Send Orders for Reprints to reprints@benthamscience.ae

Current Alzheimer Research, **2016,** *13,* **597-607**

Impact of MRI-based Segmentation Artifacts on Amyloid- and FDG-PET Quantitation

Marcus Högenauer[#], Matthias Brendel[#], Andreas Delker, Sonja Därr, Mayo Weiss, Peter Bartenstein, Axel Rominger[*], for the Alzheimer's Disease Neuroimaging Initiative[**]

Department of Nuclear Medicine, University of Munich, Germany

Axel Rominger

Abstract: ***Introduction:*** Magnet resonance image (MRI)-based segmentations are widely used for clinical brain research, especially in conjunction with positron-emission-tomography (PET). Although artifacts due to segmentation errors arise commonly, the impact of these artifacts on PET quantitation has not yet been investigated systematically. Therefore, the aim of this study was to assess the effect of segmentation errors on [^{18}F]-AV45 and [^{18}F]-FDG PET quantitation, with and without correction for partial volume effects (PVE). ***Material and Methods:*** 119 subjects with both [^{18}F]-AV45, and [^{18}F]-FDG PET as well as T1-weighted MRI at baseline and at two-year follow-up were selected from the ADNI cohort, and their MRI brain images were segmented using PMOD 3.5. MRIs with segmentation artifacts were masked with the corresponding [^{18}F]-FDG PET standard-uptake-value (SUV) images to elucidate and quantify the impact of artifacts on PET analyses for six defined volumes-of-interest (VOI). Artifact volumes were calculated for each VOI, together with error-[%] and root-mean-square-errors (RMSE) in uncorrected and PVE corrected SUV results for the two PET tracers. We also assessed the bias in longitudinal PET data. ***Results:*** Artifacts occurred most frequently in the parietal cortex VOI. For [^{18}F]-AV45 and [^{18}F]-FDG PET, the percentage-errors were dependent on artifact volumes. PVEC SUVs were consequently more distorted than were their uncorrected counterparts. In static and longitudinal assessment, a small subgroup of subjects with large artifacts (≥1500 voxels; ≙5.06 cm³) accounted for much of the PET quantitation bias. ***Conclusion:*** Large segmentation artifacts need to be detected and resolved as they considerably bias PET quantitation, especially when PVEC is applied to PET data.

Keywords: Amyloid-PET, artifacts, FDG-PET, masking, MRI, PVEC, segmentation.

1. INTRODUCTION

In recent years, magnetic resonance imaging (MRI)-based segmentations of brain into grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) are increasingly important for imaging studies in diverse disorders, including Alzheimer's disease (AD) [1, 2], multiple sclerosis [3], epilepsy [4] or schizophrenia [5, 6]. Moreover, molecular brain imaging techniques such as positron-emission-tomography (PET) [7] and single-photon-emission-computed-tomography (SPECT) [8] benefit from MRI-based segmentations, such that molecular changes can be investigated in specific target regions.

Indeed, MRI-based segmentations are essential for PET quantitation, especially in dementia studies entailing partial volume effect (PVE) corrected quantitation of amyloid [9] and/or [^{18}F]-fluorodesoxyglucose ([^{18}F]-FDG) [10, 11]. Although partial volume effect correction (PVEC) is also feasible, independent of MRI (*e.g.* by deconvolution), optimal results are obtained when using anatomically-oriented segmentation based on individual structural MRI [12]. As such, widely used PVEC methods such as the geometric transfer matrix (GTM) [13] or Müller-Gärtner methods [14] depend critically on correct segmentation of individual MR images. Volume artifacts due to segmentation errors arise commonly, with consequences for PET quantitation. This problem has led to intensive efforts aiming to enhance the existing automated segmentation methods such as Multi-Atlas Propagation and Segmentation [15], Bridge Burner [16], Brain Extraction Tool [17] and Hybrid Watershed Algorithm [18]. Despite considerable progress, entirely faultless segmentation procedures are not available; the impact of such segmentation artifacts on PET quantitation, and on error propagation with PVEC, has not yet been investigated systematically. Therefore, the aim of this study was to assess the effect of MRI segmentation errors on uncorrected and PVE-corrected [^{18}F]-AV45 and [^{18}F]-FDG PET quantitation in a large dataset of scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database.

*Address correspondence to this author at the Department of Nuclear Medicine, University of Munich, Germany; Tel: +49 (0)89 4400 74650; Fax: +49 (0)89 4400 77646; E-mail: axel.rominger@med.uni-muenchen.de

***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*

[#]both authors contributed equally

1875-5828/16 $58.00+.00

2. MATERIAL AND METHODS

2.1. Alzheimer's Disease Neuroimaging Initiative

Data used in the preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California – San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-2. To date these three protocols have recruited over 1500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

2.2. Image Data Acquisition

2.2.1. ADNI [^{18}F]-AV45 PET and [^{18}F]-FDG PET Acquisition and Pre-Processing

The [^{18}F]-AV45 PET images had been acquired using Siemens, GE and Philips PET scanners according to a standard dynamic protocol with emission recording at 50-70 min following the intravenous injection of 370 ± 37 MBq [^{18}F]-AV45. [^{18}F]-FDG PET images had been acquired using Siemens, GE and Philips PET scanners according to one of three standard emission recording protocols (30–60 minute dynamic, 30–60 minute static, 0–60 minute dynamic) following the intravenous injection of 185 ± 19 MBq [^{18}F]-FDG. Data were corrected for both scatter and measured attenuation, which was determined using the CT scan for PET/CT scanners, or a transmission scan with [^{68}Ge] or [^{137}Cs] rotating rod sources for PET-only scanners. Images were reconstructed using scanner-specific algorithms, and sent to the University of Michigan, where they were reviewed for artifacts, anonymized, and transmitted to the Laboratory of Neuro Imaging (LONI) for storage. Further details are available in the ADNI PET technical procedures manual (http://www.adni-info.org/Scientists/Pdfs/ ADNI2_ PET_Tech_Manual_0142011.pdf).

Downloaded [^{18}F]-FDG PET images in DICOM format had been pre-processed in four steps: 1) motion correction by coregistration of single five minute frames; 2) time frame averaging (30-60 min p.i.); 3) coregistration of longitudinal data to the baseline scan and reorientation in a standardized 160x160x96 matrix consisting of 1.5 mm cubic voxels; 4) smoothing with a scanner-specific filter function to an isotropic resolution of 8 mm. Details can be found at (http://adni.loni.usc.edu/methods/pet-analysis/pre-processing/) [19, 20].

2.2.2. ADNI MRI Acquisition and Pre-Processing

T1-weighted (T1w) MRI scans had been acquired using Siemens, GE or Philips MRI scanners according to a standard protocol [21] involving acquisitions of two 3-D MPRAGE imaging sequences per subject. Of the two images acquired per subject and time-point, the ADNI quality assurance team selected the qualitatively better image for preprocessing.

MRI preprocessing involved: 1) application of a scanner-specific correction for gradient nonlinearity distortion (Gradwarp) [22]; 2) correction for image intensity non-uniformity (B1) [21]; 3) application of a histogram peak sharpening algorithm for bias field correction (N3) [23]; 4) application of spatial scaling factors obtained by phantom measurements. For images acquired on Philips scanners, B1 correction was already implemented, and the gradient systems with this instrument tended to be inherently linear [21].

