

# Reproducibility of automated simplified voxel-based analysis of PET amyloid ligand [ $^{11}\text{C}$ ]PIB uptake using 30-min scanning data

Sargo Aalto · Noora M. Scheinin ·  
Nina M. Kemppainen · Kjell Någren ·  
Marita Kailajärvi · Mika Leinonen · Mika Scheinin ·  
Juha O. Rinne

Received: 31 December 2008 / Accepted: 8 May 2009 / Published online: 4 June 2009  
© Springer-Verlag 2009

## Abstract

**Purpose** Positron emission tomography (PET) with  $^{11}\text{C}$ -labelled Pittsburgh compound B ( $^{11}\text{C}$ ]PIB) enables the quantitation of  $\beta$ -amyloid accumulation in the brain of patients with Alzheimer's disease (AD). Voxel-based image analysis techniques conducted in a standard brain space provide an objective, rapid and fully automated method to

analyze [ $^{11}\text{C}$ ]PIB PET data. The purpose of this study was to evaluate both region- and voxel-level reproducibility of automated and simplified [ $^{11}\text{C}$ ]PIB quantitation when using only 30 min of imaging data.

**Methods** Six AD patients and four healthy controls were scanned twice with an average interval of 6 weeks. To evaluate the feasibility of short scanning (convenient for AD patients), [ $^{11}\text{C}$ ]PIB uptake was quantitated using 30 min of imaging data (60 to 90 min after tracer injection) for region-to-cerebellum ratio calculations. To evaluate the reproducibility, a test-retest design was used to derive absolute variability (VAR) estimates and intra-class correlation coefficients at both region-of-interest (ROI) and voxel level.

**Results** The reproducibility both at the region level (VAR 0.9–5.5%) and at the voxel level (VAR 4.2–6.4%) was good to excellent. Based on the variability estimates obtained, power calculations indicated that 90% power to obtain statistically significant difference can be achieved using a sample size of five subjects per group when a 15% change from baseline (increase or decrease) in [ $^{11}\text{C}$ ]PIB accumulation in the frontal cortex is anticipated in one group compared to no change in another group.

**Conclusion** Our results showed that an automated analysis method based on an efficient scanning protocol provides reproducible results for [ $^{11}\text{C}$ ]PIB uptake and appears suitable for PET studies aiming at the quantitation of amyloid accumulation in the brain of AD patients for the evaluation of progression and treatment effects.

---

S. Aalto · N. M. Scheinin · N. M. Kemppainen · K. Någren ·  
J. O. Rinne  
Turku PET Centre,  
University of Turku and Turku University Hospital,  
Turku, Finland

S. Aalto  
Department of Psychology, Åbo Akademi University,  
Turku, Finland

N. M. Kemppainen  
Department of Neurology, Turku University Hospital,  
Turku, Finland

M. Kailajärvi · M. Scheinin  
Department of Pharmacology and Clinical Research Services  
Turku (CRST), University of Turku,  
Turku, Finland

M. Kailajärvi  
Turku Imanet, GE Healthcare,  
Turku, Finland

M. Leinonen  
4-Pharma Ltd,  
Turku, Finland

S. Aalto (✉)  
Turku PET Centre, University of Turku,  
P.O. Box 52, FI-20521 Turku, Finland  
e-mail: sargo.aalto@utu.fi

**Keywords** Alzheimer's disease · PET · PIB ·  
Reproducibility · Test-retest · Voxel-based

## Introduction

Positron emission tomography (PET) with  $^{11}\text{C}$ -labelled Pittsburgh compound B ( $^{11}\text{C}$ ]PIB) has been shown to enable the quantitation and visualization of the accumulation of  $\beta$ -amyloid ( $\text{A}\beta$ ) in the brain of patients with Alzheimer's disease (AD) [1–4]. However, often in human PET studies with  $^{11}\text{C}$ ]PIB, data analysis has been based on traditional manual region-of-interest (ROI) analysis (see, for example, references [1] and [4]), which is laborious and liable to operator-induced bias. Using modern analysis techniques including voxel-based quantitation and statistical parametric mapping (SPM), the regions of statistically significant differences in tracer uptake can be visualized as a demonstrative brain map, and using additional automated ROI analysis of parametric images, tracer uptake can be regionally quantified, making individual manual ROI analysis unnecessary.

It has recently been demonstrated that automated voxel-based image analysis methods are feasible for  $^{11}\text{C}$ ]PIB imaging data [2, 5]. In addition, a recent study has demonstrated that the region-level reproducibility of automated ROI analysis with several modelling approaches of  $^{11}\text{C}$ ]PIB quantitation is generally excellent [6]. Knowledge of the region-level reproducibility is important when evaluating whether conventional quantitation using manual ROI analysis can be replaced by automated ROI analysis conducted in a standardized brain space [2, 5, 6]. On the other hand, the test-retest reproducibility of the quantitation of  $^{11}\text{C}$ ]PIB uptake at the voxel level, that is, at the level of individual volume elements (voxels) of images, is crucial when evaluating the feasibility of voxel-based statistical image analysis methods such as SPM, where statistical testing is carried out independently for each voxel. Automated voxel-based image analysis methods have been utilized in several studies on  $^{11}\text{C}$ ]PIB accumulation (see, for example, references [2] and [5]), but the test-retest reproducibility of these analysis methods at the voxel level have not been studied previously. Automated ROI analysis has certain advantages over traditional manual ROI delineation. It avoids investigator-induced bias and also reduces workload especially in studies with large numbers of subjects.

