Akhter K, Zhang J, Hsu JT, Miller MI, van Zijl PCM, Albert M, Lyketsos CG, Woods R, Toga AW, Pike GB, Rosa-Neto P, Evans A, Mazziotta J, Mori S (2009): Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: Application to normal elderly and Alzheimer's disease participants. Neuroimage 46:486–499.

Petersen RC, Roberts RO, Knopman DS, Geda YE, Cha RH, Pankratz VS, Boeve BF, Tangalos EG, Ivnik RJ, Rocca WA (2010): Prevalence of mild cognitive impairment is higher in men. The Mayo Clinic Study of Aging. Neurology 75:889–897.

Raniga P, Bourgeat P, Ourselin S, Villemagne V, O'Keefe G, Rowe C (2007): PIB-PET segmentation for automatic SUVR

normalisation without MR information. Biomed Imaging, Int Symp 348–351.

Roberts RO, Geda YE, Knopman DS, Cha RH, Pankratz VS, Boeve BF, Ivnik RJ, Tangalos EG, Petersen RC, Rocca WA (2008): The Mayo Clinic Study of Aging: design and sampling, participation, baseline measures and sample characteristics. Neuroepidemiology 30:58–69.

Rousset OG, Ma Y, Evans AC (1998): Correction for partial volume effects in PET: principle and validation. J Nucl Med 39: 904–911.

Rullmann M, Dukart J, Hoffmann K-T, Luthardt J, Tiepolt S, Patt M, Gertz H-J, Schroeter ML, Seibyl J, Schulz-Schaeffer WJ, Sabri O, Barthel H (2016): Partial-volume effect correction improves quantitative analysis of 18F-florbetaben-amyloid PET scans. J Nucl Med 57:198–203.

Salloway S, Sperling R, Fox NC, Blennow K, Klunk W, Raskind M, Sabbagh M, Honig LS, Porsteinsson AP, Ferris S, Reichert M, Ketter N, Nejadnik B, Guenzler V, Miloslavsky M, Wang D, Lu Y, Lull J, Tudor IC, Liu E, Grundman M, Yuen E, Black R, Brashear HR (2014): Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. N Engl J Med 370:322–333.

Schain M, Varnäs K, Cselényi Z, Halldin C, Farde L, Varrone A (2014): Evaluation of two automated methods for PET region of interest analysis. Neuroinformatics 12:551–562.

Scheinin NM, Aalto S, Koikkalainen J, Lötjönen J, Karrasch M, Kemppainen N, Viitanen M, Någren K, Helin S, Scheinin M, Rinne JO (2009): Follow-up of [11C]PIB uptake and brain volume in patients with Alzheimer disease and controls. Neurology 73:1186–1192.

Schmidt ME, Chiao P, Klein G, Matthews D, Thurfjell L, Cole PE, Margolin R, Landau S, Foster NL, Mason NS, De Santi S, Suhy J, Koeppe RA, Jagust W (2015): The influence of biological and technical factors on quantitative analysis of amyloid PET: Points to consider and recommendations for controlling variability in longitudinal data. Alzheim Dement 11:1050–1068.

Schwarz CG, Gunter JL, Wiste HJ, Przybelski SA, Weigand SD, Ward CP, Senjem ML, Vemuri P, Murray ME, Dickson DW, Parisi JE, Kantarci K, Weiner MW, Petersen RC, Jack CRJ (2016a): A large-scale comparison of cortical thickness and volume methods for measuring Alzheimer's disease severity. NeuroImage Clin 11:802–812.

Schwarz CG, Senjem ML, Gunter JL, Tosakulwong N, Weigand SD, Kemp BJ, Spychalla AJ, Vemuri P, Petersen RC, Lowe VJ, Jack CRJ (2016b): Optimizing PiB-PET SUVR Change-Over-Time Measurement by a large-scale analysis of Longitudinal Reliability, Plausibility, Separability, and Correlation with MMSE. Neuroimage 144:113–127.

Siemers ER, Sundell KL, Carlson C, Case M, Sethuraman G, Liu-Seifert H, Dowsett SA, Pontecorvo MJ, Dean RA, DeMattos R (2016): Phase 3 solanezumab trials: Secondary outcomes in mild Alzheimer's disease patients. Alzheim Dement 12:110–120.

Su Y, Blazey TM, Snyder AZ, Raichle ME, Marcus DS, Ances BM, Bateman RJ, Cairns NJ, Aldea P, Cash L, Christensen JJ, Friedrichsen K, Hornbeck RC, Farrar AM, Owen CJ, Mayeux R, Brickman AM, Klunk W, Price JC, Thompson PM, Ghetti B, Saykin AJ, Sperling RA, Johnson KA, Scho PR, Buckles V, Morris JC, Benzinger TLS, Alzheimer I (2015): Partial volume correction in quantitative amyloid imaging. Neuroimage 107:55–64.

Thomas BA (2012): Improved brain PET quantification using partial volume correction techniques. Ph.D. Thesis; University College London.

Thurfjell L, Lilja J, Lundqvist R, Buckley C, Smith A, Vandenberghe R, Sherwin P (2014): Automated quantification of 18F-Flutemetamol PET activity for categorizing scans as negative or positive for brain amyloid: concordance with visual image reads. J Nucl Med 55:1623–1628.

Tolboom N, Yaqub M, Boellaard R, Luurtsema G, Windhorst AD, Scheltens P, Lammertsma AA, Van Berckel BNM (2009): Test-retest variability of quantitative [11C]PIB studies in Alzheimer's disease. Eur J Nucl Med Mol Imaging 36:1629–1638.

Wong DF, Rosenberg PB, Zhou Y, Kumar A, Raymont V, Ravert HT, Dannals RF, Nandi A, Brašić JR, Ye W, Hilton J, Lyketsos C, Kung HF, Joshi AD, Skovronsky DM, Pontecorvo MJ (2010): In vivo imaging of amyloid deposition in Alzheimer disease using the radioligand 18F-AV-45 (Flobetapir F 18). J Nucl Med 51:913–920.

Xia C-F, Arteaga J, Chen G, Gangadharmath U, Gomez LF, Kasi D, Lam C, Liang Q, Liu C, Mocharla VP, Mu F, Sinha A, Su H, Szardenings AK, Walsh JC, Wang E, Yu C, Zhang W, Zhao T, Kolb HC (2013): [(18)F]T807, a novel tau positron emission tomography imaging agent for Alzheimer's disease. Alzheim Dement 9:666–676.

Zasadny KR, Wahl RL (1993): Standardized uptake values of normal tissues at PET with 2-[fluorine-18]-fluoro-2-deoxy-D-glucose: variations with body weight and a method for correction. Radiology 189:847–850.

Zhou L, Salvado O, Dore V, Bourgeat P, Raniga P, Macaulay SL, Ames D, Masters CL, Ellis KA, Villemagne VL, Rowe CC, Fripp J (2014): MR-less surface-based amyloid assessment based on 11C PiB PET. PLoS One 9:e84777.

Zhou Y, Resnick SM, Ye W, Fan H, Holt DP, Klunk WE, Mathis CA, Dannals R, Wong DF (2007): Using a reference tissue model with spatial constraint to quantify [11C]Pittsburgh compound B PET for early diagnosis of Alzheimer's disease. Neuroimage 36:298–312.