2.3. Patient Selection and Study Design

138 subjects with both [^{18}F]-AV45 and [^{18}F]-FDG PET as well as T1w MPRAGE MRI at baseline (BL) and two-year follow-up (FU) were selected from ADNI-GO/2. Their clinical diagnoses at BL were healthy control (HC) (N=42), MCI (N=86) and AD (N=10). Pre-processed [^{18}F]-AV45 PET, [^{18}F]-FDG PET brain images and temporally corresponding T1w MPRAGE images were downloaded from the database as available on September 2013. After processing, all images were visually checked and MRI segmentation fails were identified as cases of complete mislabelling of the compartments (GM/WM/CSF) or misregistration resistant to manual adjustment. By these criteria, images of 13 subjects failed either at BL or FU, while images of six subjects failed at both BL and FU. Thus, a total of 25/276 (9.1%) failed images were excluded from further analysis, corresponding to a total of 19 subjects (HC: N=6, MCI: N=12, AD: N=1) (Fig. **1**).

2.4. Image Analyses

2.4.1. Coregistration and Segmentation

[^{18}F]-AV45 and [^{18}F]-FDG PET images were rigidly coregistered to their individual MRI to obtain a linear PET-to-MRI transformation in PMOD. Non-rigid coregistration of individual MRIs to the standard Montreal Neurological Institute (MNI)-space was used to obtain a nonlinear MRI-to-MNI transformation. Next, the concatenation of PET-to-MRI and MRI-to-MNI transformations was used to transform the PET images in MNI-space. Meanwhile, the T1w MRIs were segmented to GM, WM and CSF compartments [24]. To this end we implemented the segmentation algorithm with the following parameters: Sampling distance,

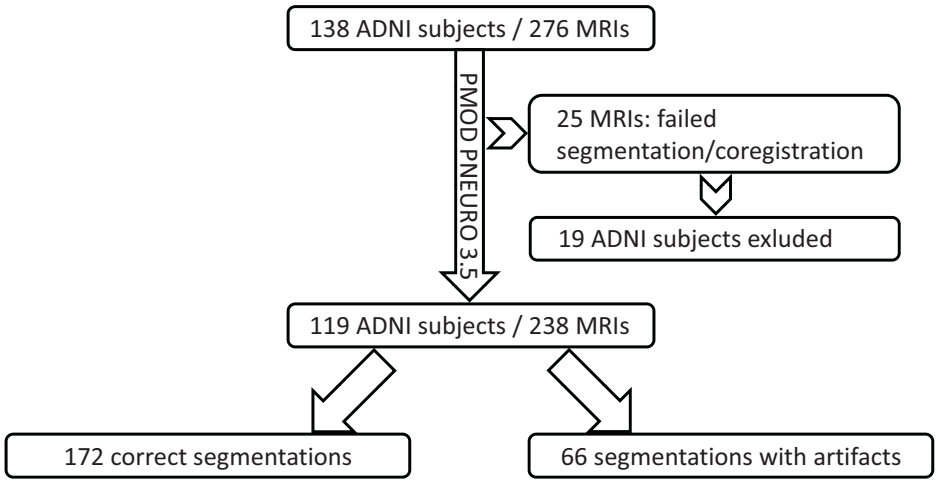

Fig. (1). Schematic overview of study design and subject selection. Image data from 138 ADNI subjects, undergoing baseline and two-year follow-up MRIs, [^{18}F]-AV45 PET and [^{18}F]-FDG PET, were processed by PMOD PNEURO 3.5. Subsequent visual identification of completely failed segmentation or coregistration resulted in exclusion of 19 subjects, and a final cohort of 238 MRIs totalling 119 subjects (HC: N=36, MCI: N=74, AD: N=9) for further analysis. 172 artifact-free segmentations and 66 segmentations with at least one artifact were obtained.

which determined the density of voxels considered in the calculation, was used at a standard 6 mm setting. Moderate bias regularisation served to compensate for variations of the image intensity across the field-of-view, and a moderate clean-up procedure rectified the incorrect segmentation along the compartment boundaries. To guarantee the best possible uniformity of segmentations, all parameters were used unadjusted in standardized manner, even when some mislabelling consequently occurred. In such cases, the segmentations were excluded, as described above. Volumes-of-interest (VOIs) of 83 cortex areas from the Hammers atlas were defined in individual GM templates for each subject in the MNI-space [25]. All image MR processings were performed with the PMOD PNEURO 3.5 tool.

2.4.2 Artifact Assessment and Masking

BL and FU segmentations of 119 subjects (HC: N=36; MCI: N=74; AD: N=9) were visually checked slice-by-slice for artifacts. In cases when an artifact was visually identified, as occurred for about 20 voxels per plane, the segmentation was masked by the individual [^{18}F]-FDG PET in order to exclude the artifact voxels. The mask binarization of [^{18}F]-FDG PET was performed by an individual intensity threshold (20-37% of total activity) optimally fitting the MRI brain border, in order to guarantee exclusion of all artifact voxels. In cases where these requirements could not be fulfilled, as occurred for <5% of all [^{18}F]-FDG PET images, the threshold was selected, which came most close to rescuing the assignment of all real brain voxels. Subsequently, the remaining artifact regions were cropped by setting to zero voxels outside the binary mask **(**Fig. **2**). For further analyses, the corrected MRI segmentations without artifacts were used for processing of the corresponding [^{18}F]-AV45 and [^{18}F]-FDG PET in PMOD PNEURO 3.5 tool again.

2.4.3. VOI-based PVEC

For each VOI defined in the Hammers atlas [25], we calculated PVE-corrected values [13, 26, 27]. The algorithm included an extra-cerebral background VOI implemented in the PNEURO 3.5 tool. We used a GM threshold 0.3 and a full-width-at-half-maximum (FWHM) defined of 8 mm; details of the PVEC algorithm are provided in Supplement **1**.

2.4.4. Analyses of Artifact Volumes and PET-VOI Data

Six GM VOIs were defined by combining bilateral Hammers atlas VOIs, i.e. frontal, parietal, temporal, and posterior cingulate cortex (PCC), a composite of all these cortical regions, as well as an entire cerebellum GM VOI. Uncorrected and PVE-corrected standard-uptake-values (SUV) were calculated for [^{18}F]-AV45 and [^{18}F]-FDG PET using non-masked (SUV_{art}) and masked (SUV_{true}) VOI sets. With these data the following analyses were performed:

2.4.4.1. Artifact Volume Quantitation

Artifact sizes were quantified by subtracting cropped from uncropped volumes. A volume threshold of </≥5.06 cm³ (≙1500 voxels) was defined for further subanalyses. Artifact volumes were estimated by a hot 3D VOI tool and 1500 voxels were the approximated lower bound of these artifacts volumes.

2.4.4.2. Error-[%] in PET SUVs and Correlation with Artifact Volumes

For each of the six VOIs, we calculated error-[%] in the uncorrected and PVE-corrected SUVs according to equation 1. Error-[%] in SUVs were then correlated with artifact volumes.

$$Error\text{-}[\%] = \frac{SUV_{art} - SUV_{true}}{SUV_{true}} \times 100 \qquad \text{Eq. 1}$$

2.4.4.3. Root-mean-square-error (RMSE)

RMSE were calculated for uncorrected and PVE-corrected SUVs according to equation 2:

$$RMSE\text{-}[\%] = \frac{\sqrt{\sum_{i=1}^{n}(SUV_{art} - SUV_{true})^2}}{\sqrt{\sum_{i=1}^{n}(SUV_{true})^2}} \times 100 \qquad \text{Eq. 2:}$$