$^{11}\text{C}$ ]PIB PET is currently used in follow-up studies of AD progression and in therapeutic trials aiming to reduce brain amyloid load or accumulation. For such studies with repeated  $^{11}\text{C}$ ]PIB measurements, information on the reproducibility of the analysis and quantitation method is of great importance. Furthermore, during disease progression, AD patients often become restless and it may be difficult for them to lie still in the PET scanner. Therefore, a short and simple imaging protocol would increase feasibility and subject comfort and compliance, and also lead to better quality of the imaging data. Recent studies indicate

that simplified quantitation using region-to-cerebellum ratios of late-scan standardized uptake values (SUVs) is valid for  $^{11}\text{C}$ ]PIB PET [7, 8], and reports have also shown that the region-to-cerebellum ratio is a useful approach in clinical group comparison studies [1–3, 9].

The aim of the present study was to evaluate the reproducibility and validity of fully automated image analysis techniques, including voxel-based quantitation of  $^{11}\text{C}$ ]PIB uptake, fitting of imaging data in a standard brain space and automated ROI analysis in a test-retest setting with repeated  $^{11}\text{C}$ ]PIB PET scans. Traditional manual ROI analysis was used as a comparison method. To evaluate the potential use of short and cost-effective imaging protocols, which would be convenient for AD patients, the accumulation of  $^{11}\text{C}$ ]PIB was quantitated as region-to-cerebellum ratios during only 30 min of data acquisition (60 to 90 min after tracer injection). In addition, based on the estimated variability in the quantitation of  $^{11}\text{C}$ ]PIB uptake, power calculations were performed to estimate the number of subjects required to reveal statistically significant differences in  $^{11}\text{C}$ ]PIB uptake changes in a parallel-group design with repeated measurements.

## Materials and methods

### Subjects

The study was approved by the Ethics Committee of Southwest Finland Hospital District. Six patients with clinically diagnosed AD and four healthy control subjects were recruited for repeated PET brain scans with  $^{11}\text{C}$ ]PIB. An interview and a physical examination were performed in order to exclude other conditions such as neurological or psychiatric disorders. Structural MRI of the brain was also performed to find possible brain lesions and to aid in the image analysis. All subjects and their carers, when applicable, gave written informed consent prior to the study. An intravenous line was inserted into each subject's forearm for the  $^{11}\text{C}$ ]PIB injection on both PET scanning occasions.

Four of the six AD patients and three of the four control subjects were female. At the time of the first scan, the average age of the AD patients was 71 years (range 59–79 years, SD 7.6), and that of the healthy controls was 64 years (range 51–71 years, SD 9.2). The average Mini-Mental State Examination (MMSE) score of the AD patients was 24.8 (range 23–27, SD 1.8), and that of the healthy controls was 28.3 (range 28–29, SD 0.5). All AD patients were on cholinesterase inhibitor treatment throughout the study (two on galantamine and four on rivastigmine). AD had been diagnosed 1.5–4 years (mean 2.6 years, SD 1.0) prior to the study.

## Radiochemistry

[<sup>11</sup>C]PIB was produced by the reaction of 6-OH-BTA-0 and <sup>11</sup>C-methyl triflate as described previously [2]. The mean  $\pm$ SD specific radioactivity of [<sup>11</sup>C]PIB at the time of administration was 31.9 $\pm$ 4.7 MBq/nmol for the AD patients, and 31.3 $\pm$ 5.0 MBq/nmol for the healthy controls. The radiochemical purity of the tracer was more than 95% in all PET studies, with an average of 97%.

## Scanning protocol

The imaging protocol has been described in detail previously [2]. In brief, all subjects underwent an established 90-min dynamic [<sup>11</sup>C]PIB PET scan with a GE Advance PET scanner (General Electric Medical Systems, Milwaukee, WI) in the 3-D scanning mode (septa retracted), providing 35 slices of 4.25 mm thickness covering the whole brain. The full-length scan was performed so that no data that could be of interest in the future would be discarded, although the plan was to use only the last 30 min of scanning data for the image analyses. Before injection of the radiotracer, an 8-min transmission scan with <sup>68</sup>Ge rod sources was performed for attenuation correction. [<sup>11</sup>C]PIB was injected into an antecubital vein and flushed with saline. The mean injected activities for the AD patients were 322 MBq (range 295–345 MBq, SD 19.2) for the first scan and 301 MBq (range 265–331 MBq, SD 23.4) for the second scan. For the healthy controls, the mean injected activities were 301 MBq (range 289–311 MBq, SD 10.1) and 275 MBq (range 236–309 MBq, SD 30.8), respectively.

The second PET brain scan was performed within 102 days (mean 34 days, SD 30, for the AD group; mean 58 days, SD 27, for the control group) of the first scan in each subject. Each 28-frame scan sequence consisted of four 30-s frames, nine 1-min frames, three 3-min frames, ten 5-min frames and two 10-min frames, but as planned, only the last 30 min (60–90 min) were used in the image analyses of this study. Although the transmission scan was performed before tracer injection, it was considered to anatomically correspond well with the late-scan data. This assumption was reasonable because head movement was controlled during scanning by monitoring landmarks on the subject's face. No blood sampling was performed during the scan. Imaging data were reconstructed into a 128 $\times$ 128 matrix using a transaxial Hann filter with a 4.6-mm cut-off and an axial ramp filter with an 8.5-mm cut-off.

Brain MRI scanning was performed with a Philips Intera 1.5-T scanner (Philips, Best, The Netherlands). T1-weighted 3-D scans with a voxel size of 0.50 $\times$ 0.50 $\times$ 1.00 mm were obtained.

