

## RESEARCH ARTICLE

## Amyloid burden quantification depends on PET and MR image processing methodology

Guilherme D. Kolinger<sup>1\*</sup>, David Vález García<sup>1</sup>, Antoon T. M. Willemsen<sup>1</sup>, Fransje E. Reesink<sup>2</sup>, Bauke M. de Jong<sup>2</sup>, Rudi A. J. O. Dierckx<sup>1</sup>, Peter P. De Deyn<sup>2,3</sup>, Ronald Boellaard<sup>1,4</sup>

**1** Medical Imaging Center, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands, **2** Department of Neurology, Alzheimer Research Centre, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands, **3** Laboratory of Neurochemistry and Behaviour, Institute Born-Bunge, University of Antwerp, Antwerp, Belgium, **4** Department of Radiology and Nuclear Medicine, Amsterdam University Medical Center, VU Medical Center, Amsterdam, The Netherlands

\* [g.domingues.kolinger@umcg.nl](mailto:g.domingues.kolinger@umcg.nl)



## Abstract

Quantification of amyloid load with positron emission tomography can be useful to assess Alzheimer's Disease *in-vivo*. However, quantification can be affected by the image processing methodology applied. This study's goal was to address how amyloid quantification is influenced by different semi-automatic image processing pipelines. Images were analysed in their *Native Space* and *Standard Space*; non-rigid spatial transformation methods based on maximum a posteriori approaches and tissue probability maps (TPM) for regularisation were explored. Furthermore, grey matter tissue segmentations were defined before and after spatial normalisation, and also using a population-based template. Five quantification metrics were analysed: two intensity-based, two volumetric-based, and one multi-parametric feature. Intensity-related metrics were not substantially affected by spatial normalisation and did not significantly depend on the grey matter segmentation method, with an impact similar to that expected from test-retest studies ( $\leq 10\%$ ). Yet, volumetric and multi-parametric features were sensitive to the image processing methodology, with an overall variability up to 45%. Therefore, the analysis should be carried out in *Native Space* avoiding non-rigid spatial transformations. For analyses in *Standard Space*, spatial normalisation regularised by TPM is preferred. Volumetric-based measurements should be done in *Native Space*, while intensity-based metrics are more robust against differences in image processing pipelines.

## OPEN ACCESS

**Citation:** Kolinger GD, Vález García D, Willemsen ATM, Reesink FE, de Jong BM, Dierckx RAJO, et al. (2021) Amyloid burden quantification depends on PET and MR image processing methodology. PLOS ONE 16(3): e0248122. <https://doi.org/10.1371/journal.pone.0248122>

**Editor:** Pierpaolo Alongi, Fondazione Istituto G. Giglio di Cefalu, ITALY

**Received:** December 18, 2020

**Accepted:** February 19, 2021

**Published:** March 5, 2021

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0248122>

**Copyright:** © 2021 Kolinger et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are available at the Zenodo repository (<https://doi.org/10.5281/zenodo.4353853>).

## Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disease and the most common pathology leading to dementia in the elderly. Past [1] and current [2] guidelines provide means for the diagnosis of probable AD via clinical assessment of symptomatic patients. While proved pathological AD is demonstrated via *post-mortem* assessment [2,3], a research framework was proposed in which AD is defined by the pathological process underlying the patient's

**Funding:** GDK, DVG, and RB have received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie [grant agreement No 764458].

**Competing interests:** The authors have declared that no competing interests exist.

symptoms [4]. This definition is achieved with an *in-vivo* assessment of AD biomarkers, where positron emission tomography (PET) imaging plays an important role in the quantification of abnormal  $\beta$ -amyloid ( $A\beta$ ) plaques depositions in the brain [5–7], one of the hallmarks of AD [8]. With the combination of PET/computed tomography (CT) with magnetic resonance imaging (MRI), it is possible to obtain detailed regional quantitative information about AD biomarkers in the brain. This hybrid imaging approach provides powerful means for the characterisation of early-stage AD and disease progression [9–11].

PET imaging can be assessed using fully quantitative methods applied to dynamic scans or semi-quantitatively with standardized uptake value ratios (SUVR) in static images that have shown to achieve high diagnostic accuracy, success for early disease detection, and differential diagnosis of AD [5,6,12–17]. While PET imaging provides quantitative data about the biological process under investigation, the anatomical volumes of interest cannot always easily be defined on the PET image due to its restricted spatial resolution and the fact that the PET signal is not always directly related to anatomy. On the other hand, MRI provides highly detailed anatomical images that can be aligned with PET images to perform a hybrid analysis, combining anatomical and functional data. Thus, it becomes possible to define anatomically derived volumes of interest on the MRI and perform PET quantification using these volumes.

Additionally, it is also possible to spatially normalise the images from the *Native Space* (defined by the MRI or the PET, and specific for each subject) to a *Standard Space*, commonly defined by the Montreal Neurological Institute (MNI) Space [18] that was created based on healthy young volunteers. With images in the *Standard Space*, it is possible to perform voxel-wise analysis to compare subjects, while regional analysis (based on anatomical volumes of interest) is possible in both *Native Space* and *Standard Space*. These spatial transformations are commonly calculated with the subject's anatomical MRI as reference. However, there is no consensus on the best approach for PET+MR image processing and only a few studies are exploring the impact of image processing approaches on PET quantification results [19–23]. Notably, there has been an initial exploration of a MR-less approach for spatial transformation of brain PET images into the MNI space using the low-dose CT acquired for attenuation correction during the PET/CT scanning [24]. However, in our study we focussed on pipelines for which both PET and MRI data are available, including the use of grey and white matter segmentations that are most commonly derived from T1-weighted MR images.

Non-rigid spatial transformations change the voxel volume, while the voxel intensities can be modulated. Some spatial transformation methods preserve the PET voxel concentrations and are expected to have less effects on voxel-wise analysis and metrics related to the activity concentration measured in the PET image. Other approaches preserve the total signal in the image and are expected to have less influence on measurements of total regional activity [25–28]. Furthermore, when carrying out analysis in *Standard Space* the tissue segmentation can be performed on the MR images before or after their non-rigid spatial transformation. This is not expected to strongly affect metrics, however, when pre-defined templates in *Standard Space* are used to delineate regions of interest on the PET images, a significant impact on metrics is expected. Unfortunately, the methodology applied to process the images and to define the volumes of interest is frequently underreported in literature and vary depending on the specific scientific question under investigation, rendering reproducibility and comparison between studies difficult.

The current study aimed to assess the influence of different image processing pipelines on quantification of amyloid burden with PET and to what extent various quantitative tracer uptake and volumetric measurements are affected by performing the feature extraction in *Native* or *Standard Space*, by using different spatial transformation methods, and by employing different grey matter tissue segmentation approaches.

## Materials and methods

### Subjects

Twenty-eight subjects underwent dynamic [ $^{11}\text{C}$ ]PiB PET/CT scans for amyloid imaging, as well as a T1-Weighted 3D MRI scan. Twelve subjects were clinically diagnosed with AD following the National Institute on Aging and Alzheimer's Association (NIA-AA) criteria [2] and classified as  $\beta$ -amyloid positive based on visual inspection of the [ $^{11}\text{C}$ ]PiB images, while sixteen subjects were healthy volunteers (HC) with no cognitive complaints and classified as  $\beta$ -amyloid negative. Healthy volunteers were recruited via advertisement on local newspapers and within the University Medical Center Groningen (UMCG). All subjects were considered competent (MMSE > 18) to give informed consent to participate in the study and a written informed consent was obtained. Subjects' demographics can be seen in Table 1. No statistically significant difference was found for age, weight, BMI, and injected activity between HC and AD subjects, while the MMSE difference was statistically significant (two-tailed t-test). The current cohort is a subset of a larger study at the memory clinic, Alzheimer Centre of the UMCG, Groningen, The Netherlands. Ethical approval was obtained from the Medical Ethical Committee of the UMCG (2014/320).

### Image acquisition and pre-processing

All PET/CT scans were performed with Siemens Biograph 40mCT or 64mCT scanners (Siemens Medical Solutions, USA) at the UMCG. Both systems were from the same vendor and generation; the acquisition and reconstruction protocols were harmonized, and the systems were (cross-)calibrated. No significant differences between the images provided by these two different scanners were found, as shown also in a previously published study using the same dataset [29]. Scans were performed under standard resting conditions with eyes closed. PET acquisition started at tracer injection and lasted at least 60 minutes. List-mode PET data was reconstructed with 3D OSEM (3i24s) including point spread function resolution modelling and time-of-flight. More details on the PET acquisition and reconstruction can be found elsewhere [30]. Using PMOD (version 3.8; PMOD Technologies LLC), the 40 to 60 minutes interval of the dynamic PET images was averaged to obtain a static PET image. MRI acquisitions were performed at the UMCG and nearby hospitals following local clinical routine 3D T1-weighted MRI acquisitions. MR images were re-sliced to a voxel size of  $1.0 \times 1.0 \times 1.0 \text{ mm}^3$ , since some clinical scans had different voxel spacing ( $0.98 \times 0.98 \times 1.0 \text{ mm}^3$ ,  $1.01 \times 1.01 \times 1.0 \text{ mm}^3$ , and  $0.49 \times 0.49 \times 1.0 \text{ mm}^3$ ).

### Native Space and Standard Space

In this study, the quantification of images was performed in two different reference spaces:

**Table 1. Subject demographics (mean  $\pm$  standard deviation).**

	AD	HC
Gender (M:F)	8:4	11:5
Age (years)	64 $\pm$ 8	69 $\pm$ 5
Weight (kg)	81 $\pm$ 14	77 $\pm$ 12
BMI	26 $\pm$ 5	25 $\pm$ 3
MMSE	24 $\pm$ 3	30 $\pm$ 1
Minimum MMSE	19	28
Injected activity (MBq)	352 $\pm$ 75	386 $\pm$ 46

<https://doi.org/10.1371/journal.pone.0248122.t001>

1. *Native Space*: the T1-MRI image was used to define the reference space of each specific subject. In this *Native Space*, all the subject's brains are anatomically unique, and they all have their individual coordinate systems.
2. *Standard Space*: the T1-MRI of each subject was used to estimate the non-rigid spatial transformations needed to align and warp each subject-specific brain image into the MNI space. In this *Standard Space*, all the subject's brains are anatomically equivalent, and they share a single common coordinate system.