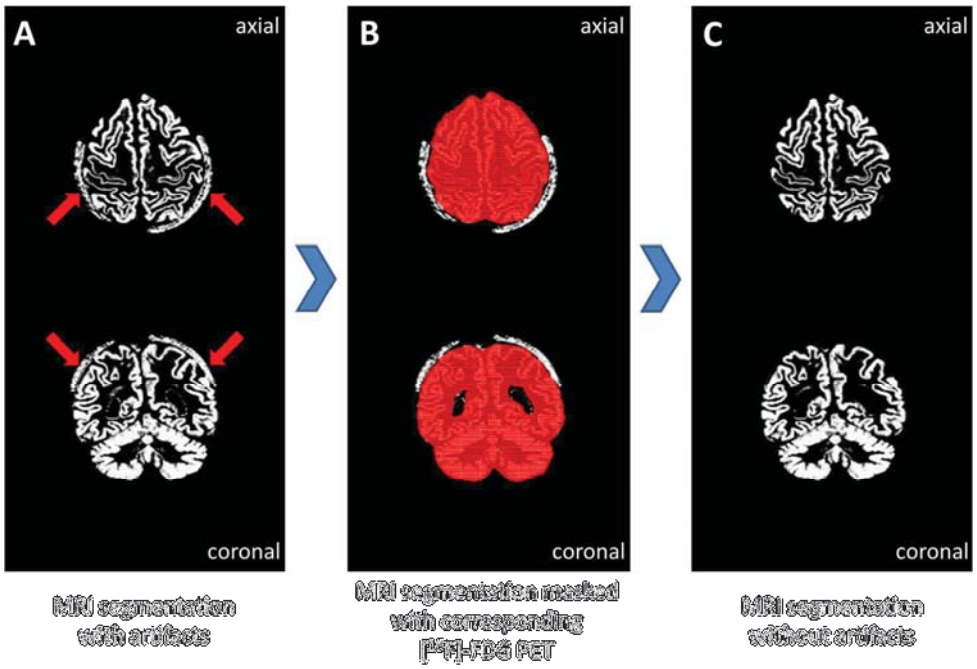

Fig. (2). Workflow of artifact masking. **(A)** Example of a segmentation artifact (highlighted by arrows) in axial and coronal plane. **(B)** Binarization of the corresponding [^{18}F]-FDG PET was performed by the individual threshold (20-37% of total activity) give the best fit to the MRI brain border. **(C)** Resulting artifact-free MRI segmentation after cropping process.

2.4.4.4. Longitudinal PET Assessment

Distribution of artifacts relatively to the different brain regions was assessed in the BL and FU study. Percentage changes in longitudinal SUVs (Δ-SUV-[%]) between BL and FU were calculated for the SUV data with artifacts (equation 3) or MR-artifact-free (equation 4) PET results.

$$\Delta\text{-}SUV\text{-}[\%]_{art} = \frac{SUV_{art}FU - SUV_{art}BL}{SUV_{art}BL} \times 100 \quad \text{Eq. 3}$$

$$\Delta\text{-}SUV\text{-}[\%]_{true} = \frac{SUV_{true}FU - SUV_{true}BL}{SUV_{true}BL} \times 100 \quad \text{Eq. 4}$$

The consecutive resultant bias in longitudinal PET data was calculated by Equation 5:

$$Bias\text{-}\Delta\text{-}SUV\text{-}[\%] = \Delta\text{-}SUV\text{-}[\%]_{true} - \Delta\text{-}SUV\text{-}[\%]_{art} \quad \text{Eq. 5}$$

2.4.5. Semi-automatic Masking Approach

In order to investigate a possibility method for facilitating and simplifying the masking process, we examined template masks deriving from the complete data set. In this approach, the masked segmentations from all 238 individual sets of VOIs deemed without artifacts were binarized as unity for all voxels included in any VOI. The 238 binarized images were then averaged to obtain a probability value for each voxel of "inclusion in the VOI-set" within the whole population, as has been done previously for probabilistic atlas generation [28]. Standardized masks were generated with either 100% (100%-voxel-mask) or 50% (50%-voxel-mask) probability level threshold. Thus, the 100%-voxel-mask contained every voxel of all VOI sets, while the much more stringent 50%-voxel-mask contained all voxels present in the VOI-set of at least half of the subjects. These masks were subsequently used to test the statistical template approach in the subjects for whom there were more than 1500 artifact voxels.

2.4.6. Statistics

Means (± SD) and ranges of volume as well as error-[%] in uncorrected and PVE-corrected SUVs were assessed for both tracers for all six VOI regions. Additionally Pearson's coefficient of correlation (r) was assessed for uncorrected and PVE-corrected SUV error-[%] with volumes of artifacts. Differences between uncorrected and PVE-corrected arising from MR segmentation artifacts were assessed by comparing single square errors with a paired student's *t*-test. Furthermore distinctions between different mask settings (raw, 50%- and 100%-voxel-mask) were as well assessed by a paired students *t*-test. The null hypothesis was rejected for $p<0.05$. In the longitudinal PET assessments the moduli of the mean (± SD) for Bias-Δ-SUV-[%] and Bias-Δ-VOL-[cm^3] were calculated and reported with the respective ranges.

3. RESULTS

3.1. Demographics

Detailed demographics of the study cohort are provided in Table **1**.

3.2. Artifact Volume Assessment and Error-[%] in SUV

In 66/238 (27.7%) of the MRIs, at least one artifact was identified by visual inspection, leading to a total of 276 artifacts, which were most frequent in the parietal region (N=84). Details of artifact volumes of all six assessed regions are presented in Table **2**, and the highest artifact volumes were found in parietal (mean: 2.9 ± 5.6 cm^3; range: 0 - 26.7 cm^3) and temporal regions (mean: 2.4 ± 2.9 cm^3; range:

Table 1. Demographics of healthy controls (HC), patients with mild cognitive impairment (MCI), and patients with Alzheimer's disease (AD).

Diagnosis	N	Age	Gender
(Baseline)		(y ± SD)	(f / m)
HC	36	77.4 ± 6.4	13 (36%) f / 23 (64%) m
MCI	74	70.2 ± 7.9	33 (45%) f / 41 (55%) m
AD	9	73.4 ± 8.2	6 (67%) f / 3 (33%) m

Table 2. Mean (±SD) and ranges (MIN / MAX) of artifact volumes, as well as counts (N) of artifacts in different brain regions. Correlation coefficients (r) are given for the relationship between artifact volumes and the error-[%] in SUV of [^{18}F]-AV45 and [^{18}F]-FDG PET, for uncorrected and PVE-corrected data.

Region	Artifact			Error-[%] in [^{18}F]-AV45 PET						Error-[%] in [^{18}F]-FDG PET					
	Volume [cm^3]			uncorrected			PVEC			uncorrected			PVEC		
	MEAN ± SD	MIN / MAX	N	MEAN ± SD	MIN / MAX	r	MEAN ± SD	MIN / MAX	r	MEAN ± SD	MIN / MAX	r	MEAN ± SD	MIN / MAX	r
Frontal	0.9 ± 1.5	0.0 / 11.5	51	-0.1 ± 0.5	-4.1 / 0.4	-0.88	-0.5 ± 1.2	-9.2 / 0.2	-0.94	-0.3 ± 0.6	-4.7 / 0.0	-0.99	-0.2 ± 0.9	-6.8 / 0.5	-0.88
Parietal	2.9 ± 5.6	0.0 / 26.7	84	-1.2 ± 2.8	-14.8 / 0.4	-0.94	-3.0 ± 7.0	-36.6 / 0.8	-0.90	-1.9 ± 3.7	-17.9 / 0.0	-0.99	-2.4 ± 5.9	-28.7 / 0.4	-0.98
Temporal	2.4 ± 2.9	0.0 / 17.5	67	-0.4 ± 0.8	-6.3 / 0.4	-0.82	-0.6 ± 1.3	-9.6 / 1.3	-0.68	-0.8 ± 1.2	-8.5 / 0.0	-0.96	-0.6 ± 1.5	-11.1 / 0.3	-0.84
PCC	0.1 ± 0.2	0.0 / 0.2	5	0.0 ± 0.0	-0.1 / 0.1	-0.25	0.4 ± 1.3	-0.5 / 6.6	-0.07	0.0 ± 0.0	-0.1 / 0.0	-0.94	0.4 ± 1.0	-0.1 / 5.6	-0.03
Composite	6.2 ± 8.2	0.1 / 38.9	207	-0.5 ± 1.0	-5.2 / 0.3	-0.89	-1.1 ± 2.2	-10.3 / 0.1	-0.84	-0.9 ± 1.3	-6.3 / 0.0	-0.98	-0.9 ± 1.9	-9.2 / 0.2	-0.94
Cerebellum	1.4 ± 1.3	0.0 / 5.4	69	-0.3 ± 0.5	-2.1 / 0.1	-0.82	-0.5 ± 0.9	-3.8 / 1.4	-0.76	-0.7 ± 0.6	-2.8 / 0.0	-0.98	-0.5 ± 0.6	-2.6 / 0.1	-0.90