## Image preprocessing

The preprocessing of the [<sup>11</sup>C]PIB PET imaging data was performed using the Statistical Parametric Mapping [10] software version 2 (SPM2) and Matlab 6.5 for Windows (Mathworks, Natick, MA). For the spatial normalization of the imaging data, a ligand-specific template for [<sup>11</sup>C]PIB was generated from 14 [<sup>11</sup>C]PIB scans from another study sample using MRI-aided procedures [11] as described in more detail previously [2], using SPM99. The template was created using [<sup>11</sup>C]PIB scans from age-matched healthy control subjects ( $n=5$ ), patients with mild cognitive impairment ( $n=4$ ), and patients with AD ( $n=5$ ) in an effort to eliminate group-related bias in spatial normalization. The spatial normalization would be most accurate when using group-specific templates. This would, however, lead to systematic differences in spatial normalization between groups and would invalidate group comparisons. Furthermore, we were interested in evaluating the use and reproducibility of a simple, fast and cost-effective image analysis procedure.

The template was created with the following steps. First, individual MR images were coregistered to summated (60–90 min) [<sup>11</sup>C]PIB images and then the MR images were normalized using a T1-weighted MRI template delivered with SPM to obtain normalization parameters. Using the acquired parameters, [<sup>11</sup>C]PIB summated images were normalized and written onto the template bounding box. To ensure the symmetry of the [<sup>11</sup>C]PIB template, the mean image of these normalized images was averaged with its mirror copy (left–right flipped image). Finally, this symmetrical [<sup>11</sup>C]PIB template was smoothed using an 8-mm gaussian kernel, as 8-mm smoothing is used for the estimation of the normalization parameters in SPM.

## Quantitation of [<sup>11</sup>C]PIB uptake

The uptake of [<sup>11</sup>C]PIB was quantitated at the voxel level using the region-to-cerebellum ratio which is identical to the SUV ratio (SUVR; [7]). This simplified quantitation enables the utilization of a short 30-min image acquisition, which has been shown to be valid for [<sup>11</sup>C]PIB [7]. The calculation of parametric images was performed using a fully automated procedure conducted in a common stereotactic space with the following phases. At first, summated (60 to 90 min) [<sup>11</sup>C]PIB images were normalized using a [<sup>11</sup>C]PIB template and written using bilinear interpolation. To calculate parametric images, the radioactivity concentration over 60 to 90 min in the cerebellar cortex was calculated from these normalized images using automated ROI analysis as described below. Finally, parametric images representing [<sup>11</sup>C]PIB region-to-cerebellum ratios in each voxel were calculated by dividing normalized summated images by the cerebellar radioactivity concentration values.

## Automated ROI analysis

Standardized ROIs were defined on the MRI template image representing brain anatomy in accordance with the MNI space (Montreal Neurological Institute database) utilized in SPM. To ensure a common stereotactic space, the mean image of 12 spatially normalized MR images was used instead of a single MR image. Separate gray and white matter templates were not considered necessary for the simplified methodology used in this study. The ROIs were delineated using Imadeus software (version 1.50; Forima, Turku, Finland) on the anterior and posterior cingulate cortices, lateral prefrontal and lateral temporal cortices, mesial temporal lobe (including amygdala, hippocampus and parahippocampal gyrus), occipital and inferior parietal cortices, caudate nucleus and putamen in the MR image. The ROIs are shown in Fig. 1. The left and right sides of each brain region were averaged. For reference purposes a ROI was drawn on the cerebellar cortex (to avoid contamination by white matter). The average regional region-to-cerebellum ratios were calculated using these ROIs on spatially normalized parametric images. In order to avoid quantitative bias (underestimation) in the estimation of ratio values, automated ROI analysis was performed without any smoothing of the images. Finally, the regional values were subjected to evaluation of reproducibility as described below (Fig. 2).

## Manual ROI analysis

Traditional manual ROI analysis was performed for purposes of comparison. [ $^{11}\text{C}$ ]PIB PET image pairs were realigned (within-subject) with SPM2, and the individual MR images were coregistered with the individual mean [ $^{11}\text{C}$ ]PIB images (mean of realigned images). ROIs were drawn on the same anatomical regions as listed in the section [Automated ROI analysis](#). Imadeus software was utilized for manual delineation of ROIs on the coregistered MR images and for calculating the radioactivity concentrations from realigned [ $^{11}\text{C}$ ]PIB PET images. The regional uptake of [ $^{11}\text{C}$ ]PIB was quantitated as region-to-cerebellum ratios during the 60–90 min interval by dividing regional radioactivity concentration values by the cerebellar values.

## Evaluation of reproducibility

To assess the reproducibility of [ $^{11}\text{C}$ ]PIB uptake measurements, absolute variability (VAR) values were calculated as follows:

$$VAR = \frac{|\text{scan2} - \text{scan1}|}{0.5(\text{scan1} + \text{scan2})} \times 100\%$$

Where scan1 and scan2 refer to parameter estimates of interest (region-to-cerebellum ratios) in the test and retest scans, respectively. In addition, test-retest reproducibility was evaluated in terms of intraclass correlation coefficients (ICC):