### Spatial transformations

Using Statistical Parametric Mapping version 12 (SPM12, Wellcome Trust Centre for Neuroimaging, UK), static [ $^{11}\text{C}$ ]PiB images were spatially aligned to the subjects' MRI (i.e. *Native Space*) using rigid transformations with a Normalised Mutual Information objective function and linear interpolation. Then, non-linear transformations were estimated to adjust each subject's MRI to the *Standard Space* using three different methods (details on the settings used can be found in Tables A and B in [S1 File](#)). Two methods are based on a maximum a posteriori (MAP) approach [25,26] that provides the same spatial transformation parameters while treating the voxel values differently: one method modulates the voxels to preserve the total amount of signal in the image (MAPa), while the other preserves the voxel concentrations (voxel intensities) of the original image (MAPc). The third spatial transformation method is a unified segmentation method (USM) that uses tissue probability maps (TPM) as deformable spatial priors for regularisation of the non-linear deformations and that preserves the voxel concentrations of the original image [27,28].

### Volumes of interest: Grey matter and cerebellum

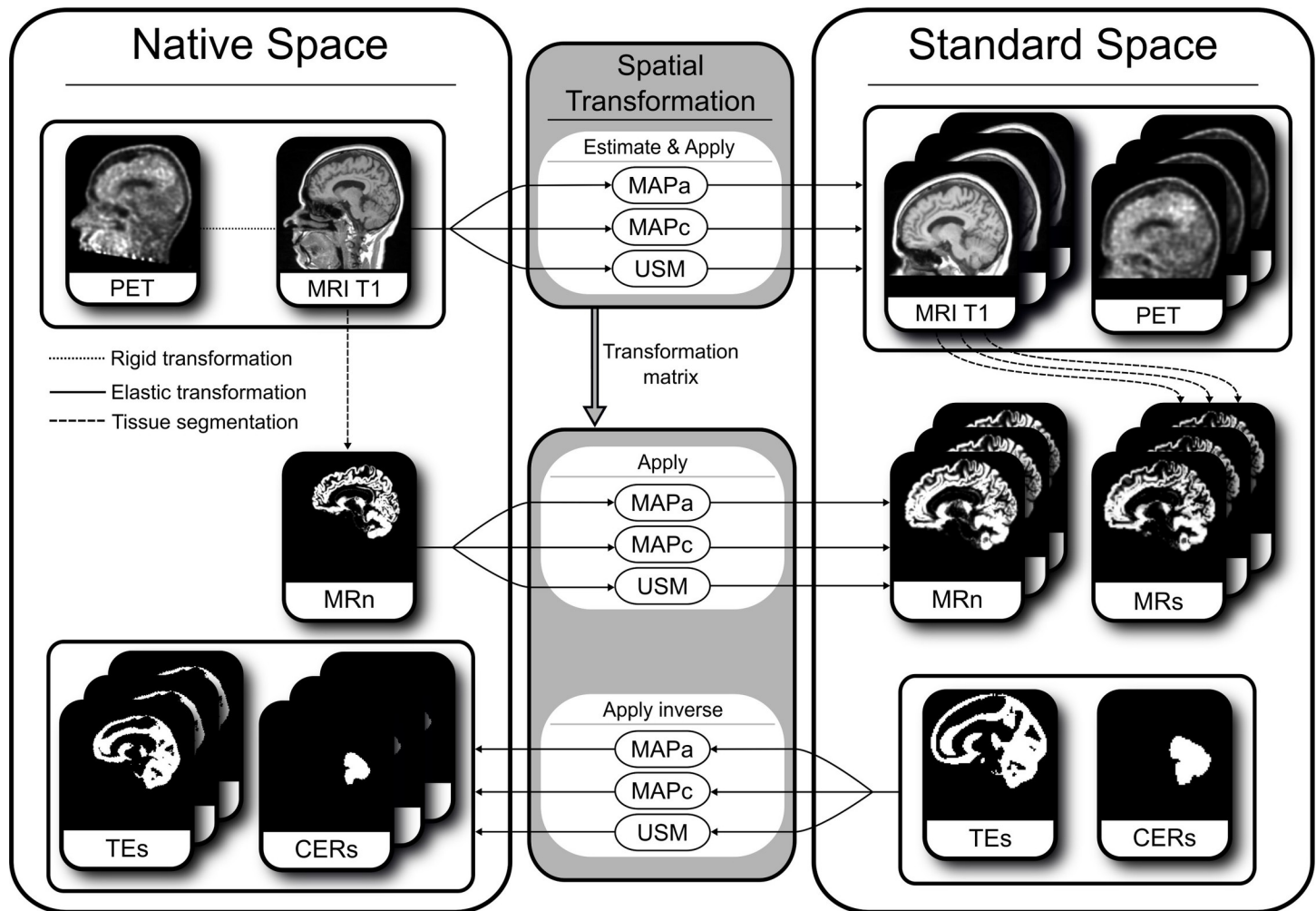
For each subject, different volumes of interest (VOI) were created for the grey matter (GM) defined by a binary mask based on a 50% threshold of the GM TPM. The GM TPM was determined by either segmenting the subject's MRI (using the tissue segmentation tool in SPM12) or by using a GM template [31]. As such, three methods for defining GM tissue were used:

1. GM TPM in *Native Space* obtained from the tissue segmentation of the subject's MRI in its *Native Space* (MRn)
2. GM TPM in *Standard Space* obtained from the subject's spatially transformed MRI to *Standard Space* (MRs)
3. GM TPM from a template available in *Standard Space* (TEs)

After the definition of GM tissue, (inverse) spatial transformations were applied to move GM TPMs between *Native* and *Standard Space* using the three different methods (MAPa, MAPc, and USM) as described above. Thus, five sets of GM tissue maps were available: GM from MRn in both *Native* and *Standard Space*, GM from MRs in *Standard Space*, and GM from TEs in both *Native* and *Standard Space*. Notice that since spatial transformations were carried out with three different methods, three GM TPMs were available after moving to a different space. In this study, the cerebellum VOI was defined based on the Hammers atlas [31], which is available in the *Standard Space* (CERs) and can be transformed into *Native Space* using each of the three inverse spatial transformations.

### Image processing pipelines

The extraction of metrics was performed for each possible combination of reference space, spatial transformation method, and GM definition. Each of these combinations will be referred



**Fig 1. Image processing pipelines scheme.** Scheme of the different image processing pipelines applied to the PET image, MRI, and grey matter (GM) tissue maps. The rigid transformation is shown by a dotted line, elastic transformations by continuous lines, and tissue segmentation by dashed lines. Images were transformed from *Native Space* into *Standard Space* using one of three possible spatial normalisation methods (upper and middle section). Spatial transformation matrices were calculated based on the MRI T1 image and these matrices were then applied to the other relevant images, while the inverse matrix was applied to transform the images from *Standard Space* to *Native Space*, for example when moving the GM derived from a standard template image or the cerebellum (CER) volume-of-interest (bottom section). Note that subscripts 's' and 'n' denote the space in which the tissue images were defined.

<https://doi.org/10.1371/journal.pone.0248122.g001>

to as an *image processing pipeline*. Fig 1 shows a diagram of the different image processing pipelines assessed in the present study. Notice that spatial transformations can impact even the results obtained in the *Native Space* since the inverse transformations are applied to move the GM and the cerebellum VOI from the *Standard Space* into the *Native Space*. Two pipelines were considered as the main/reference pipelines: *Native Space* + MRn + USM and *Standard Space* + MRn + USM. In the first main pipeline, all images and pre-processing were performed in the patient *Native Space*. The second main pipeline involves a rigid transformation of the PET and MRI in *Native Space*, a GM segmentation using USM in *Native Space* followed by a non-rigid transformation of the PET, MRI, and GM images to *Standard Space*. These main pipelines represent the most commonly used ones for PET+MR data analysis in either reference space, as the tissue segmentation is performed before transformations, and a state-of-the-art method is used for spatial normalisation.

## Metrics

At each image processing pipeline, a set of SUVR images were generated by dividing each voxel value of the [<sup>11</sup>C]PiB image by the average SUV obtained from the whole cerebellum's VOI. Then, the voxels with  $SUVR \geq 1.5$  were defined as PiB positive (A $\beta$ +) [32–34] and corresponds to a high sensitivity for AD classification [35,36]. Finally, three sets of metrics were extracted from the GM VOI: intensity-based, volumetric, and multi-parametric.

Intensity metrics:

- Average SUVR of all voxels within the GM VOI ( $SUVR_{\text{mean}}$ )
- Average SUVR of A $\beta$ +

Volumetric measurements:

- Volume of A $\beta$ +
- Amyloid Fractional Volume (AFV): the ratio between the A $\beta$ +

Multi-parametric feature that combines intensity and volumetric metrics:

- Total Amyloid Burden (TAB) =  $SUVR_{\text{meanA}\beta+} \times \text{A}\beta+$  volume

The extraction of these metrics was performed by in-house developed MATLAB (version R2018b) scripts based on SPM functions.