0 - 17.5 cm^3). Volumes for PCC were negligibly affected by segmentation artifacts.

Strongest correlations of artifact volumes with error-[%] in SUV were detected in the parietal cortex region for [^{18}F]-AV45 PET (uncorrected: r = -0.94; PVEC: r = -0.90) (Fig. **3A**) as well as for [^{18}F]-FDG PET (uncorrected: r = -0.99; PVEC: r = -0.98) (Fig. **3B**). Considerable error-[%] in SUV were detected above the defined artifact volume threshold of 5.06 cm³ (≙1500 voxels), which was present in 9/66 segmentations with artifacts. However there was no relationship between artifact sizes and clinical diagnoses.

3.3. RMSE-[%] in uncorrected and PVE-corrected PET SUVs

Highest RMSE-[%] in SUV was observed in the parietal region. For both tracers, PVE-corrected data was significantly more affected by artifacts when compared to uncorrected SUVs ($p<0.05$), as illustrated in (Fig. **4A/B**), *e.g.* in the parietal region ([^{18}F]-AV45 PET: 7.1% *vs*. 2.6% / [^{18}F]-FDG PET: 6.3% *vs*. 4.2%). Furthermore, RMSE-[%] results were analysed for subgroups with </≥1500 artifact voxels (57 segmentations with <1500 artifact voxels and 9 segmentations with ≥1500 artifact voxels). The small group with artifacts of ≥1500 voxels nonetheless accounted for the largest share of total RMSE-[%] in SUV for both tracers (Fig. **4C**), while RMSE-[%] in all regions were consistently below 1.5% for the subgroup with <1500 voxels, irrespective of radiotracer or whether or not PVEC was applied.

3.4. Distribution of Artifacts and Impact on Longitudinal PET Studies

The 66 segmentations with artifacts were obtained from 39 different subjects (27 of whom had artifacts in their MR-segmentations both at BL and FU). The latter group comprised a regional concordance of 88/112 (78.6%) artifacts at both MR assessments. In the longitudinal PET measure of Bias-Δ-SUV-[%], again a small fraction of subjects (Bias-Δ-VOL-[cm^3] ≥1500 voxels) had the highest impact on PET quantitation, which increased as a function of artifact size (Fig. **5 A,B**).

Highest Bias-Δ-SUV-[%] of up to 36% was found in parietal region of PVE-corrected [^{18}F]-AV45 PET (Table **3**). For both PET tracers, longitudinal assessments were more affected by MR segmentation artifacts in PVE-corrected SUVs than for uncorrected SUVs ([^{18}F]-AV45: $p<0.001$ / [^{18}F]-FDG: $p<0.05$). The group of 34 subjects with a Bias-Δ-VOL-[cm^3] <1500 voxels revealed a negligible mean Bias-Δ-SUV-[%] to FU even in the most affected region, i.e. parietal

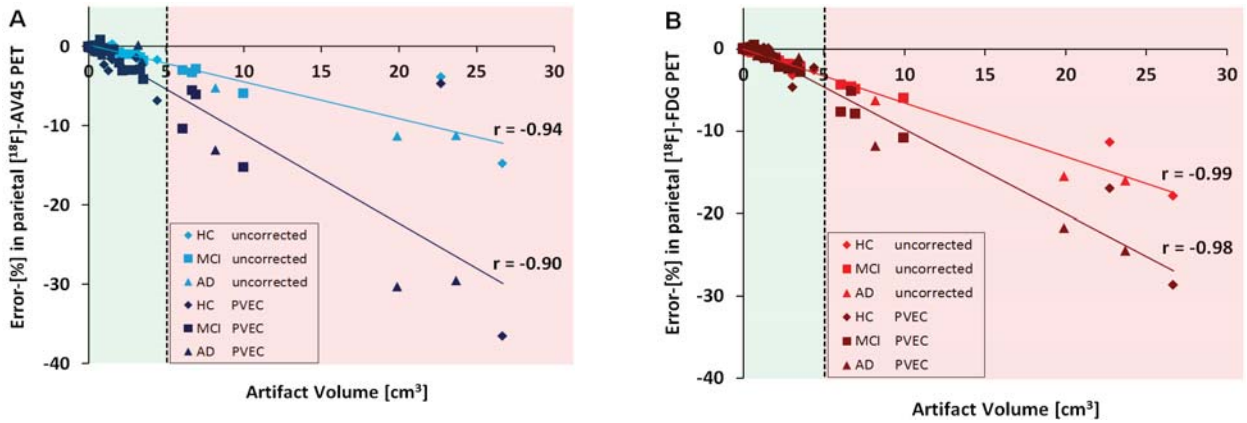

Fig. (3). (A) Correlation of artifact volumes [cm^3] with error-[%] in uncorrected (light blue) and PVEC (dark blue) for [^{18}F]-AV45 PET in parietal cortex. **(B)** Correlation of artifact volumes [cm^3] with error-[%] in uncorrected (light red) and PVEC (dark red) for [^{18}F]-FDG PET in parietal cortex. In both charts, single subjects are depicted by diamond (HC), square (MCI), and triangle (AD). Pearson's correlation coefficients (r) are provided. Dotted line represents the threshold above 5.06 cm^3 (≙1500 voxels, red area), which had considerable impact on PET quantitation for both tracers.