$$ICC = \frac{BSMS - WSMS}{BSMS + (n - 1)WSMS}$$

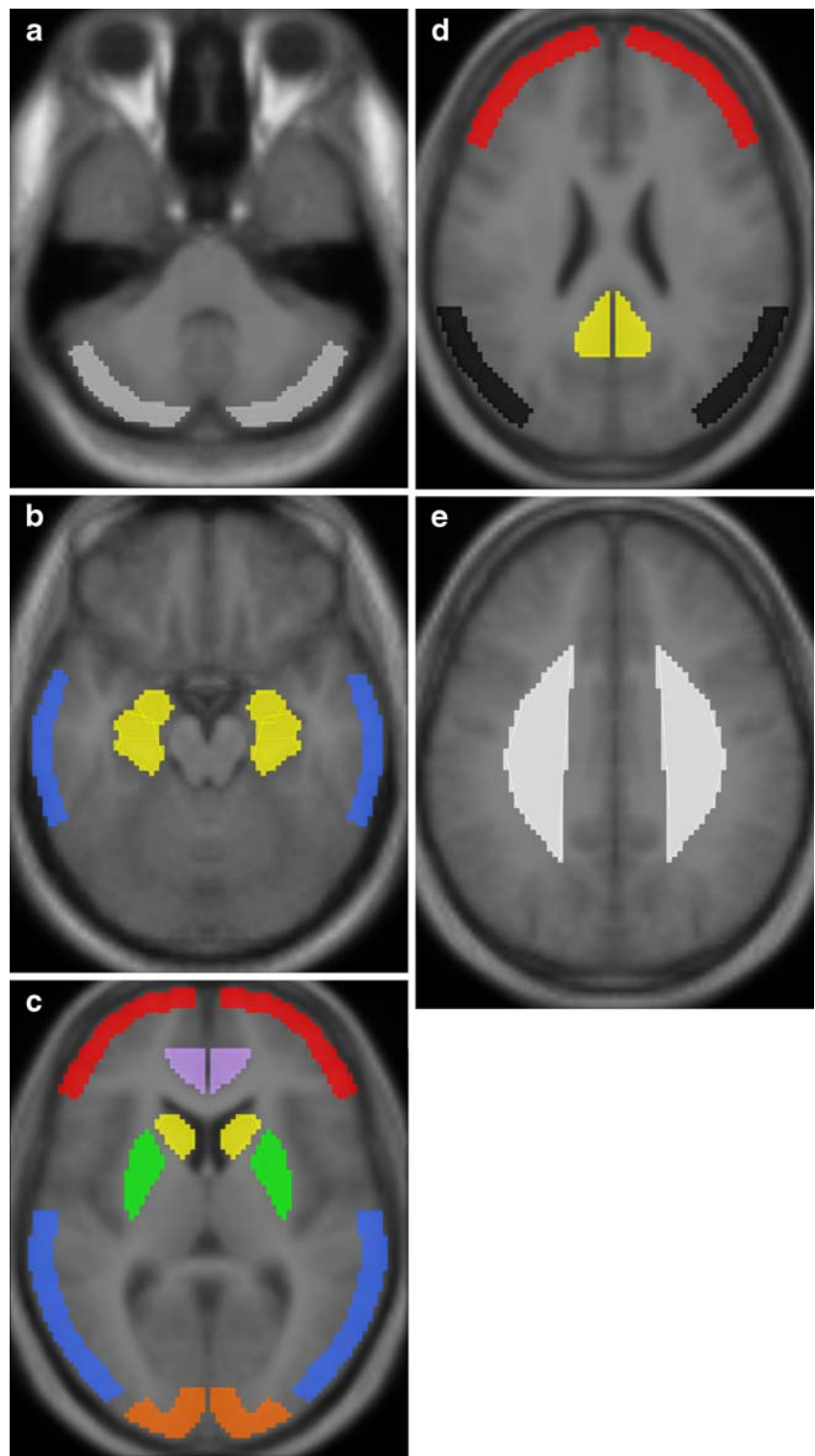
Where BSMS is the between-subject mean square, WSMS is the within-subject mean square, and  $n$  is the number of repeated observations (in this case,  $n=2$ ). The ICC can have values between  $-1$  and  $1$ ; values close to  $1$  indicate that most variance is due to between-subject rather than within-subject variation, whereas values below zero imply greater within-subject than between-subject variation.

For the evaluation of reproducibility at the voxel level, VAR and ICC values were calculated using the above formulas for each voxel in the MNI space separately from the region-to-cerebellum ratio images that were prepared for SPM analysis, that is, spatially normalized and smoothed using a 12-mm gaussian kernel. The calculations produced 3-D maps representing the spatial distribution of ICC and VAR values at the voxel level in the MNI brain space. The average values of ICC and VAR at the voxel level in each brain region were calculated using the ROIs included in the automated ROI analysis. In addition, to evaluate the reproducibility of the automated ROI analysis, region-level ICC and VAR values were calculated directly from the nonsmoothed spatially normalized parametric images. Smoothing was omitted to avoid underestimation in quantification (see [Automated ROI analysis](#)). To achieve more detailed results concerning the reproducibility in the two groups of subjects, reproducibility at the voxel level and region-level was evaluated separately for the AD patients and the controls.

Student's  $t$ -test for paired samples was used to test for differences in the parameters of interest between scan 1 and scan 2.  $P < 0.05$  was considered statistically significant. The analyses were performed using SPSS for Windows (release 13.0.1, SPSS, Chicago, IL).

To estimate sample sizes required to reveal statistically significant differences in [ $^{11}\text{C}$ ]PIB uptake in a parallel-group design with two measurements (for instance comparison of the effect of active treatment vs placebo in [ $^{11}\text{C}$ ]PIB uptake in AD patients), power calculations were performed using nQuery Advisor v7.0. (Statistical Solutions, Cork, Republic of Ireland). It was assumed according to the values of the present study, that the average baseline level of [ $^{11}\text{C}$ ]PIB uptake ratio would be 1.505 units for the lateral frontal cortex and that the standard deviation would be 0.08 units. The sample sizes were evaluated for differences between two parallel groups in the change from the