## Data analysis

Image processing pipelines were explored with boxplots and descriptive statistical analysis (mean and standard deviations are shown in the [S1 File](#)). The lower and upper whiskers of the boxplots extend to the 1.5×IQR (interquartile range), data beyond that are plotted individually. Additionally, an inferential statistical analysis was performed to explore the impact of the different pipelines on the metrics using a Generalised Estimating Equations (GEE) model [37–39]. The GEE model is known to achieve higher statistical power with small sample sizes than ANOVA [39] and does not require data normality [37]. For the present study, the GEE model was built with a linear scale response and an independent working correlation matrix. Space of analysis, spatial transformation method, and grey matter definition were included in the model as independent variables, as well as their interactions. The model was then applied independently to each metric (i.e. dependent variable). The overall results of the GEE model were reported with the mean difference, the 95% confidence interval, and the p-value in a table. These results are mentioned in the text were appropriate with the mean difference and p-value. A p-value of 0.05 was used as the threshold for considering statistical significance (Wald method, without correction for multiple comparisons in the pairwise comparisons). The results from the two main pipelines are presented in the main text in depth with the mean difference, 95% confidence interval, and p-value. The boxplot and descriptive statistical analyses were carried out with R (version 3.5.1, Rstudio version 1.1.456), while the GEE analyses were performed with SPSS (version 26.0.0.1, IBM Armonk, NY, USA).

## Results

In this section we will address the impact of using different reference spaces, spatial transformation methods, and GM definition methods separately for each of the amyloid burden uptake metrics, being intensity (SUVR), A $\beta$ +

Table 2. Overall comparisons of different pipelines.

Metric	Comparison		Difference (%)	95% Confidence interval (%)	p
SUVR <sub>mean</sub>	Reference Spaces	Standard—Native	0.9	0.3–1.6	< 0.01
	Spatial Transformations	MAPc—MAPa	0.6	-0.4–1.6	N.S.
		USM—MAPa	10.5	5.5–15.5	< 0.001
		USM—MAPc	9.9	5.2–14.6	< 0.001
	Grey Matter Definitions	MRs—MRn <sup>†</sup>	-0.2	-0.3–0.0	N.S.
		TEs—MRn	5.7	-8.0–3.4	< 0.001
TEs—MRs		5.6	-8.0–3.3	< 0.001	
SUVR <sub>meanAβ+</sub>	Reference Spaces	Standard—Native	-0.6	-0.9 – -0.3	< 0.001
	Spatial Transformations	MAPc—MAPa	-0.0	-0.5–0.5	N.S.
		USM—MAPa	4.8	2.0–7.6	< 0.01
		USM—MAPc	4.8	2.2–7.4	< 0.001
	Grey Matter Definitions	MRs—MRn <sup>†</sup>	-0.1	-0.2 – -0.0	< 0.01
		TEs—MRn	0.9	0.2–1.7	< 0.05
TEs—MRs		1.3	0.4–2.1	< 0.01	
Aβ+ Volume	Reference Spaces	Standard—Native	13.8	5.9–21.7	< 0.01
	Spatial Transformations	MAPc—MAPa	3.6	-1.8–8.9	N.S.
		USM—MAPa	36.0	16.5–55.4	< 0.001
		USM—MAPc	32.5	14.9–50.1	< 0.001
	Grey Matter Definitions	MRs—MRn <sup>†</sup>	5.3	2.9–7.7	< 0.001
		TEs—MRn	30.8	20.9–40.8	< 0.001
TEs—MRs		18.0	10.6–25.3	< 0.001	
Amyloid Fractional Volume	Reference Spaces	Standard—Native	0.3	-2.6–3.1	N.S.
	Spatial Transformations	MAPc—MAPa	1.1	-3.7–5.9	N.S.
		USM—MAPa	35.6	16.6–54.6	< 0.001
		USM—MAPc	34.6	16.8–52.3	< 0.001
	Grey Matter Definitions	MRs—MRn <sup>†</sup>	-1.5	-2.0 – -1.0	< 0.001
		TEs—MRn	-9.9	-18.1 – -1.7	< 0.05
TEs—MRs		-8.3	-16.7–0.1	N.S.	
Total Amyloid Burden	Reference Spaces	Standard—Native	13.4	5.6–21.2	< 0.01
	Spatial Transformations	MAPc—MAPa	3.8	-2.3–10.0	N.S.
		USM—MAPa	45.0	21.0–68.9	< 0.001
		USM—MAPc	41.3	19.5–63.1	< 0.001
	Grey Matter Definitions	MRs—MRn <sup>†</sup>	5.2	2.7–7.7	< 0.001
		TEs—MRn	30.4	19.9–40.8	< 0.001
TEs—MRs		17.7	10.1–25.3	< 0.001	

<sup>†</sup>These comparisons were made in *Standard Space*. Statistical significance is presented in tiers: Non-significant (N.S.),  $p < 0.05$ ,  $p < 0.01$ , and  $p < 0.001$ .

<https://doi.org/10.1371/journal.pone.0248122.t002>

pipelines (*Native Space* + MRn + USM vs. *Standard Space* + MRn + USM) are compared in Table 3. At the end of this section, an overview summary of the main results is provided.

### Intensity metrics

**Comparing reference spaces.** The global effects from the GEE model (i.e., no interaction of the independent variables) found a small statistically significant overall difference of 0.92% ( $p = 0.007$ ) in the SUVR<sub>mean</sub> values obtained in the different reference spaces, with *Native Space* having lower values than those in *Standard Space* (Table 2). Fig 2 and Table B in S1 File show SUVR<sub>mean</sub> distribution for different image processing pipelines and splits the subject

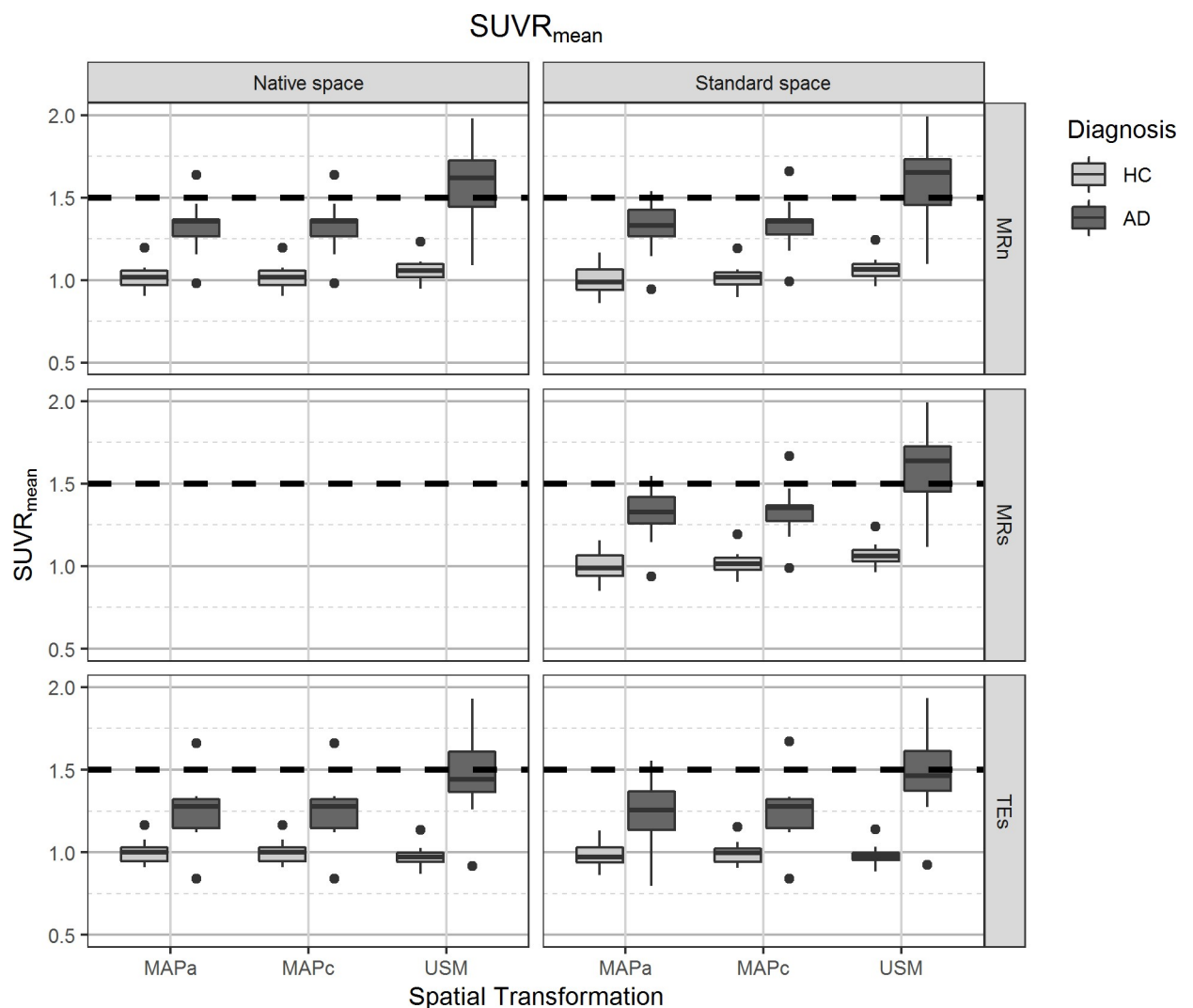
**Table 3. Comparing the main pipelines.**

	SUVR <sub>mean</sub>	SUVR <sub>meanAβ+</sub>	Volume Aβ+	AFV	TAB
Difference (%)	0.8	-0.2	19.3	1.6	19.1
95% CI (%)	0.5–1.0	-0.3 – -0.0	11.5–27.1	0.5–2.7	11.0–27.2
p	< 0.001	< 0.05	< 0.001	< 0.01	< 0.001

Comparing the two main pipelines: [Standard Space + MRn + USM]–[Native Space + MRn + USM].