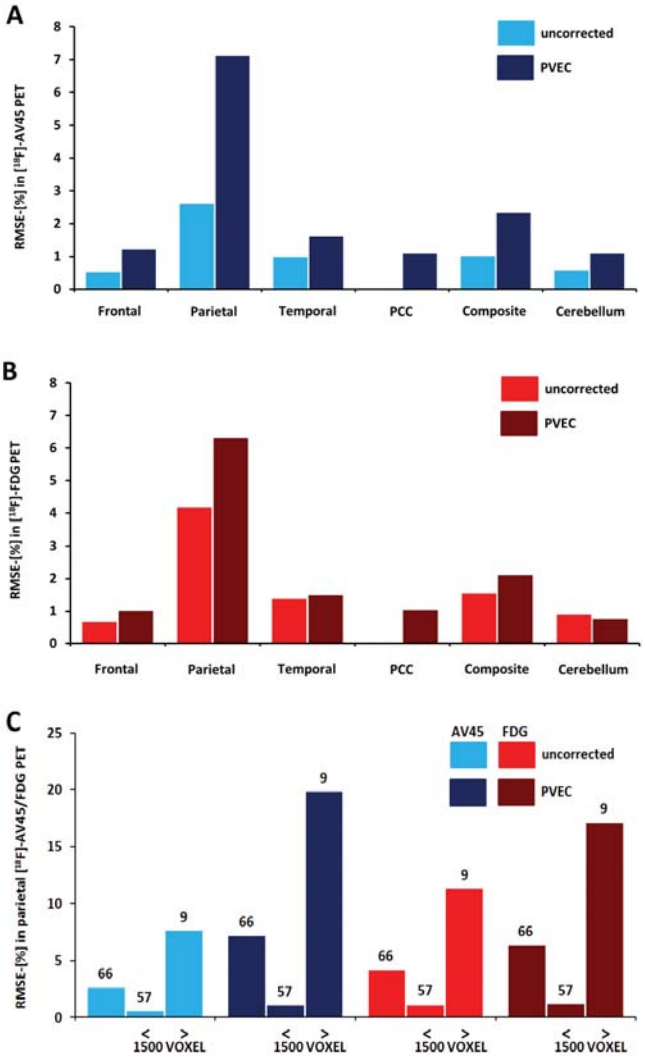

Fig. (4). (A) RMSE-[%] in [^{18}F]-AV45 PET SUV caused by artifacts for uncorrected (light blue) and PVE-corrected (dark blue) values. **(B)** RMSE-[%] in [^{18}F]-FDG PET SUV caused by artifacts for uncorrected (light red) and PVE-corrected (dark red) values. **(C)** RMSE-[%] of uncorrected (light blue/red) and PVEC (dark blue/red) values of the parietal cortex VOI before (N=66) and after stratifying in subgroups containing <1500 (N=57) and ≥1500 (N=9) artifact voxels.

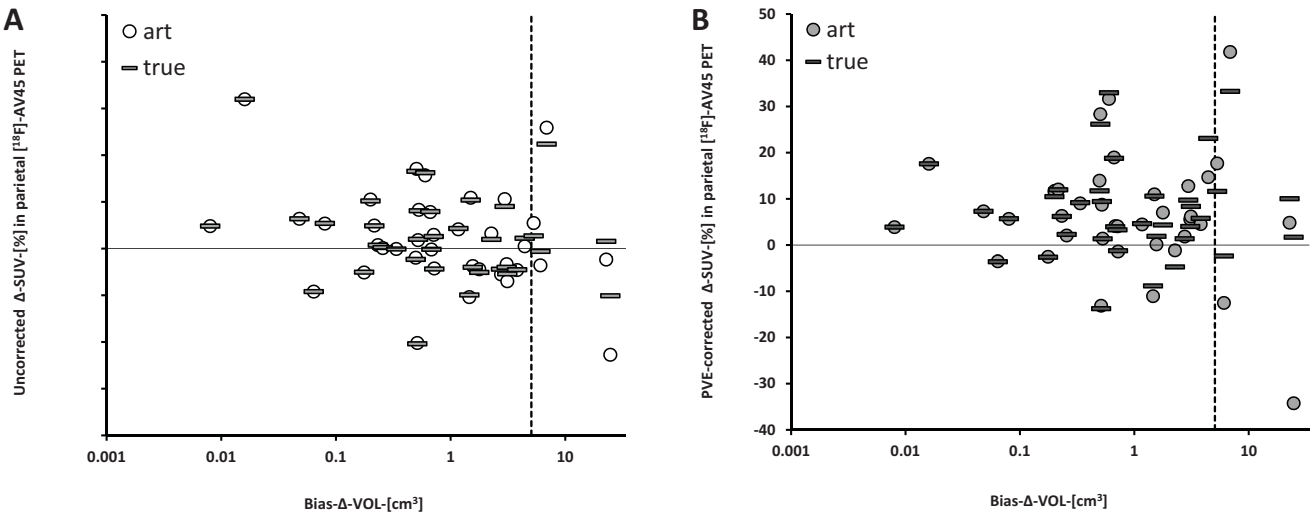

Fig. (5). (A) Discrepancies in two year follow-up PET results of single subjects as expressed by true (light grey bars) and biased (white circles) Δ-SUV-[%] (uncorrected [^{18}F]-AV45 PET) are presented as a logarithmic function of longitudinal apparent MR-volume differences (Bias-Δ-VOL-[cm^3]). **(B)** Discrepancies in two year follow-up of PET results of single subjects as expressed by true (dark grey bars) and biased (light grey circles) Δ-SUV-[%] (PVE-corrected [^{18}F]-AV45 PET) are presented as a logarithmic function of longitudinal apparent MR volume differences (Bias-Δ-VOL-[cm^3]). Dotted line represents the cut-off threshold of 5.06 cm^3 (≙1500 voxels). Parietal cortex region serves as the illustrative example.

Table 3. Mean (±SD; modulus) and MIN / MAX for two-year longitudinal Bias-Δ-VOL-[cm^3] as well as Bias-Δ-SUV-[%] of [^{18}F]-AV45 and [^{18}F]-FDG. Results are given separately for uncorrected and PVE-corrected Bias-Δ-SUV-[%]. Significant differences between uncorrected and PVE-corrected data are indicated by * p < 0.05; ** p < 0.01; *** p < 0.001.

Region	Bias-Δ-VOL-[cm^3]		Bias-Δ-SUV-[%] in [^{18}F]-AV45 PET				Bias-Δ-SUV-[%] in [^{18}F]-FDG PET			
			uncorrected		PVEC		uncorrected		PVEC	
	□MEAN ± SD□	MIN / MAX	□MEAN ± SD□	MIN / MAX	□MEAN ± SD□	MIN / MAX	□MEAN ± SD□	MIN / MAX	□MEAN ± SD□	MIN / MAX
Frontal	0.8 ± 1.4	-2.6 / 8.5	0.2 ± 0.5	-3.0 / 0.7	0.4 ± 0.9**	-5.9 / 1.4	0.3 ± 0.6	-4.1 / 0.8	0.3 ± 0.9	-5.8 / 0.6
Parietal	2.6 ± 5.2	-24.7 / 6.9	1.0 ± 2.1	-3.5 / 12.6	2.7 ± 6.0*	-8.5 / 35.9	1.6 ± 3.2	-4.9 / 17.3	2.4 ± 5.4	-7.7 / 30.2
Temporal	2.3 ± 3.0	-11.5 / 13.6	0.4 ± 0.8	-5.0 / 1.8	0.7 ± 1.2**	-7.3 / 2.2	0.8 ± 1.3	-7.2 / 3.2	0.7 ± 1.6	-9.9 / 2.9
PCC	0.1 ± 0.2	0.0 / 0.2	0.0 ± 0.0	-0.1 / 0.1	0.4 ± 0.9**	-5.5 / 1.4	0.0 ± 0.0	0.0 / 0.2	0.4 ± 0.8**	-4.3 / 1.9
Composite	5.7 ± 7.6	-36.5 / 23.9	0.4 ± 0.7	-3.2 / 3.5	0.9 ± 1.7**	-5.4 / 9.3	0.8 ± 1.1	-4.0 / 4.7	0.9 ± 1.6	-5.8 / 7.2
Cerebellum	1.3 ± 1.2	-4.9 / 4.1	0.3 ± 0.4	-1.0 / 1.9	0.7 ± 0.9***	-3.1 / 4.1	0.6 ± 0.5	-2.1 / 2.2	0.5 ± 0.5**	-2.1 / 2.1

cortex: [^{18}F]-AV45 uncorrected: -0.02%; PVE-corrected: -0.02%; [^{18}F]-FDG uncorrected: +0.09%; PVE-corrected: -0.16%.