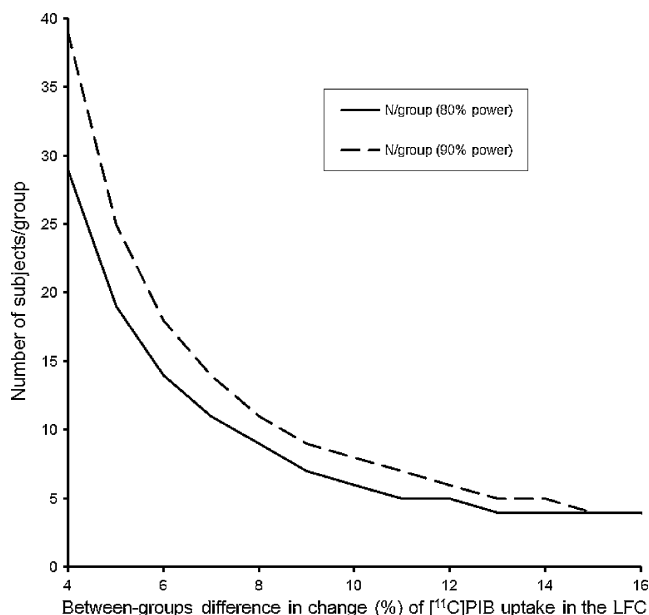
**Fig. 1** ROIs used in this study: **a** cerebellar cortex; **b** lateral temporal cortex and mesial temporal cortex; **c** lateral frontal cortex, caudate nucleus, putamen, lateral temporal cortex and occipital cortex; **d** posterior cingulate and parietal cortex; **e** white matter



baseline value to the post-intervention value (e.g. a change of 15% equal to 0.226 units in one group compared with a change of 0% in another group). The sample size calculations were made for 80% and 90% power and they were evaluated using a *t*-distribution with a two-sided significance level of 0.05.

## Results

The regional estimates of [ $^{11}\text{C}$ ]PIB uptake and the reproducibility estimates of the three analysis methods employed are given in Tables 1 and 2. Regional estimates of [ $^{11}\text{C}$ ]PIB accumulation calculated using automated ROI



**Fig. 2** Number of subjects needed per study group in relation to the between-group difference in the within-group relative change of [11C]PIB uptake in the lateral frontal cortex. The curves are for 90% power and for 80% power

analysis were generally in good agreement with values obtained using manual ROI analysis.

In general, the reproducibility of [11C]PIB uptake measurements was excellent in both groups, independent

of the analysis method. In the AD group, the reproducibility of automated ROI analysis (average VAR 4.3%, range 3.3–5.1%) was slightly better than that of manual ROI analysis (average VAR 6.6%, range 5.5–9.9%) and reproducibility at the voxel level was intermediate (average VAR 5.3%, range 4.9–6.2%). In the control group, the differences between methods were even smaller. The reproducibility of automated ROI analysis (average VAR 3.5%, range 0.9–5.5%) was practically the same as that of manual ROI analysis (average VAR 3.2%, range 1.9–4.8%) and reproducibility at the voxel level was slightly inferior (average VAR 5.3%, range 4.2–6.4%). There were no statistically significant differences between the VAR estimates obtained by the different analysis methods. The ICC of the automated ROI and voxel-based analyses varied between 0.69 and 0.97 in the AD group and 0.52 and 0.99 in the control group. The ICC values of the three different analysis methods in individual brain regions are given in Tables 1 and 2.

Power calculations indicated that 90% power can be achieved using a sample size of five subjects per group when a 15% change from baseline (increase or decrease) in [11C]PIB accumulation in the frontal cortex is anticipated in one group compared to no change in another group. A graphical illustration of the relationship between sample size and difference in the within-group relative change of [11C]PIB accumulation is presented in Fig. 2.

**Table 1** Regional estimates of [11C]PIB uptake (mean±SD) and VAR and ICC values in the AD patients (n=6)

	Caudate nucleus	Lateral frontal cortex	Occipital cortex	Posterior cingulate	Putamen	White matter	Lateral temporal cortex	Mesial temporal cortex	Striatum	Parietal cortex
<b>Auto-ROI</b>										
Mean scans 1	1.49±0.29	1.69±0.25	1.39±0.16	1.89±0.23	1.60±0.20	1.78±0.24	1.61±0.26	1.21±0.10	1.55±0.21	1.68±0.23
Mean scans 2	1.56±0.28	1.73±0.21	1.44±0.15	1.97±0.23	1.64±0.17	1.86±0.29	1.65±0.23	1.26±0.12	1.60±0.18	1.73±0.19
Mean all scans	1.52	1.71	1.42	1.93	1.62	1.82	1.63	1.23	1.58	1.70
VAR (%)	4.9	3.9	3.7	4.9	5.1	4.5	3.3	4.0	4.6	3.7
Mean difference (%)	4.8	2.8	3.8	4.5	2.6	4.2	2.8	4.1	3.4	3.6
ICC	0.96	0.95	0.94	0.88	0.89	0.90	0.97	0.85	0.93	0.94
<b>Manual ROI</b>										
Mean scans 1	1.53±0.35	1.65±0.21	1.42±0.11	1.73±0.21	1.56±0.24	1.77±0.21	1.63±0.23	1.21±0.13	1.54±0.26	1.68±0.14
Mean scans 2	1.63±0.45	1.68±0.24	1.48±0.13	1.81±0.15	1.58±0.30	1.84±0.33	1.70±0.24	1.28±0.09	1.60±0.35	1.73±0.18
Mean all scans	1.58	1.66	1.45	1.77	1.57	1.81	1.67	1.25	1.57	1.70
VAR (%)	6.0	5.7	5.5	8.4	6.0	7.0	5.9	9.9	5.7	5.7
Mean difference (%)	5.8	2.1	4.5	4.9	1.1	3.6	4.1	6.4	3.0	3.0
ICC	0.93	0.88	0.77	0.60	0.90	0.86	0.89	0.32	0.94	0.80
<b>Voxel-level</b>										
VAR (%)	5.7	5.0	5.1	5.4	5.4	5.4	4.9	6.2	5.6	4.7
ICC	0.91	0.90	0.81	0.83	0.77	0.88	0.90	0.69	0.82	0.89