<https://doi.org/10.1371/journal.pone.0248122.t003>

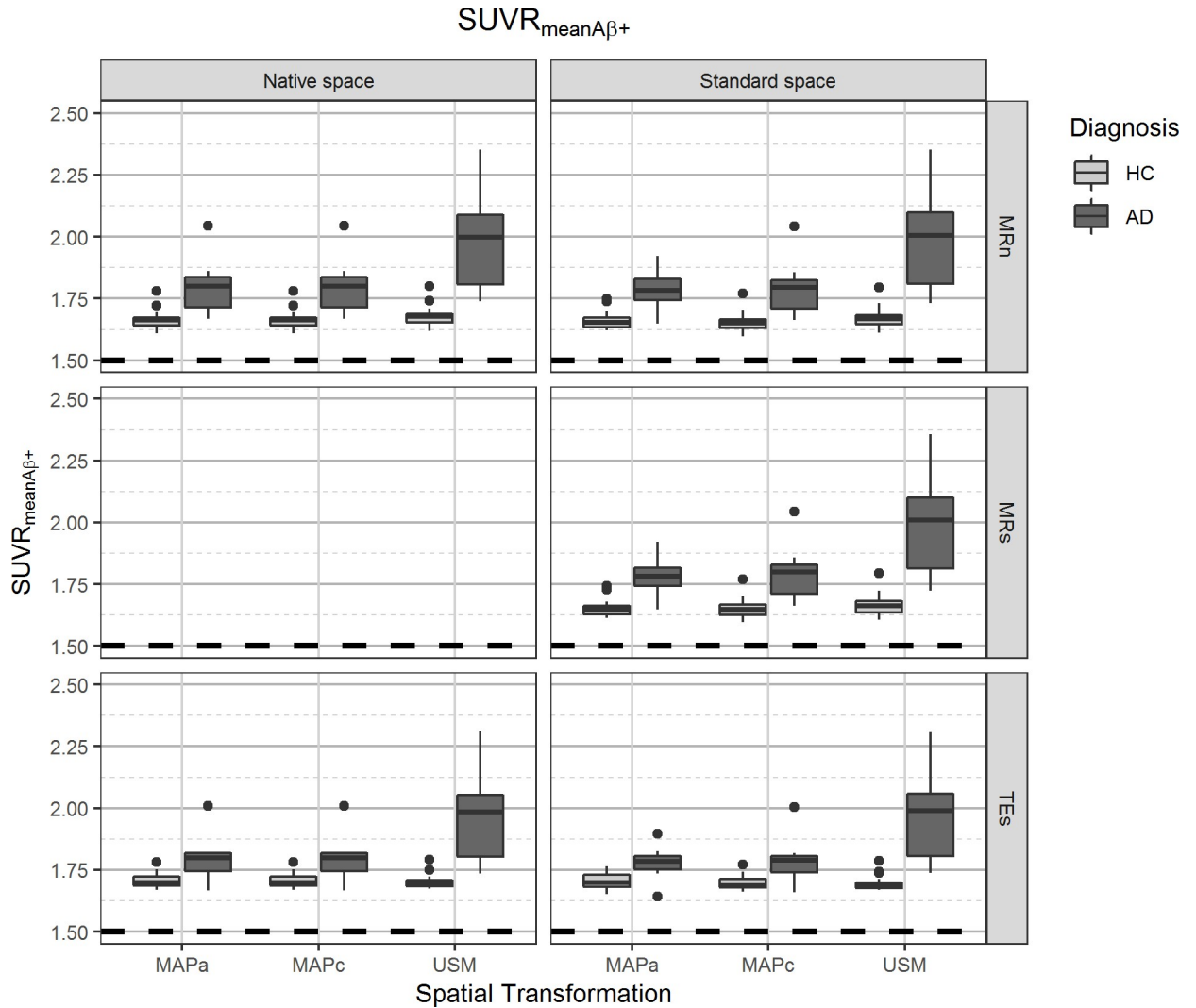
groups by diagnosis. It was also possible to see that there was not a substantial difference in SUVR<sub>mean</sub> between the two main pipelines (0.8%, p<0.001, Table 3), that is, the change was less than what is expected from a repeatability analysis, even if it was statistically significant. Similarly, SUVR<sub>meanAβ+</sub> was not substantially affected by analysis in different reference spaces



**Fig 2. Distribution of [<sup>11</sup>C]PiB mean uptake ratio in grey matter tissue (SUVR<sub>mean</sub>).** On the left panels, values from images analysed in the *Native Space*, on the right images that were spatially normalised to *Standard Space*. The different spatial normalisation methods are represented in sub-columns within each column, notice that spatial transformations for images in *Native Space* are required to transform template images from *Standard Space*. Rows show the different grey matter mask definitions. Diagnosis is indicated on the legend.

<https://doi.org/10.1371/journal.pone.0248122.g002>





**Fig 3. Distribution of Aβ+ mean uptake in all grey matter tissue (SUVR<sub>meanAβ+</sub>).** On the left panels, values from images analysed in the *Native Space*, on the right images that were spatially normalised to *Standard Space*. The different spatial normalisation methods are represented in sub-columns within each column, notice that spatial transformations for images in *Native Space* are required to transform template images from *Standard Space*. Rows show the different grey matter mask definitions. Diagnosis is indicated on the legend.

<https://doi.org/10.1371/journal.pone.0248122.g003>

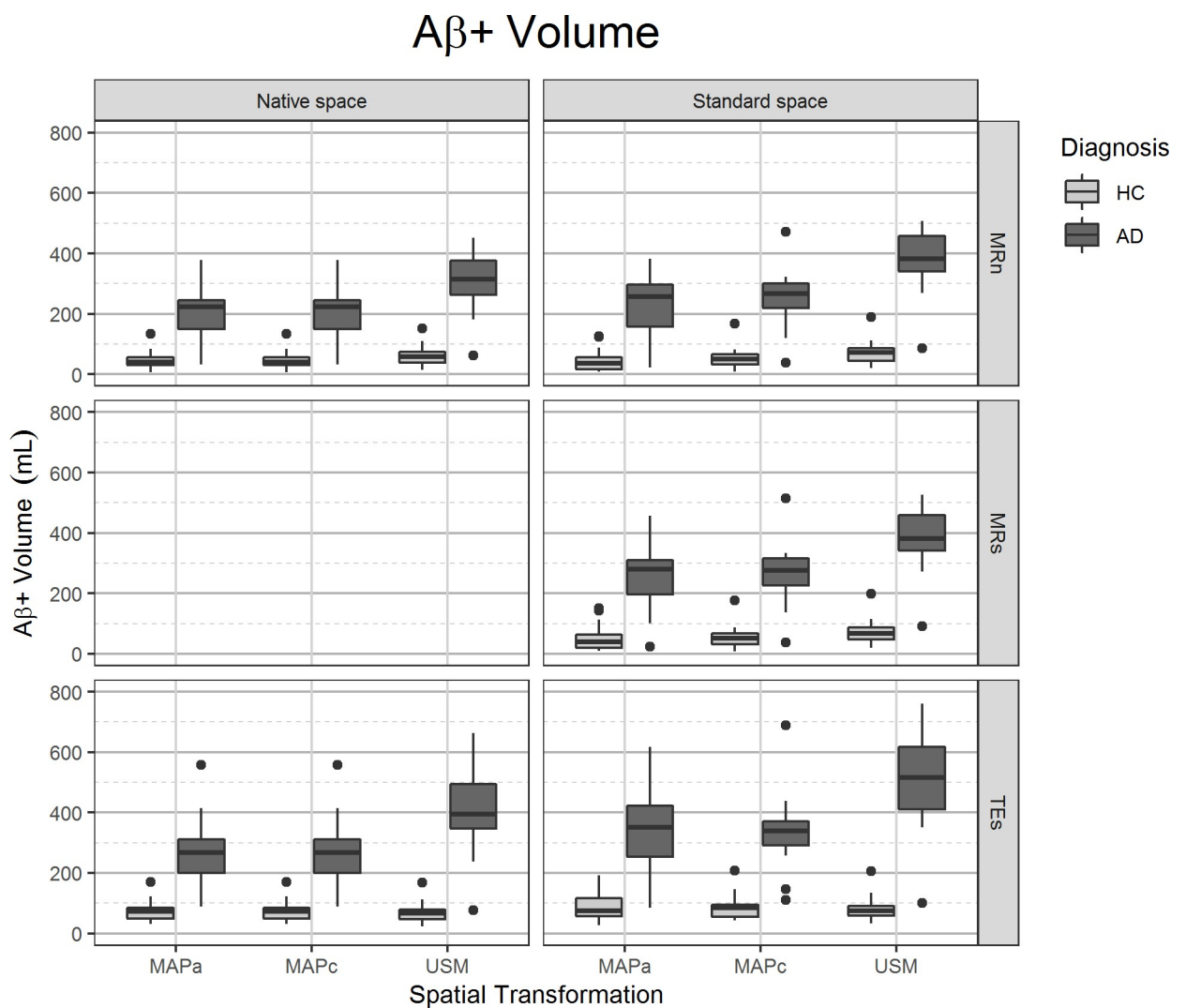
(Table 2). Fig 3 and Table B in S1 File show similar behaviour for SUVR<sub>meanAβ+</sub> as previously described for SUVR<sub>mean</sub>. Difference between the two main pipelines was statistically significant, though with a very small effect size of -0.2% (p = 0.024, Table 3).

**Comparing spatial transformation methods.** Overall, the SUVR<sub>mean</sub> values were only statistically and substantially affected by spatial transformation between the USM and the MAP methods, with USM yielding higher SUVR<sub>mean</sub> values (p<0.001, 10.5% and 9.9%, respectively for MAPa and MAPc). Fig 2 and Table B in S1 File show that there was a better separation between the subject populations when USM was used for spatial transformations, which also resulted in a larger spread of SUVR<sub>mean</sub> values. Likewise, there was not an overall statistically significant difference in using MAPa or MAPc for SUVR<sub>meanAβ+</sub> estimation. Meanwhile, USM produced slightly higher values than both MAP methods, however there was not a substantial impact in the metric (Table 2).