3.5. Semi-automatic Masking Approach

By application of the statistically defined 100%-voxel-mask, only minor effects on artifact volumes were observed, whereas the more stringent 50%-voxel-mask led to a considerable reduction of mean parietal cortex segmentation artifact volumes (14.1 cm³ vs. 3.4; p<0.001). However, the 50%-voxel-mask resulted in distinct underestimation of true brain voxels in the frontal region (-7.6 cm³) (Fig. **6A**). Despite the reduction of parietal cortex artifact volumes, SUV RMSE-[%], were reduced by only 50% in uncorrected and PVE-corrected PET of both tracers (Fig. **6B**) when using the strict 50%-voxel-mask.

4. DISCUSSION

This is the first systematic investigation of the impact of MRI segmentation-related artifacts on the quantitation of brain PET data. Longitudinal [^{18}F]-AV45 and [^{18}F]-FDG PET recordings from the ADNI database were analyzed according to uptake in GM VOIs derived from MRI segmentation, with and without correction for segmentation artifacts.

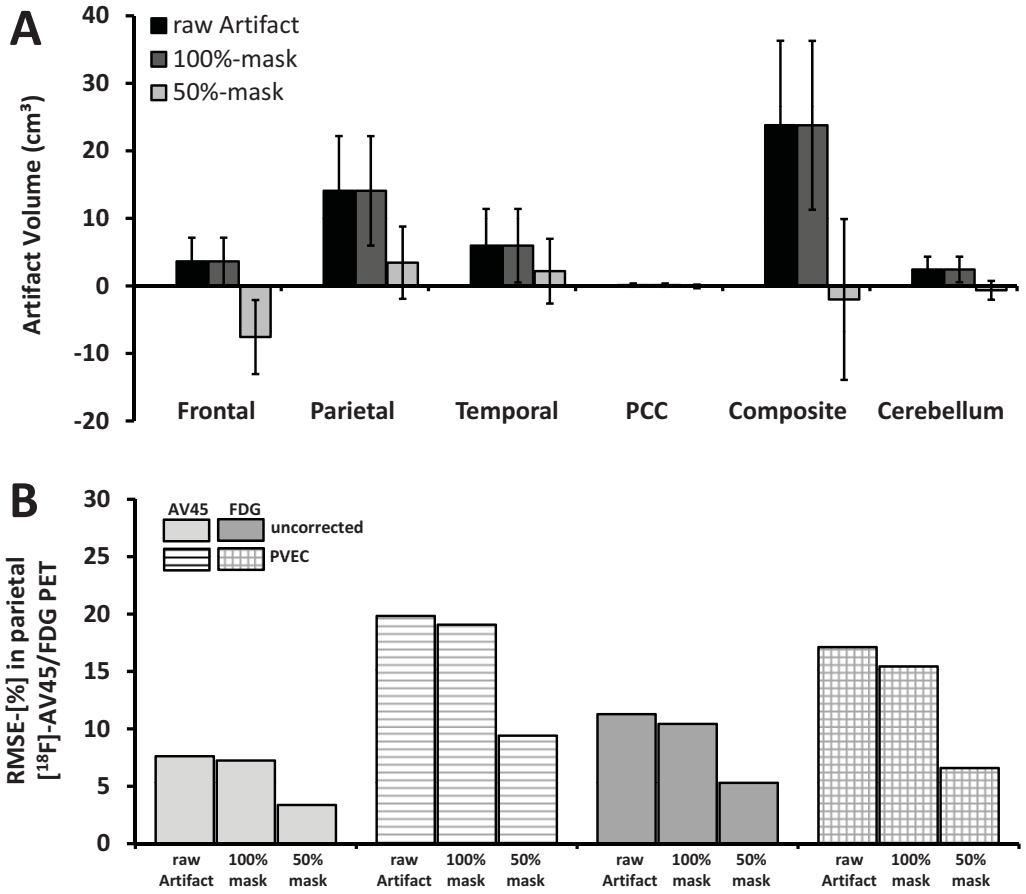

Fig. (6). (A) Mean (±SD) artifact volumes either for raw (black bars) segmentations or after standardized masking by the 100% probabilistic-mask (darker grey bars) and the more stringent 50%-mask (lighter grey bars) for six different brain regions (expressed for the nine MR segmentations with extremely large artifacts (≥1500 voxels)).**(B)** RMSE-[%] in [^{18}F]-AV45 and [^{18}F]-FDG PET SUV for the nine MR segmentations with extremely large artifacts (≥1500 voxels). RMSE-[%] for the parietal cortex region are given either for raw segmentations or after standardized masking by the 100%-mask and 50%-mask for uncorrected and PVE-corrected values.

MRI-based segmentation artifacts propagated to bias in PET quantitation in proportion to the volume of segmentation errors; this effect was most pronounced in PVE-corrected data. There was considerable variance in the magnitude of bias in individual SUVs, with substantial impact on the group results arising from the few subjects with especially large artifacts; identification of such patients is required in order to avoid biased PET quantitation in multicenter cohort studies.

4.1. Artifact Assessment and Masking

We found that one quarter of processed segmentations revealed at least one artifact by visual quality control, indicating a high frequency of relevant segmentation artifacts in this multi-center collective. The parietal cortex region proved to be most vulnerable to artifacts, most likely due to inherently greater likelihood of bridge building VOIs adjacent to this brain area. This is concordant with previous studies showing the meninges and large sinuses, which are notably present in vicinity to the parietal region, were the non-brain structures most often falsely segmented in brain MRI [16]. The rather sophisticated masking method used in this investigation afforded identification of the optimal threshold for masking of VOIs. Through careful editing, nearly perfect masking of artifacts was achievable, in a trade-off between the most rigorous possible exclusion of all artifact voxels, and preservation of all real cortical voxels in the VOI. As an alternative, we also tested a semiautomatic masking approach, but this proved without benefit, as artifact voxels were not sufficiently excluded by a moderate stringency threshold mask, or bias was introduced by elimination of real cortical voxels when a strict threshold mask was applied; we found that RMSE-[%] in the PET result was reduced by only 50% when using the strict mask. This reflects the location of most artifacts in the direct vicinity to the brain, which brings a risk of overlap between true GM and artifact voxels in the MNI space when considering subjects with differing brain shape and atrophy. Therefore, we conclude that standardized semiautomatic masks are no substitute for individual masking to resolve segmentation artifacts by post-processing. Correct masking is aided by access to the individual [^{18}F]-FDG PET image, which is not available in many brain PET studies. When [^{18}F]-FDG PET is not available, masking is also possible through time-consuming manual identification and erasing of artifact voxels using VOI or ROI post-processing tools.

Despite the presence of extra-cerebral artifacts, overestimation of GM compared to WM (mis-segmentation) is well-

known to arise in MRI segmentation algorithms. This was examined in detail by Gutierrez and colleagues in [^{18}F]-FDG PETs of 19 patients with dementia, and they concluded that statistical parametric mapping (SPM) 5, which was also implemented in the PMOD algorithm used in the current study, provides reliable results, superior to those obtained with SPM 99 or SPM 2 [29]. Nonetheless, we found 9% completely failed MRI segmentations, which is also in line with our previous findings in a larger ADNI dataset of 1300 subjects [30]; this could reflect the challenge presented by atrophy in the aged or disease brain relative to the MNI target, which represents healthy young subjects.

4.2. Influence of Segmentation Artifacts on PET Quantitation

Bias in PET quantitation correlated highly with the individual volume of artifacts. This is compelling, as artifact voxels are predominantly located outside the cerebral cortex, and thus have very low radiotracer uptake, which consequently exert a large bias on the apparent VOI activity. Therefore, spuriously low SUV values are the logical consequence of MR segmentation error, and are most pronounced when brain tracer uptake is relatively high compared to the extracerebral background activity.