**Table 2** Regional estimates of [<sup>11</sup>C]PIB uptake (mean±SD) and VAR and ICC values in the control group (n=4)

	Caudate nucleus	Lateral frontal cortex	Occipital cortex	Posterior cingulate	Putamen	White matter	Lateral temporal cortex	Mesial temporal cortex	Striatum	Parietal cortex
<b>Auto-ROI</b>										
Mean scans 1	1.40±0.17	1.33±0.09	1.39±0.09	1.34±0.08	1.40±0.12	1.84±0.21	1.29±0.08	1.22±0.06	1.40±0.11	1.25±0.05
Mean scans 2	1.37±0.06	1.33±0.08	1.39±0.11	1.34±0.07	1.44±0.16	1.89±0.16	1.29±0.08	1.24±0.08	1.41±0.11	1.25±0.04
Mean all scans	1.39	1.33	1.39	1.34	1.42	1.87	1.29	1.23	1.41	1.25
VAR (%)	5.3	2.7	3.2	0.9	4.1	5.5	2.5	4.0	4.6	2.0
Mean difference (%)	-1.5	0.3	-0.3	-0.1	2.5	3.2	0.7	2.3	1.1	0.5
ICC	0.56	0.88	0.89	0.99	0.89	0.79	0.91	0.66	0.76	0.81
<b>Manual ROI</b>										
Mean scans 1	1.28±0.18	1.19±0.11	1.29±0.10	1.16±0.05	1.25±0.09	1.70±0.23	1.19±0.08	1.12±0.12	1.26±0.11	1.18±0.08
Mean scans 2	1.26±0.09	1.15±0.08	1.27±0.06	1.15±0.06	1.22±0.07	1.72±0.13	1.17±0.04	1.11±0.06	1.24±0.07	1.17±0.07
Mean all scans	1.27	1.17	1.28	1.15	1.24	1.71	1.18	1.11	1.25	1.18
VAR (%)	4.4	3.0	2.9	3.0	2.8	4.8	2.3	4.6	2.4	1.9
Mean difference (%)	-0.3	-2.9	-1.2	-0.4	-2.1	1.6	-1.3	-0.5	-1.6	-0.5
ICC	0.84	0.80	0.86	0.71	0.89	0.87	0.71	0.78	0.86	0.92
<b>Voxel-level</b>										
VAR (%)	6.4	4.8	4.8	4.2	6.3	6.0	4.9	5.5	6.3	4.3
ICC	0.52	0.73	0.75	0.75	0.66	0.70	0.69	0.65	0.61	0.59

## Discussion

Voxel-based image analysis techniques are frequently used because they enable automated and objective image analysis and provide illustrative visualization of the results (see, for example, references [2] and [5]). Unlike conventional manual ROI analysis, voxel-based methods such as SPM can cover the whole brain space, making it very useful in explorative studies without any a priori hypotheses concerning the location of the differences or treatment effects. As manual ROI analysis is laborious and liable to operator-induced errors and bias, there is a need for automated and more objective quantitative methods. Recent studies have indicated that automated ROI analysis of spatially normalized parametric images is feasible for [<sup>11</sup>C]PIB [2, 5, 12] as well as for several other PET tracers (see, for example, references [13–15]).

Since voxel-based methods are based on image analysis conducted in a common stereotactic brain space, they require spatial normalization that fits individual images into the common space. Although the images are smoothed before statistical analysis, the accuracy of spatial normalization is crucial as statistical testing is carried out at the level of individual voxels. It has been demonstrated that spatial normalization with a ligand-specific template can provide more accurate spatial normalization than MRI-aided spatial normalization [11],

probably because inaccuracies in a between-modality (PET-MRI) coregistration can be avoided. However, it is open to discussion whether corrupting effects of gray matter atrophy or the white matter binding of [<sup>11</sup>C]PIB would possibly be better excluded by using MRI-aided normalization in studies in patients with AD. In any case, the purpose of this study was to evaluate the reproducibility of [<sup>11</sup>C]PIB uptake applying the most simplified methods of analysis, and therefore in this study spatial normalization was performed using a ligand-specific [<sup>11</sup>C]PIB template. The main aim of this work was to assess whether voxel-level simplified quantitation of 30-min [<sup>11</sup>C]PIB PET scanning data and fully automated image analysis can provide reproducible and valid [<sup>11</sup>C]PIB uptake estimates.

The regional binding estimates derived from voxel-based quantitation and automated ROI analysis were highly consistent with the region-to-cerebellum ratios achieved by manual ROI analysis. In addition, the region-level reproducibility of automated ROI analysis appeared to be at least as good as the reproducibility of manual ROI analysis. The spatial normalization of image pairs (scans 1 and 2) was performed separately using a ligand-specific [<sup>11</sup>C]PIB template, that is, without PET-PET realignment which is required in the manual ROI analysis of repeated PET measurements. The good reproducibility of automated ROI analysis indicates that

individually performed spatial normalization is feasible and the accuracy is comparable with realignment between images acquired with an interval of several weeks. Altogether, the good reproducibility and validity of the automated ROI estimation argue for the use of automated ROI analysis instead of conventional manual ROI analysis. The use of automated analysis avoids investigator-induced bias and also reduces workload.

To evaluate reproducibility at the voxel level, VAR and ICC maps were calculated from parametric [ $^{11}\text{C}$ ]PIB images representing the spatial distribution of region-to-cerebellum ratio values. The small VAR values indicated that reproducibility at the voxel level is generally equal to that obtained by manual ROI analysis. The ICC values were fairly low in the control group (Table 2). However, this does not mean that within-subject variability would be greater in the controls, but rather that when there is practically no specific tracer uptake, between-subject variation is *also* very low, which causes the ICC value to decline. It has to be emphasized, therefore, that the VAR values should be considered the primary measure of reproducibility, especially in the control group. The good to excellent reproducibility inferred from the low VAR values indicates that voxel-based analysis with SPM is sensitive even in fairly small samples. Indeed, it has already been shown that voxel-based analysis is feasible for demonstrating increased [ $^{11}\text{C}$ ]PIB uptake in AD patients as compared to controls [2, 5, 12].