**Comparing GM definition methods.** Overall (global effects from the GEE model), slightly lower  $SUVR_{mean}$  values were obtained using TEs than when the subject's MRI was used to define the grey matter tissue, either from MRn (-5.7%,  $p < 0.001$ ) or MRs (-5.6%,  $p < 0.001$ ). For analysis in *Standard Space*, there was no statistical difference on  $SUVR_{mean}$  values between MRn and MRs. This can be seen in Table 2, Fig 3 and Table B in S1 File, which also show that all pipelines involving TEs make the separation between the two subject populations more difficult and set the average  $SUVR_{mean}$  value below the  $A\beta$  positivity threshold ( $SUV = 1.5$ ). Differences in overall  $SUVR_{meanA\beta+}$  values estimated from different GM tissue segmentations were not substantial, that is, smaller than that expected from a repeatability assessment (Table 2).

**A $\beta$ + volume**

**Comparing reference spaces.** The overall mean  $A\beta$ + volume obtained in *Standard Space* was 13.8% ( $p = 0.001$ ) higher than that from *Native Space*. Fig 4 and Table 3 in S1 File show



**Fig 4. Volume of high [ $^{11}C$ ]PiB deposition region in grey matter ( $A\beta$ + volume in GM).** On the left panels, values from images analysed in the *Native Space*, on the right images that were spatially normalised to *Standard Space*. The different spatial normalisation methods are represented in sub-columns within each column, notice that spatial transformations for images in *Native Space* are required to transform template images from *Standard Space*. Rows show the different grey matter mask definitions. Diagnosis is indicated on the legend.

<https://doi.org/10.1371/journal.pone.0248122.g004>

that, as expected, AD subjects presented a higher  $A\beta+$  volume than HC subjects for any image processing pipeline with analysis in either reference space, and indicates that  $A\beta+$  volume measurements on AD subjects were more susceptible to spatial normalisation than HC subjects. Additionally, there was a significant difference between the two main image processing pipelines, with 19% ( $p < 0.001$ ) larger  $A\beta+$  volumes in *Standard Space* than in *Native Space*.

**Comparing spatial transformation methods.** Overall larger  $A\beta+$  volumes were obtained using with USM than with the MAPa and MAPc (36% and 33%, respectively;  $p < 0.001$  for both), and there was no significant difference between the two MAP methods themselves (Table 2). Spatial transformation with USM noticeably increased the separation between the two subject populations when compared with the MAP methods (Fig 4 and Table C in S1 File).

**Comparing GM definition methods.** Overall,  $A\beta+$  volumes estimated with GM from TEs were 31% ( $p < 0.001$ ) larger than those that used MRn and 18% ( $p < 0.001$ ) larger than MRs. For analysis in *Standard Space*, there was no substantial difference between tissue segmentation from the T1-MRI before spatial transformation (MRn) or after (MRs).

## Amyloid fractional volume

**Comparing reference spaces.** Overall, the AFV values were robust and not statistically significantly affected by the estimation at *Native Space* or *Standard Space*. Fig 5 and Table C in S1 File show that the invariability of AFV against estimation on different reference spaces was not dependent on subject diagnosis. Comparing the two main image processing pipelines, only a 1.6% ( $p = 0.005$ ) difference was found, with higher values in *Standard Space*.

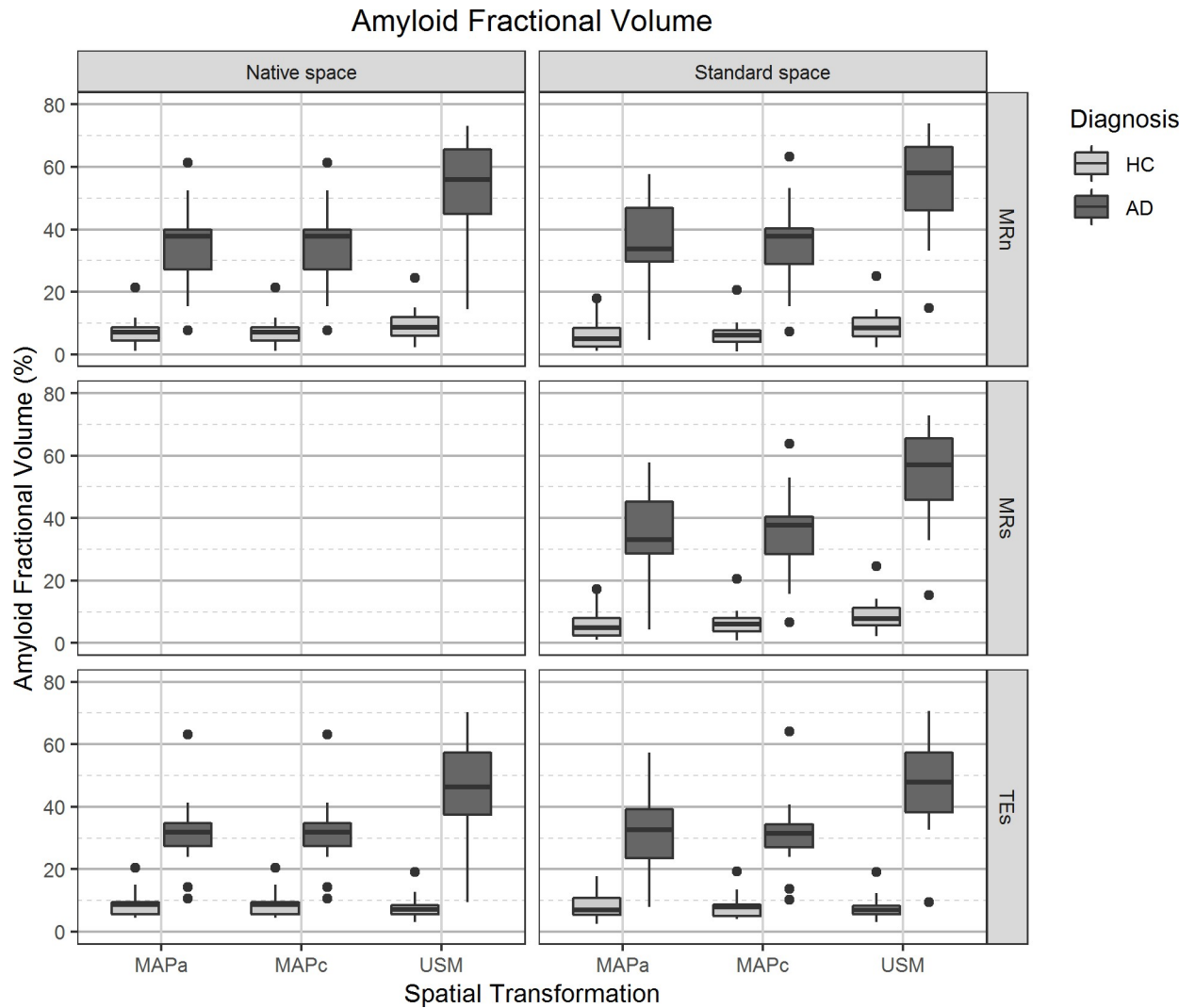
**Comparing spatial transformation methods.** AFV was strongly affected by the spatial transformation method, with USM showing higher values than MAPa and MAPc methods (36% and 35%, respectively;  $p < 0.001$  for both). However, there was no statistically significant difference between the MAP methods (Table 2). Subject group separation was larger when using USM for spatial transformations when compared against the MAP methods (Fig 5 and Table C in S1 File).

**Comparing GM definition methods.** In general, AFV was only substantially and statistically significantly affected by GM definition when comparing TEs and MRn segmentation methods (-9.9%,  $p = 0.017$ , with lower values using TEs). In *Standard Space*, there was no substantial difference between GM segmentation before or after spatial normalisation of the MRI (Table 2).

## Total amyloid burden

**Comparing reference spaces.** TAB estimation was affected by analysis in different reference spaces, with 13% ( $p = 0.001$ ) higher values in *Standard Space* than in *Native Space*. Fig 6 and Table D in S1 File show that TAB of AD subjects was slightly more impacted by the analysis in different spaces than that of HC subjects. Comparing the TAB estimated from the two main pipelines, 19% ( $p < 0.001$ ) higher values were found in *Native Space* than in *Standard Space*.

**Comparing spatial transformation methods.** There was no significant difference in TAB estimations between images processed with the MAP methods (Table 2). However, USM resulted in overall higher TAB values than MAPa (45%,  $p < 0.001$ ) and MAPc (41%,  $p < 0.001$ ). Fig 6 and Table D in S1 File indicate that there was a larger subject group separation when using USM than one of the MAP methods, regardless of reference space or GM tissue segmentation approach. However, USM did show a larger spread in TAB values for AD subjects than the MAP methods (TAB descriptive statistics for each pipeline available in the Table D in S1 File).