One objective of this study was to establish the size of artifacts leading to relevant errors in PET quantitation, with likely consequences for between-group comparisons. This is a matter of interest as the extensive quality control performed in this investigation is labour intensive, and might not be feasible in planned studies with many hundreds of subjects. We found artifacts comprising <1500 voxels to be of minor relevance in PET quantitation at a single time point, whereas artifacts larger than 1500 voxels (5 cm^3) in specific brain VOIs had a considerable impact on PET quantitation. Thus, it is very important to take into account the target region sizes of PET studies, as quantitation in small brain regions is inherently more vulnerable to MR segmentation artifacts. Thus for example, amyloid-PET [^{18}F]-AV45 will scarcely be affected if the end point is semi-quantitative assessment of amyloid positivity or negativity within the entire cortical GM VOI, which comprises 143,000 voxels. In contrast, [^{18}F]-FDG PET analysis in a parietal subregion comprising as few as 4,500 voxels is highly vulnerable to the presence of a large artifact.

From our experience, all cases with segmentation artifacts exceeding the 1500 voxel threshold were easily discerned by simply scrolling through the axial layer (< 1 min), which should enable rapid screening of segmentations requiring correction.

4.3. Impact of Segmentation Artifacts on PVEC

Accounting for effects of atrophy is crucial in brain PET studies of healthy aging and AD, as the loss of cortical volume will bias observations of metabolism and ligand binding due to increasing PVEC (12). Particular implications of PVEC on [^{18}F]-flutemetamol/amyloid PET quantitation were recently assessed in a phantom study employing a new region-based voxel-wise technique, in comparison to well-established VOI-based methods [9]. The authors concluded that PVEC enhances precision of regional amyloid PET quantitation, and should be considered as obligatory in clinical trials with PET endpoints. In the present work, we found that PET quantitation with PVEC was more affected (than was native, uncorrected PET data) by segmentation artifacts, regardless of the subregion or the radiotracer. This is expected from the methodological principles of VOI-based PVEC, since the presence of (extracerebral) artifacts will tend to create the appearance of hypertrophy in the GM. In this circumstance, a spuriously low correction PVEC is performed by the algorithm in artifact-contaminated regions, which was reflected by the greater underestimation of PET SUV values with PVEC when compared to uncorrected data (Figs. **3 A, B**). This result is in line with findings in previous works, where (other than misregistration) mis-segmentation (over/underestimation of GM) had the most pronounced impact on region-based [31] and/or voxel-based [32, 33] PVEC accuracy. Efforts towards detection and elimination of segmentation artifacts are consequentially justified when PVEC is implemented in the data analysis.

4.4. Longitudinal Studies

Monitoring of therapeutic trials in neurodegenerative diseases with well-established non-invasive biomarkers like MRI and [^{18}F]-FDG PET is of burgeoning importance [34]. However, these longitudinal investigations tend to suffer from low effect sizes magnitudes, even for therapeutic interventions lasting for more than two years of disease progression. This is exemplified in a multimodal image study of a trial of memantine in mild to moderate AD [35]; in the course of one year, cerebral glucose metabolism decreased by a mean of only -1.8% in memantine-treated patients versus -3.0% in the placebo group. It is self-evident that a longitudinal single subject bias (over two years' follow-up) of 17% in the uncorrected [^{18}F]-FDG PET uptake, as we observed in parietal cortex, the region most vulnerable to segmentation artifact of our study, would detrimentally influence such results, if not resolved or excluded. Nevertheless, the overall impact of smaller artifacts should normally be compensated by sufficient sample sizes; we found the mean Bias-Δ-SUV-[%] in [^{18}F]-FDG PET to be less than 0.2% (for the highest affected parietal region) in the sub-group of 34 subjects with a Bias-Δ-VOL-[cm^3] <1500 voxels.

We found a high concordance of regional artifact distribution to baseline and follow-up segmentations in the two-year longitudinal setting (79%). Thus, we conclude that individual brain anatomy contributes to the occurrence of artifacts. We did not systematically investigate the possible influence of specific imaging sites on the magnitude of artifacts in this multi-center study, as we processed MRIs only when dual tracer PET imaging was also present. We did note that some sites tended to produce more MRIs that were affected by segmentation artifacts in consecutive post-processing (max. 50%; 10/20), but this effect of the site was not statistically significant.

4.5. Limitations

In the interests of clarity, we restricted our investigation only to VOI-based PVEC as implementing another PVEC approach (*e.g.* Müller-Gärtner method). However, pilot data (not shown) suggested that the effects of artifacts on voxel-

wise MRI-based PVEC tended to mirror our findings with VOI-based PVEC. This is unsurprising, as the bias of enlarged GM volumes remains constant for both approaches. The aim of this study was rather to emphasize the direct impact and consequences of MR segmentation artifacts in commonly used multi-center data by application of a standard (and widely used) PMOD algorithm. However, segmentation of other MRI sequences and use of different segmentation algorithms need to be explored in more detail. More sophisticated segmentation approaches would potentially lower the frequency of relevant artifacts, as specialized artifact-reducing features are increasingly implemented in those algorithms (*e.g.* bridge burner).

CONCLUSION

MRI-based segmentation artifacts have a size-dependent effect on PET quantitation in this multitracer follow-up study of HC, MCI and AD subjects. Individual SUVs were affected to a highly variable extent, with very substantial impact in those few subjects with especially large artifacts. When applying PVEC algorithms, the impact is even more pronounced than for uncorrected studies. Screening and editing of suspect MR segmentations is necessary to avoid biased PET quantitation. If available, [^{18}F]-FDG-PET is suitable for masking segmentation-based artifacts in such cases, whereas a standardized template based masking approach was inferior.

ABBREVIATIONS

AD	=	Alzheimer's disease
ADNI	=	Alzheimer's disease neuroimaging initiative
BL	=	baseline
CSF	=	cerebrospinal fluid
FDA	=	food and drug administration
[^{18}F]-FDG	=	[^{18}F]-fluorodesoxyglucose
FU	=	follow-up
FWHM	=	full-width-at-half-maximum
GTM	=	geometric transfer matrix
GM	=	grey matter
HC	=	healthy control
LONI	=	laboratory of neuro imaging
MCI	=	mild cognitive impairment
MNI	=	montreal neurological institute
MRI	=	magnetic resonance imaging
NIA	=	national institute on aging
NIBIB	=	national institute of biomedical imaging and bioengineering
PCC	=	posterior cingulate cortex
PET	=	positron-emission-tomography
PVE(C)	=	partial volume effect (correction)
RMSE	=	root-mean-square-error
SPECT	=	single-photon-emission-computed-tomography
SPM	=	statistical parametric mapping
SUV	=	standard-uptake-value
Δ-SUV-[%]	=	change in longitudinal SUV
T1w	=	T1-weighted
VOI	=	volume-of-interest
VOL	=	volume
WM	=	white matter

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's web site along with the published article.

CONFLICT OF INTEREST

M.H., M.B., A.D., S.D., M.W., P.B. and A.R. report no disclosures.

ACKNOWLEDGEMENTS

Data collection and sharing for 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: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; 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.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. 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. We acknowledge profession textual revisions by Inglewood Biomedical Editing.

REFERENCES

[1] Anderson VM, Schott JM, Bartlett JW, Leung KK, Miller DH, Fox NC. Gray matter atrophy rate as a marker of disease progression in AD. Neurobiol Aging 33(7): 1194-202 (2012).