In our study, using manual ROI delineation, the VAR values varied from 1.9% to 9.9% between brain areas (mean 6.6%, range 5.5–9.9%, in AD patients; mean 3.2%, range 1.9–4.8%, in controls). Two previous studies have evaluated the reproducibility of manual ROI analysis [7, 16]. In a comprehensive study that compared various simplified [ $^{11}\text{C}$ ]PIB uptake quantitation methods it was found that the test-retest variability of the region-to-cerebellum ratio was among the smallest of the examined analysis methods. The ratios were determined using two time intervals, 40–60 min and 40–90 min after tracer injection, and the average variability values of manual ROI analysis were 5.3% and 5.0%, respectively [7], comparable with the variability estimates of the present study. In another study with a 2-year follow-up of AD patients [16], four AD patients were rescanned with [ $^{11}\text{C}$ ]PIB PET to evaluate the reproducibility of the region-to-cerebellum ratio obtained at 40 to 60 min after injection. The results were calculated with an equation identical to the one used in the present study. Mean VAR values in the cortical areas examined were between 3.2% and 7.3%, and slightly greater in the striatum (12.7%).

In a recent study that compared different parametric reference tissue models and their reproducibility in the quantitation of [ $^{11}\text{C}$ ]PIB uptake, the test-retest variability of the region-to-cerebellum ratios (5.2% in controls and 7.9%

in AD patients) was greater than those obtained with other parametric methods [6]. In the present study, however, we were able to evaluate and report the good reproducibility of region-to-cerebellum ratios not only at the regional level but also at the voxel level, and the reproducibilities were of the same magnitude as was seen in the previous study at the regional level for simplified or multilinear reference tissue models [6]. Undoubtedly some parametric methods not based on the region-to-cerebellum ratio may be more accurate in the quantitation of [ $^{11}\text{C}$ ]PIB uptake. However, the previously evaluated reference-input Logan method [6] and the promising recently evaluated masked-volume-wise PCA method without any modelling assumptions [17], for example, require data to be collected from the time of injection to the end of the scan, whereas a scan planned for region-to-cerebellum ratio uptake analysis can be conducted with (a) no blood sampling and (b) even with a shorter scanning protocol. These differences in the required amount of input data underline the practical value of the methods evaluated in this study.

Disease modifying therapies in AD are under clinical investigation. To verify disease-modifying effects using clinical measures (neuropsychological testing, global impression of change, assessment of behavioural symptoms) requires long follow-up and large numbers of patients. Therefore surrogate markers of treatment efficacy would be desirable. The good reproducibility presented here can be interpreted as indicating that [ $^{11}\text{C}$ ]PIB PET could potentially be reliably used in study designs with repeated measurements, such as follow-up studies or therapeutic trials, as one method to monitor therapeutic efficacy. The post hoc power calculations performed by us provide preliminary information on sample sizes required in a parallel-group design. For example, for the comparison of two groups with a treatment response resulting in a 15% change in [ $^{11}\text{C}$ ]PIB uptake in the frontal cortex in one group and no change in another group, samples of only five subjects per group would seem to be sufficient (90% power). This small number of required subjects highlights the relevance of the adequate reproducibility of automated image analysis of [ $^{11}\text{C}$ ]PIB PET data.

The present findings show that good reproducibility can be obtained using only 30 min of [ $^{11}\text{C}$ ]PIB PET data. We emphasize that only simplified quantitation, which was used in the present study, is possible when using only 30 min of PET imaging data. A short imaging protocol would also have important practical implications since it enhances cost-effectiveness of PET scanner and tracer use, as the tracer may be injected and scanning would take place from 60 to 90 min after tracer injection, enabling the division of the same tracer batch among several patients. In addition, a short scan time will improve patient comfort and



increase compliance, possibly contributing to better data quality from subjects with more advanced AD. This is important in follow-up studies, for example, where patients become more demented as time goes by. Advanced dementia is accompanied by increasing restlessness and agitation which may make it difficult for the patient to lie still in the scanner for a long period of time. Systematically poorer scanning data at the second imaging time point could lead to bias in follow-up studies.

The choice of image acquisition timing and other methodological issues should still be weighed separately for each study and clinical practice. Ideally, of course, at each PET imaging site using [ $^{11}\text{C}$ ]PIB a test-retest study should be done to evaluate the reproducibility being achieved at that site with the instrumentation at hand, and investigators should also estimate acquired sample sizes separately for each study design. The timing of the transmission scan should also be optimized in future when scanning 60 to 90 min after tracer injection. We note that we performed a dynamic 90-min scan preceded by a preinjection transmission scan and used the last 30 min data for analysis. Ideally one would perform the transmission immediately before scanning session or after it. A comparison of pre- and postinjection transmission scan for brain FDG studies has shown that the results obtained using a preinjection transmission scan agree well with those obtained using a postinjection transmission scan despite the time gap between the preinjection transmission scan and the emission scan [18]. However, such a comparison has not been performed for [ $^{11}\text{C}$ ]PIB. In addition, disregarding the early frames may result in loss of valuable information in some cases, since early frames contain information about blood flow [19].

In conclusion, the results of the present study indicate that the reproducibility of [ $^{11}\text{C}$ ]PIB region-to-cerebellum ratios using simplified fully automated analysis is good to excellent, even when applying only 30 min of imaging data. The reproducibility and validity of the estimates are maintained when quantitation is performed at the voxel level and the analysis is conducted in a standard brain space. Our results suggest that with [ $^{11}\text{C}$ ]PIB PET data, automated voxel-based image analysis is a quick and cost-effective alternative to conventional ROI analysis.