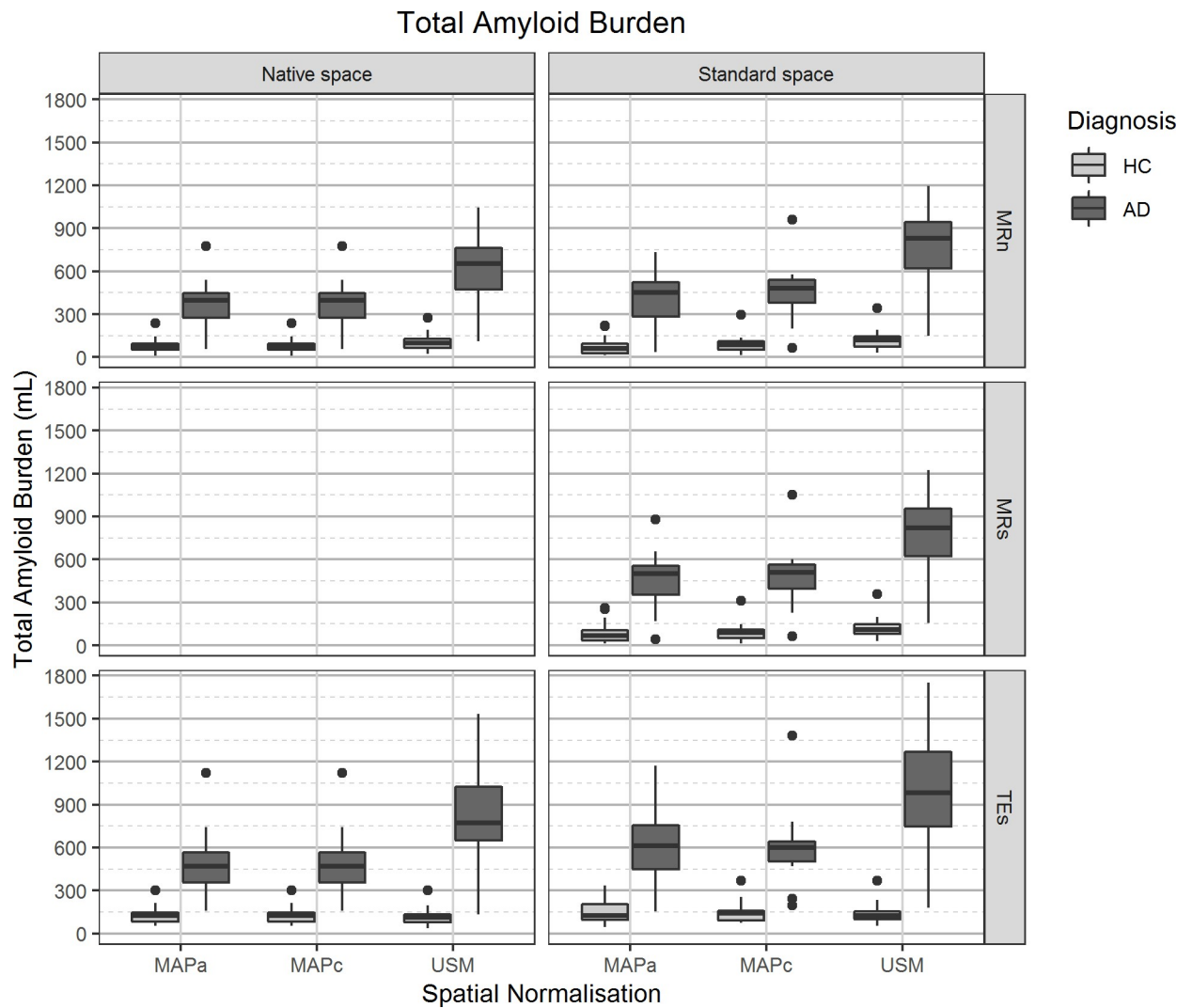


**Fig 5. Amyloid fractional volume (percentage of voxels within grey matter classified as  $A\beta+$ ) values distribution.** On the left panels, values from images analysed in the *Native Space*, on the right images that were spatially normalised to *Standard Space*. The different spatial normalisation methods are represented in sub-columns within each column, notice that spatial transformations for images in *Native Space* are required to transform template images from *Standard Space*. Rows show the different grey matter mask definitions. Diagnosis is indicated on the legend.

<https://doi.org/10.1371/journal.pone.0248122.g005>

**Comparing GM definition methods.** TAB values were highly dependent of GM tissue definition, with overall TAB values from TESs being higher than those from MRn (30%,  $p < 0.001$ ) and MRs (18%,  $p < 0.001$ ). For the analysis in *Standard Space*, there was a statistically significant difference on TAB estimated from MRn and MRs, though without a substantial effect size (Table 2). Fig 6 and Table D in S1 File show that there was a larger spread of TAB values from AD subjects when TESs was used in comparison to MRn and MRs.

In summary, it was found that the intensity metrics were not substantially affected by changes in the image processing pipeline, with a variability smaller than that expected from test-retest studies (~10%). Meanwhile, the volumetric metrics were heavily dependent on image processing pipeline (up to 36% overall differences). AFV showed more robustness to pipeline changes than  $A\beta+$  volume, and only differences in spatial transformation methods led to substantial differences. TAB was also substantially impacted by image processing pipeline



**Fig 6. Total amyloid burden value distribution.** On the left panels, values from images analysed in the *Native Space*, on the right images that were spatially normalised to *Standard Space*. The different spatial normalisation methods are represented in sub-columns within each column, notice that spatial transformations for images in *Native Space* are required to transform template images from *Standard Space*. Rows show the different grey matter mask definitions. Diagnosis is on the legend.

<https://doi.org/10.1371/journal.pone.0248122.g006>

with overall variability reaching 45%. However, there was no substantial difference from MAPa to MAPc and MRn to MRs when estimating either  $A\beta+$  volume or TAB. The two main pipelines resulted in statistically significantly different estimations for all metrics, but only  $A\beta+$  volume and TAB were substantially different, presenting 19% ( $p < 0.001$ ) higher values in *Standard Space*.

## Discussion

In this study, the impact of different PET/CT+MR imaging processing pipelines on the quantification of intensity and volumetric-derived metrics was explored. These metrics were extracted after processing the images with different pipelines, that is, choosing a reference space (*Native Space* or *Standard Space*), a spatial transformation method (MAPa, MAPc, or USM) and a grey matter delineation, i.e. based on the subject's MRI (MRn and MRs) or using

a standard template (TEs). Overall, it was found that intensity metrics were not substantially affected (i.e., changes were less than the expected test-retest (<10%) by the use of different image processing pipelines, while features using volumetric information were susceptible to the specific image processing pipeline used.

Quantification of intensity metrics, such as  $SUVR_{\text{mean}}$  and  $SUVR_{\text{meanA}\beta+}$ , provided equivalent values when extracted in *Native Space* or *Standard Space*. A higher influence on the intensity metrics was observed when different spatial transformations were used, with overall differences reaching a difference of 10% when USM was used instead of one of the MAP methods. Furthermore, the definition of the grey matter tissue did not affect the intensity metrics in a substantial way (that is, less than what is expected from test-retest studies). As expected, there was no substantial difference between the two methods that used the subject's MRI for GM definition (MRn and MRs), though they did result in slightly different intensity metrics (<6%) than a GM definition from a standard template (TEs). These (pairwise) differences were found to be statistically significant, however, the measured effect size was within the expected variability from [ $^{11}\text{C}$ ]PiB test-retest studies [40–42] and, therefore, cannot be reliably measured. Comparing the two main image processing pipelines (*Native Space* + MRn + USM vs. *Standard Space* + MRn + USM), there were no clinically relevant differences in intensity metrics between them. Based on these results, it seems reasonable to conclude that  $SUVR_{\text{mean}}$  and  $SUVR_{\text{meanA}\beta+}$  are not notably affected by the image processing pipeline used, as the effect size observed was much smaller than the test-retest variability and will, therefore, have little impact on group studies.

The two volumetric measurements studied behaved differently. On one hand,  $\text{A}\beta+$  volume was significantly affected by analysis in different reference spaces, showing that the images spatially transformed to *Standard Space* resulted in larger brain volumes than those in their *Native Space*. This can have an effect on the analysis of non-healthy brains in *Standard Space*, since the spatial transformations warp the brains into the MNI space [18], which was built with young healthy brains. Therefore, differences between healthy and non-healthy brains (that have smaller grey matter volumes due to atrophy, for example) are reduced on spatially transformed images. Hence, analysis of volumetric measurements should preferably be carried out in *Native Space* since spatial transformations may lead to bias, especially for AD subjects with strong atrophy. As such, studies reporting volumes or using the volumetric extend to determine disease stage should be interpreted with caution when comparing their results with data extracted in a different reference space. On the other hand, AFV values were not notably different from *Native Space* to *Standard Space* ( $\leq 4\%$  difference, although statistically significant). Those volume-ratio metrics were more robust than the  $\text{A}\beta+$  volume measurement when analysed in different reference spaces, and it is not surprising since the  $\text{A}\beta+$  and GM VOI volumes will be scaled and spatially transformed similarly. It is important to notice, however, that localised brain atrophy information may be lost when spatial transformations into *Standard Space* are performed. Moreover, the use of spatial transformations based on USM method resulted in roughly 33–36% higher  $\text{A}\beta+$  volumes and AFV values as compared to those obtained using the MAP methods, indicating that image transformation regularisation with TPM has a strong and significant impact on volumetric measurements, even when the ratio between volumes is considered. Note that USM is the most up-to-date spatial transformation method tested in the present study and previous literature suggest that it is an improvement over MAP methods [23]. Furthermore, the grey matter tissue definition heavily affected the estimation of the  $\text{A}\beta+$  volume, with overall differences between pipelines on the range of 18–31% if a template GM was used instead of tissue segmentation directly from the subject's T1-MRI. This was expected as these GM templates are made such to guarantee the full inclusion of the actual GM regions of the subject and thereby overestimating the actual GM. AFV

values were somewhat less influenced by the grey matter definition, with around 9–10% difference between the GM estimation from the subject's MRI against a standard template. This reflects another consequence of the GM volume overestimation by these predefined templates, i.e. including voxels outside the actual subject's GM and that had a lower tracer uptake ( $SUVR < 1.5$ ). Finally, comparing the two main image processing pipelines there was a significant and measurable difference in  $A\beta+$  volume (19%,  $p < 0.001$ ), but the volumetric normalisation aspect of AFV made it robust between these pipelines (1.6%,  $p = 0.005$ ). These results highlight the importance of harmonising the order in which the spatial transformations and the grey matter tissue segmentation are performed to obtain reproducible volumetric estimations for amyloid burden volumes.

Total amyloid burden combines  $SUVR_{\text{mean}A\beta+}$  with  $A\beta+$  volume and could potentially provide similar information to what Total Lesion Glycolysis gives in oncological [ $^{18}\text{F}$ ]FDG studies, such as being a prognostic factor for disease progression, recurrence and death [43,44]. However, the variability due to changes in image processing pipeline observed for  $A\beta+$  volume was reflected in TAB, which was significantly different between pipelines: overall significant changes from the analysis in different reference spaces was at 13%, due to different spatial transformations at 4–45%, and due to GM definitions on the range of 5–30%. By processing the images with the two main pipelines, a strong effect on TAB estimation between the two reference spaces (19%,  $p < 0.001$ ) was found. As such, the multi-parametric aspect of TAB did not compensate the variability from its volumetric component, but the other way around, it was amplified by its intensity component. Thus, it is uncertain to confirm the usefulness for an amyloid burden assessment with [ $^{11}\text{C}$ ]PiB using a metric that is so susceptible to image processing methodology.