[2] Whitwell JL, Przybelski SA, Weigand SD, Knopman DS, Boeve BF, Petersen RC, *et al*. 3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. Brain 130(Pt 7): 1777-86 (2007).

[3] Sepulcre J, Sastre-Garriga J, Cercignani M, Ingle GT, Miller DH, Thompson AJ. Regional gray matter atrophy in early primary progressive multiple sclerosis: a voxel-based morphometry study. Arch Neurol 63(8): 1175-80 (2006).

[4] Ciumas C, Savic I. Structural changes in patients with primary generalized tonic and clonic seizures. Neurology 67(4): 683-6 (2006).

[5] Hutcheson NL, Clark DG, Bolding MS, White DM, Lahti AC. Basal ganglia volume in unmedicated patients with schizophrenia is associated with treatment response to antipsychotic medication. Psychiatr Res 221(1): 6-12 (2014).

[6] Oertel-Knochel V, Knochel C, Matura S, Rotarska-Jagiela A, Magerkurth J, Prvulovic D, *et al.* Cortical-basal ganglia imbalance in schizophrenia patients and unaffected first-degree relatives. Schizophr Res 138(2-3): 120-7 (2012).

[7] Kawachi T, Ishii K, Sakamoto S, Sasaki M, Mori T, Yamashita F, *et al.* Comparison of the diagnostic performance of FDG-PET and VBM-MRI in very mild Alzheimer's disease. Eur J Nucl Med Mol Imaging 33(7): 801-9 (2006).

[8] Goto H, Ishii K, Uemura T, Miyamoto N, Yoshikawa T, Shimada K, *et al*. Differential diagnosis of dementia with Lewy Bodies and Alzheimer Disease using combined MR imaging and brain perfusion single-photon emission tomography. AJNR Am J Neuroradiol 31(4): 720-5 (2010).

[9] Thomas BA, Erlandsson K, Modat M, Thurfjell L, Vandenberghe R, Ourselin S, *et al.* The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease. Eur J Nucl Med Mol Imaging 38(6): 1104-19 (2011).

[10] Samuraki M, Matsunari I, Chen WP, Yajima K, Yanase D, Fujikawa A, *et al*. Partial volume effect-corrected FDG PET and grey matter volume loss in patients with mild Alzheimer's disease. Eur J Nucl Med Mol Imaging 34(10): 1658-69 (2007).

[11] Forster S, Yousefi BH, Wester HJ, Klupp E, Rominger A, Forstl H, *et al*. Quantitative longitudinal interrelationships between brain metabolism and amyloid deposition during a 2-year follow-up in patients with early Alzheimer's disease. Eur J Nucl Med Mol Imaging 39(12): 1927-36 (2012).

[12] Erlandsson K, Buvat I, Pretorius PH, Thomas BA, Hutton BF. A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. Phys Med Biol 57(21): R119-59 (2012).

[13] Rousset OG, Ma Y, Evans AC. Correction for partial volume effects in PET: principle and validation. J Nucl Med 39(5): 904-11 (1998).

[14] Muller-Gartner HW, Links JM, Prince JL, Bryan RN, McVeigh E, Leal JP, *et al*. Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects. J Cereb Blood Flow Metab 2(4): 571-83 (1992).

[15] Leung KK, Barnes J, Modat M, Ridgway GR, Bartlett JW, Fox NC, *et al*. Brain MAPS: an automated, accurate and robust brain extraction technique using a template library. NeuroImage 2011 55(3): 1091-108 (2011).

[16] Mikheev A, Nevsky G, Govindan S, Grossman R, Rusinek H. Fully automatic segmentation of the brain from T1-weighted MRI using Bridge Burner algorithm. J Magn Reson Imaging 2008 27(6): 1235-41 (2008).

[17] Smith SM. Fast robust automated brain extraction. Hum Brain Mapp 17(3): 143-55 (2002).

[18] Segonne F, Dale AM, Busa E, Glessner M, Salat D, Hahn HK, *et al*. A hybrid approach to the skull stripping problem in MRI. NeuroImage 22(3): 1060-75 (2004).

[19] Gray KR, Wolz R, Heckemann RA, Aljabar P, Hammers A, Rueckert D, *et al*. Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease. Neuroimage 60(1): 221-9 (2012).

[20] Sanabria-Diaz G, Martinez-Montes E, Melie-Garcia L. Alzheimer's Disease Neuroimaging I. Glucose metabolism during resting state reveals abnormal brain networks organization in the Alzheimer's disease and mild cognitive impairment. PLoS One 8(7): e68860 (2013).

[21] Jack CR, Jr., Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, *et al.* The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Mag Res Imaging 27(4): 685-91 (2008).

[22] Jovicich J, Czanner S, Greve D, Haley E, van der Kouwe A, Gollub R, *et al*. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. NeuroImage 30(2): 436-43 (2006).

[23] Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans Med Imaging 17(1): 87-97 (1998).

[24] Ashburner J, Friston KJ. Unified segmentation. NeuroImage 26(3): 839-51 (2005).

[25] Hammers A, Allom R, Koepp 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 19(4): 224-47 (2003).

[26] Rousset OG, Collins DL, Rahmim A, Wong DF. Design and implementation of an automated partial volume correction in PET: application to dopamine receptor quantification in the normal human striatum. J Nucl Med 49(7): 1097-106 (2008).

[27] Rousset OG, Deep P, Kuwabara H, Evans AC, Gjedde AH, Cumming P. Effect of partial volume correction on estimates of the influx and cerebral metabolism of 6-[(18)F]fluoro-L-dopa studied with PET in normal control and Parkinson's disease subjects. Synapse 37(2): 81-9 (2000).

[28] Watanabe H, Andersen F, Simonsen CZ, Evans SM, Gjedde A, Cumming P, *et al.* MR-based statistical atlas of the Gottingen minipig brain. Neuroimage 14(5): 1089-96 (2001).

[29] Gutierrez D, Montandon ML, Assal F, Allaoua M, Ratib O, Lovblad KO, *et al*. Anatomically guided voxel-based partial volume effect correction in brain PET: impact of MRI segmentation. Com Med Imaging Graph 36(8): 610-9 (2012).

[30] Brendel M, Hogenauer M, Delker A, Sauerbeck J, Bartenstein P, Seibyl J, *et al.* Improved Longitudinal [18F]-AV45 Amyloid PET by White Matter Reference and VOI-based Partial Volume Effect Correction. Neuroimage 108: 450-9 (2015).

[31] Frouin V, Comtat C, Reilhac A, Gregoire MC. Correction of partial-volume effect for PET striatal imaging: fast implementation and study of robustness. J Nucl Med 43(12): 1715-26 (2002).

[32] Strul D, Bendriem B. Robustness of anatomically guided pixel-by-pixel algorithms for partial volume effect correction in positron emission tomography. J Cereb Blood Flow Metabol 19(5): 547-59 (1999).

[33] Quarantelli M, Berkouk K, Prinster A, Landeau B, Svarer C, Balkay L, *et al.* Integrated software for the analysis of brain PET/SPECT studies with partial-volume-effect correction. J Nucl Med 45(2): 192-201 (2004).

[34] Jack CR, Jr., 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 9(1): 119-28 (2010).

[35] Schmidt R, Ropele S, Pendl B, Ofner P, Enzinger C, Schmidt H, *et al*. Longitudinal multimodal imaging in mild to moderate Alzheimer disease: a pilot study with memantine. J Neurol Neurosurg Psychiatry 79(12): 1312-7 (2008).

Received: February 05, 2015 **Revised: October 10, 2015** **Accepted: October 16, 2015**