**Acknowledgments** The assistance of the personnel of Turku PET Centre and CRST is gratefully acknowledged. This study was financially supported by grants from the Academy of Finland (project # 205954), the Sigrid Jusélius Foundation and Turku University Hospital (EVO).

**Conflicts of Interest** Marita Kailajärvi works currently for Turku Imanet, GE Healthcare. The other authors have no conflicts of interest.

## References

1. Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Ann Neurol* 2004;55:306–19. doi:10.1002/ana.20009.
2. Kemppainen NM, Aalto S, Wilson IA, Nägren K, Helin S, Brück A, et al. Voxel-based analysis of PET amyloid ligand [ $^{11}\text{C}$ ]PIB uptake in Alzheimer disease. *Neurology* 2006;67:1575–80. doi:10.1212/01.wnl.0000240117.55680.0a.
3. Edison P, Archer HA, Hinz R, Hammers A, Pavese N, Tai YF, et al. Amyloid, hypometabolism, and cognition in Alzheimer disease: an [ $^{11}\text{C}$ ]PIB and [ $^{18}\text{F}$ ]FDG PET study. *Neurology* 2007;68:501–8. doi:10.1212/01.wnl.0000244749.20056.d4.
4. 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. doi:10.1212/01.wnl.0000261919.22630.ea.
5. Mikhno A, Devanand D, Pelton G, Cuasay K, Gunn R, Upton N, et al. Voxel-based analysis of  $^{11}\text{C}$ -PIB scans for diagnosing Alzheimer's disease. *J Nucl Med* 2008;49:1262–9. doi:10.2967/jnumed.107.049932.
6. Yaqub M, Tolboom N, Boellaard R, Berckel BNM, van Tilburg EW, Luurtsema G, et al. Simplified parametric methods for [ $^{11}\text{C}$ ]PIB studies. *Neuroimage* 2008;42:76–86. doi:10.1016/j.neuroimage.2008.04.251.
7. Lopresti BJ, Klunk WE, Mathis CA, Hoge JA, Ziolkko SK, Lu X, et al. Simplified quantification of Pittsburgh Compound B amyloid imaging PET studies: a comparative analysis. *J Nucl Med* 2005;46:1959–72.
8. Price JC, Klunk WE, Lopresti BJ, Lu X, Hoge JA, Ziolkko SK, et al. Kinetic modeling of amyloid binding in humans using PET imaging and Pittsburgh Compound-B. *J Cereb Blood Flow Metab* 2005;25:1528–47. doi:10.1038/sj.jcbfm.9600146.
9. Kemppainen NM, Aalto S, Wilson IA, Nägren K, Helin S, Brück A, et al. PET amyloid ligand [ $^{11}\text{C}$ ]PIB uptake is increased in mild cognitive impairment. *Neurology* 2007;68:1603–6. doi:10.1212/01.wnl.0000260969.94695.56.
10. Friston KJ, Worsley KJ, Poline JB, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional brain imaging. A general approach. *Hum Brain Mapp* 1995;2:189–210. doi:10.1002/hbm.460020402.
11. Meyer JH, Gunn RN, Myers R, Grasby PM. Assessment of spatial normalization of PET ligand images using ligand-specific templates. *Neuroimage* 1999;9:545–53. doi:10.1006/nimg.1999.0431.
12. Ziolkko SK, Weissfield LA, Klunk WE, Mathis CA, Hoge JA, Lopresti BJ, et al. Evaluation of voxel-based methods for the statistical analysis of PIB PET amyloid imaging studies in Alzheimer's disease. *Neuroimage* 2006;33:94–102. doi:10.1016/j.neuroimage.2006.05.063.
13. Brück A, Aalto S, Nurmi E, Bergman J, Rinne JO. Cortical 6-[ $^{18}\text{F}$ ]fluoro-L-dopa uptake and frontal cognitive functions in early Parkinson's disease. *Neurobiol Aging* 2005;26:891–8. doi:10.1016/j.neurobiolaging.2004.07.014.
14. Kemppainen J, Aalto S, Fujimoto T, Kallioikoski KK, Långsjö J, Oikonen V, et al. High intensity exercise decreases global brain glucose uptake in humans. *J Physiol* 2005;568:323–32. doi:10.1113/jphysiol.2005.091355.
15. Salmi E, Aalto S, Hirvonen J, Långsjö JW, Maksimov AT, Oikonen V, et al. Measurement of GABAA receptor binding in vivo with [ $^{11}\text{C}$ ]flumazenil: a test-retest study in healthy subjects. *Neuroimage* 2008;41:260–9. doi:10.1016/j.neuroimage.2008.02.035.
16. Engler H, Forsberg A, Almkvist O, Blomqvist G, Larsson E, Savitcheva I, et al. Two-year follow-up of amyloid deposition in patients with Alzheimer's disease. *Brain* 2006;129:2856–66. doi:10.1093/brain/awl178.

17. Razifar P, Ringheim A, Engler H, Hall H, Långström B. Masked-volume-wise PCA and "reference Logan" illustrate similar regional differences in kinetic behavior in human brain PET study using [11C]-PIB. *BMC Neurol* 2009;9:2. doi:[10.1186/1471-2377-9-2](https://doi.org/10.1186/1471-2377-9-2).
18. Hooper PK, Meikle SR, Eberl S, Fulham MJ. Validation of postinjection transmission measurements for attenuation correction in neurological FDG-PET studies. *J Nucl Med* 1996;37:128–36.
19. Blomquist G, Engler H, Nordberg A, Ringheim A, Wall A, Forsberg A, et al. Unidirectional influx and net accumulation of PIB. *Open Neuroimaging J* 2008;2:114–25. doi:[10.2174/187444000802010114](https://doi.org/10.2174/187444000802010114).