The effects of the image processing pipeline on the diagnostic discrimination power of each feature were not statistically evaluated since the HC and AD populations in the present study were chosen to have maximum separation and represent rather extreme cases, that is, no subjects with a doubtful classification or mild-cognitive impairment were included. As such, the statistical analysis was carried out pooling together both diagnoses to investigate the impact of different image processing pipelines on metrics regardless of clinical diagnosis. In any case, it was observed in the boxplots and descriptive statistics of each population that, in general, metrics from HC subjects were less affected by changes in image processing pipeline than measurements from AD subjects (Figs 2–6 and Supplementary Tables B–D in S1 File). This was expected as most features only included voxels with  $SUVR$  above 1.5 and AD subjects have more voxels that fit this criterion and a larger range of values above this threshold. A previous study [45] suggested that a lower  $SUVR$  threshold ( $SUVR \geq 1.2$ ) for amyloid positivity and the use of the grey matter cerebellum as the reference tissue could improve sensitivity for diagnosis without impairing specificity. In the current study, the metrics were also assessed with a lower positivity threshold ( $SUVR \geq 1.2$ ; see available data) and the general behaviour of the metrics was similar as the ones already presented (i.e.  $SUVR \geq 1.5$ ). Therefore, changes in the  $SUVR$  threshold for amyloid burden positivity do not seem to overcome the variability observed in quantification from different image processing pipelines. That same study provides an answer about the use of different reference regions by stating that the use of the grey matter cerebellum as the reference region (instead of the whole cerebellum) increases the [ $^{11}\text{C}$ ]PiB  $SUVR$  values by approximately 6%, thus simply rescaling the positivity threshold. Consequently, the impact of different image processing pipelines on metrics should remain similar to the one presented here if the GM cerebellum was considered as reference region.

The results obtained in the present study are in line with those published by Nørgaard et al. [19,20], where the influence of acquisition (PET scanner, acquisition protocol, reconstruction, and MR field strength), pre-processing steps (motion correction, co-registration, volume of

interest delineation strategy, and partial volume correction), and analysis (kinetic model and statistical analysis) on [ $^{11}\text{C}$ ]DASB studies was explored. However, that study used a different tracer and metrics. The results of Nørgaard et al. and those observed in the present study demonstrate that different image processing methodologies can have a significant effect on PET quantification and support the idea that these findings are not restricted to a specific radio-tracer or disease. Other studies [22,23] explored different methods for the spatial transformation of brain images and observed that methods which do not use TPM for regularisation of the transformations had worse performance than the regularised transformations, supporting the observed difference between the MAP and the USM methods in the current study. Furthermore, rigid transformation alignment errors can have an effect on SUVR quantification and lead to a bias related to clinical severity [46]. This may also be the case for non-linear transformations, possibly explaining why data from AD subjects were more affected by differences in image processing pipeline than those of HC subjects in the present study.

A possible limitation of this work is that the reference region and grey matter template masks were spatially transformed from *Standard Space* to *Native Space* using the inverse transformation calculated from the subjects' MRI. Therefore, MR-less pipelines which would benefit from template masks were not explored. However, it is uncommon to have MR-less dementia PET studies. Regional metrics were not evaluated, and target regions could be more sensitive to image processing than reference regions or global regions (as explored here).

In summary, the main finding of the present study is that intensity metrics (SUVRs) are rather robust against variations in image processing pipelines, and regional SUVR variability was within that expected from test-retest studies. In contrast, volumetric measurements are more sensitive to the image processing pipeline used. Based on the current results, regional-based studies should preferably be conducted in *Native Space* and tissue segmentation should be estimated from each subject's MRI when available, while standard VOIs (available in *Standard Space*) should be transformed into *Native Space*. Nevertheless, an inter-subject voxel-based analysis must be conducted in *Standard Space* and, for that matter, a standardised approach for spatial transformations must be employed, preferably using the Unified Segmentation Method (USM) since it uses tissue probability maps for transformations' regularisation. In either case, volumetric-based metrics should be avoided for studies in *Standard Space*. When comparing data and results with previous reports in the literature, interpretations should be made carefully since different conclusions could arise from methodological differences rather than from the biological process under investigation.

## Supporting information

**S1 File. Supplementary tables.** Include tables with additional information about parameters for spatial transformations, descriptive statistics of the metrics presented in the manuscript, and a table of abbreviations.  
(DOCX)

## Author Contributions

**Conceptualization:** Guilherme D. Kolinger, David Vallez Garca, Ronald Boellaard.

**Data curation:** Guilherme D. Kolinger, David Vallez Garca.

**Formal analysis:** Guilherme D. Kolinger.

**Funding acquisition:** Peter P. De Deyn, Ronald Boellaard.

**Investigation:** Guilherme D. Kolinger, David Vallez Garca, Ronald Boellaard.



**Methodology:** Guilherme D. Kolinger, David Vallez Garca, Ronald Boellaard.

**Project administration:** Fransje E. Reesink, Bauke M. de Jong, Rudi A. J. O. Dierckx, Peter P. De Deyn, Ronald Boellaard.

**Resources:** Bauke M. de Jong, Rudi A. J. O. Dierckx.

**Software:** Guilherme D. Kolinger.

**Supervision:** David Vallez Garca, Ronald Boellaard.

**Validation:** Guilherme D. Kolinger.

**Writing – original draft:** Guilherme D. Kolinger, David Vallez Garca, Antoon T. M. Willemssen, Fransje E. Reesink, Bauke M. de Jong, Rudi A. J. O. Dierckx, Peter P. De Deyn, Ronald Boellaard.

**Writing – review & editing:** Guilherme D. Kolinger, David Vallez Garca, Antoon T. M. Willemssen, Fransje E. Reesink, Bauke M. de Jong, Rudi A. J. O. Dierckx, Peter P. De Deyn, Ronald Boellaard.

## References

1. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group\* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984; 34: 939–939. <https://doi.org/10.1212/wnl.34.7.939> PMID: 6610841
2. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011; 7: 263–269. <https://doi.org/10.1016/j.jalz.2011.03.005> PMID: 21514250
3. Baal M, Braak H, Coleman P, Dickson D, Duyckaerts C, Gambetti P, et al. Consensus Recommendations for the Postmortem Diagnosis of Alzheimer's Disease. *Neurobiol Aging*. 1997; 18: S1–S2. [https://doi.org/10.1016/S0197-4580\(97\)00057-2](https://doi.org/10.1016/S0197-4580(97)00057-2) PMID: 9330978
4. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018; 14: 535–562. <https://doi.org/10.1016/j.jalz.2018.02.018> PMID: 29653606
5. Rowe CC, Ng S, Ackermann U, Gong SJ, Pike K, Savage G, et al. Imaging beta-amyloid burden in aging and dementia. *Neurology*. 2007; 68: 1718–25. <https://doi.org/10.1212/01.wnl.0000261919.22630.ea> PMID: 17502554
6. Tolboom N, Yaqub M, van der Flier WM, Boellaard R, Luurtsema G, Windhorst AD, et al. Detection of Alzheimer pathology in vivo using both 11C-PIB and 18F-FDDNP PET. *J Nucl Med Off Publ Soc Nucl Med*. 2009; 50: 191–7. <https://doi.org/10.2967/jnumed.108.056499> PMID: 19164243
7. Firouzian A, Whittington A, Searle GE, Koychev I, Zamboni G, Lovestone S, et al. Imaging A $\beta$  and tau in early stage Alzheimer's disease with [18F]AV45 and [18F]AV1451. *EJNMMI Res*. 2018; 8. <https://doi.org/10.1186/s13550-018-0371-y> PMID: 29500717
8. Jack CR, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol*. 2010; 9: 119–128. [https://doi.org/10.1016/S1474-4422\(09\)70299-6](https://doi.org/10.1016/S1474-4422(09)70299-6) PMID: 20083042
9. Nordberg A, Rinne JO, Kadir A, Lngstrom B. The use of PET in Alzheimer disease. *Nat Rev Neurol*. 2010; 6: 78–87. <https://doi.org/10.1038/nrneurol.2009.217> PMID: 20139997
10. Cohen AD, Klunk WE. Early detection of Alzheimer's disease using PiB and FDG PET. *Neurobiol Dis*. 2014; 72: 117–122. <https://doi.org/10.1016/j.nbd.2014.05.001> PMID: 24825318
11. Zwan MD, Bouwman FH, Konijnenberg E, Van Der Flier WM, Lammertsma AA, Verhey FRJ, et al. Diagnostic impact of [18F]flutemetamol PET in early-onset dementia. *Alzheimers Res Ther*. 2017; 9: 1–8. <https://doi.org/10.1186/s13195-016-0227-5> PMID: 28073379
12. Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, et al. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Ann Neurol*. 2004; 55: 306–319. <https://doi.org/10.1002/ana.20009> PMID: 14991808

13. Ng SY, Villemagne VL, Masters CL, Rowe CC. Evaluating Atypical Dementia Syndromes Using Positron Emission Tomography With Carbon 11–Labeled Pittsburgh Compound B. *Arch Neurol*. 2007; 64: 1140. <https://doi.org/10.1001/archneur.64.8.1140> PMID: 17698704
14. Pike KE, Savage G, Villemagne VL, Ng S, Moss SA, Maruff P, et al.  $\beta$ -amyloid imaging and memory in non-demented individuals: evidence for preclinical Alzheimer's disease. *Brain*. 2007; 130: 2837–2844. <https://doi.org/10.1093/brain/awm238> PMID: 17928318
15. van Berckel BNM, Ossenkuppele R, Tolboom N, Yaqub M, Foster-Dingley JC, Windhorst AD, et al. Longitudinal Amyloid Imaging Using 11C-PiB: Methodologic Considerations. *J Nucl Med*. 2013; 54: 1570–1576. <https://doi.org/10.2967/jnumed.112.113654> PMID: 23940304
16. Mintun MA, LaRossa GN, Sheline YI, Dence CS, Lee SY, Mach RH, et al. [11C]PiB in a nondemented population: Potential antecedent marker of Alzheimer disease. *Neurology*. 2006; 67: 446–452. <https://doi.org/10.1212/01.wnl.0000228230.26044.a4> PMID: 16894106
17. Rabinovici GD, Furst AJ, O'Neil JP, Racine CA, Mormino EC, Baker SL, et al. 11C-PIB PET imaging in Alzheimer disease and frontotemporal lobar degeneration. *Neurology*. 2007; 68: 1205–1212. <https://doi.org/10.1212/01.wnl.0000259035.98480.ed> PMID: 17420404
18. Collins DL, Zijdenbos AP, Kollokian V, Sled JG, Kabani NJ, Holmes CJ, et al. Digital Brain Phantom. *IEEE Trans Med Imaging*. 1998; 17: 463–468. <https://doi.org/10.1109/42.712135> PMID: 9735909
19. Nørgaard M, Ganz M, Svarer C, Feng L, Ichise M, Lanzenberger R, et al. Cerebral serotonin transporter measurements with [11C]DASB: A review on acquisition and preprocessing across 21 PET centres. *J Cereb Blood Flow Metab*. 2018. <https://doi.org/10.1177/0271678X18770107> PMID: 29651896
20. Nørgaard M, Greve DN, Svarer C, Strother SC, Knudsen GM, Ganz M, et al. The Impact of Preprocessing Pipeline Choice in Univariate and Multivariate Analyses of PET Data. 2018 Int Workshop Pattern Recognit Neuroimaging PRNI 2018. 2018; 2–5. <https://doi.org/10.1109/PRNI.2018.8423962>
21. Schwarz CG, Jones DT, Gunter JL, Lowe VJ, Vemuri P, Senjem ML, et al. Contributions of imprecision in PET-MRI rigid registration to imprecision in amyloid PET SUVR measurements. *Hum Brain Mapp*. 2017; 38: 3323–3336. <https://doi.org/10.1002/hbm.23622> PMID: 28432784
22. Martino ME, De Villoria JG, Lacalle-Aurioles M, Olazarán J, Cruz I, Navarro E, et al. Comparison of different methods of spatial normalization of FDG-PET brain images in the voxel-wise analysis of MCI patients and controls. *Ann Nucl Med*. 2013; 27: 600–609. <https://doi.org/10.1007/s12149-013-0723-7> PMID: 23585159
23. Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang M-CC, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*. 2009; 46: 786–802. <https://doi.org/10.1016/j.neuroimage.2008.12.037> PMID: 19195496
24. Presotto L, Iaccarino L, Sala A, Vanoli EG, Muscio C, Nigri A, et al. Low-dose CT for the spatial normalization of PET images: A validation procedure for amyloid-PET semi-quantification. *NeuroImage Clin*. 2018; 20: 153–160. <https://doi.org/10.1016/j.nicl.2018.07.013> PMID: 30094164
25. Ashburner J, Neelin P, Collins DL, Evans A, Friston K. Incorporating Prior Knowledge into Image Registration. *NeuroImage*. 1997; 6: 344–352. <https://doi.org/10.1006/nimg.1997.0299> PMID: 9417976
26. Ashburner J, Friston KJ. Nonlinear spatial normalization using basis functions. 13.
27. Ashburner J, Friston KJ. Unified segmentation. *NeuroImage*. 2005; 26: 839–851. <https://doi.org/10.1016/j.neuroimage.2005.02.018> PMID: 15955494
28. D'Agostino E, Maes F, Vandermeulen D, Suetens P. Non-rigid Atlas-to-Image Registration by Minimization of Class-Conditional Image Entropy. In: Barillot C, Haynor DR, Hellier P, editors. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2004*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2004. pp. 745–753. [https://doi.org/10.1007/978-3-540-30135-6\\_91](https://doi.org/10.1007/978-3-540-30135-6_91)
29. Peretti DE, Reesink FE, Doorduyn J, de Jong BM, De Deyn PP, Dierckx RAJO, et al. Optimization of the  $k_2'$  Parameter Estimation for the Pharmacokinetic Modeling of Dynamic PiB PET Scans Using SRTM2. *Front Phys*. 2019; 7. <https://doi.org/10.3389/fphy.2019.00212>
30. Peretti DE, Vázquez García D, Reesink FE, van der Goot T, De Deyn PP, de Jong BM, et al. Relative cerebral flow from dynamic PiB scans as an alternative for FDG scans in Alzheimer's disease PET studies. *Plos One*. 2019; 14: e0211000. <https://doi.org/10.1371/journal.pone.0211000> PMID: 30653612
31. Hammers A, Allom R, Koeppe MJ, Free SL, Myers R, Lemieux L, et al. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp*. 2003; 19: 224–47. <https://doi.org/10.1002/hbm.10123> PMID: 12874777
32. Jack CR, Lowe VJ, Weigand SD, Wiste HJ, Senjem ML, Knopman DS, et al. Serial PiB and MRI in normal, mild cognitive impairment and Alzheimer's disease: implications for sequence of pathological events in Alzheimer's disease. *Brain*. 2009; 132: 1355–1365. <https://doi.org/10.1093/brain/awp062> PMID: 19339253

33. Villemagne VL, Pike KE, Chételat G, Ellis KA, Mulligan RS, Bourgeat P, et al. Longitudinal assessment of A $\beta$  and cognition in aging and Alzheimer disease. *Ann Neurol*. 2011; 69: 181–192. <https://doi.org/10.1002/ana.22248> PMID: 21280088
34. Villemagne VL, Burnham S, Bourgeat P, Brown B, Ellis KA, Salvado O, et al. Amyloid  $\beta$  deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol*. 2013; 12: 357–367. [https://doi.org/10.1016/S1474-4422\(13\)70044-9](https://doi.org/10.1016/S1474-4422(13)70044-9) PMID: 23477989
35. Jack CR, Knopman DS, Weigand SD, Wiste HJ, Vemuri P, Lowe V, et al. An operational approach to National Institute on Aging-Alzheimer's Association criteria for preclinical Alzheimer disease. *Ann Neurol*. 2012; 71: 765–75. <https://doi.org/10.1002/ana.22628> PMID: 22488240
36. Rowe CC, Ellis KA, Rimajova M, Bourgeat P, Pike KE, Jones G, et al. Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiol Aging*. 2010; 31: 1275–1283. <https://doi.org/10.1016/j.neurobiolaging.2010.04.007> PMID: 20472326
37. Wang M. Generalized Estimating Equations in Longitudinal Data Analysis: A Review and Recent Developments. *Adv Stat*. 2014; 2014: 1–11. <https://doi.org/10.1155/2014/303728>
38. Zeger SL, Liang K-Y. Longitudinal Data Analysis for Discrete and Continuous Outcomes. *Biometrics*. 1986; 42: 121–130. PMID: 3719049
39. Ma Y, Mazumdar M, Memtsoudis SG. Beyond repeated-measures analysis of variance: Advanced statistical methods for the analysis of longitudinal data in anesthesia research. *Reg Anesth Pain Med*. 2012; 37: 99–105. <https://doi.org/10.1097/AAP.0b013e31823ebc74> PMID: 22189576
40. Tolboom N, Yaqub M, Boellaard R, Luurtsema G, Windhorst AD, Scheltens P, et al. Test-retest variability of quantitative [ $^{11}\text{C}$ ]PIB studies in Alzheimer's disease. *Eur J Nucl Med Mol Imaging*. 2009; 36: 1629–1638. <https://doi.org/10.1007/s00259-009-1129-6> PMID: 19384547
41. Lopresti BJ, Klunk WE, Mathis CA, Hoge JA, Ziolkowski SK, Lu X, et al. Simplified quantification of Pittsburgh Compound B amyloid imaging PET studies: a comparative analysis. *J Nucl Med Off Publ Soc Nucl Med*. 2005; 46: 1959–1972. PMID: 16330558
42. Engler H, Forsberg A, Almkvist O, Blomquist G, Larsson E, Savitcheva I, et al. Two-year follow-up of amyloid deposition in patients with Alzheimer's disease. *Brain*. 2006; 129: 2856–2866. <https://doi.org/10.1093/brain/awl178> PMID: 16854944
43. Lee P, Weerasuriya DK, Lavori PW, Quon A, Hara W, Maxim PG, et al. Metabolic Tumor Burden Predicts for Disease Progression and Death in Lung Cancer. *Int J Radiat Oncol*. 2007; 69: 328–333. <https://doi.org/10.1016/j.ijrobp.2007.04.036> PMID: 17869659
44. La TH, Filion EJ, Turnbull BB, Chu JN, Lee P, Nguyen K, et al. Metabolic Tumor Volume Predicts for Recurrence and Death in Head-and-Neck Cancer. *Int J Radiat Oncol*. 2009; 74: 1335–1341. <https://doi.org/10.1016/j.ijrobp.2008.10.060> PMID: 19289263
45. Villeneuve S, Rabinovici GD, Cohn-Sheehy BI, Madison C, Ayakta N, Ghosh PM, et al. Existing Pittsburgh Compound-B positron emission tomography thresholds are too high: statistical and pathological evaluation. *Brain*. 2015; 138: 2020–2033. <https://doi.org/10.1093/brain/awv112> PMID: 25953778
46. Schwarz CG, Jones DT, Gunter JL, Lowe VJ, Vemuri P, Senjem ML, et al. Contributions of imprecision in PET-MRI rigid registration to imprecision in amyloid PETSUVR measurements. *Hum Brain Mapp*. 2017; 38: 3323–3336. <https://doi.org/10.1002/hbm.23622> PMID: 28432784