ORIGINAL ARTICLE

Test-retest variability of quantitative [11C]PIB studies in Alzheimer's disease

Nelleke Tolboom · Maqsood Yaqub · Ronald Boellaard · Gert Luurtsema ·
Albert D. Windhorst · Philip Scheltens · Adriaan A. Lammertsma ·
Bart N. M. van Berckel

Abstract

Purpose The aim of this study was to assess the test-retest variability of [11C]PIB studies in patients with Alzheimer's disease (AD) and healthy controls using several tracer kinetic models and to assess the suitability of the cerebellum as reference tissue.

Methods [11C]PIB studies with arterial sampling were performed in eight AD patients and eight healthy controls. Retest scans were performed in six controls and six AD patients. Data were analysed using plasma input and reference tissue models, together with simple ratios.

Results Test-retest variability was best (~3%) for SRTM2, a parametric method based on the simplified reference tissue model. Highest values (~10%) were found for plasma input models. Cerebellar V_T values did not differ significantly between AD and controls.

Conclusion Parametric SRTM2 with the cerebellum as reference tissue is the method of choice for quantitative analysis of [11C]PIB PET studies.

Keywords Pittsburgh compound B · PIB · Arterial sampling · Cerebellum · Parametric images · Amyloid · PET

N. Tolboom · M. Yaqub · R. Boellaard · G. Luurtsema ·
A. D. Windhorst · A. A. Lammertsma · B. N. M. van Berckel
Departments of Nuclear Medicine & PET Research,
VU University Medical Centre,
Amsterdam, The Netherlands

N. Tolboom (✉) · P. Scheltens
Department of Neurology & Alzheimer Centre,
VU University Medical Centre,
P.O. Box 7057, 1007 Amsterdam, MB, The Netherlands
e-mail: n.tolboom@vumc.nl

Introduction

Amyloid beta (Aβ) is the main constituent of amyloid plaques in the brain of Alzheimer's disease (AD) patients. It is thought to be a key component of AD pathology [1]. With the development of new treatments for AD, aimed at either preventing Aβ deposition or increasing its clearance, the role of quantification of Aβ burden in vivo becomes crucial [2, 3].

Over the past two decades several positron emission tomography (PET) tracers have been developed for imaging and quantifying fibrillar Aβ accumulation in AD in vivo. Of these, *N*-methyl-[11C]2-(4'methylaminophenyl)-6-hydroxybenzothiazole ([11C]PIB) [4] has been used most widely. Several tracer kinetic models for quantification of [11C]PIB binding have been reported [5, 6], although Logan graphical analysis [7] and standardized uptake value ratios (SUV$_r$) have been used most frequently [5].

The reliability of [11C]PIB tracer kinetic models can be assessed using test-retest studies. Good test-retest variability is particularly important when evaluating the natural time course of Aβ depositions and for monitoring therapeutic effects of novel drugs designed to reduce Aβ accumulation in the brain. Data on test-retest variability of [11C]PIB tracer kinetic models are confined to three studies, mostly dedicated to simplified approaches [5, 6, 8]. In addition, most clinical studies published to date have used simplified analysis approaches with cerebellum grey matter as reference region. Whilst the cerebellum is reported to be free of Congo red and thioflavin S-positive plaques, diffuse amyloid deposits are commonly observed [9, 10], especially in AD. Although these diffuse amyloid deposits in the cerebellum were not detectable in a 6-CN-PiB staining study [11], more data on cerebellar [11C]PIB binding in vivo are necessary. This is especially important for those

analyses where the cerebellum is used as reference tissue, as binding of [11C]PIB to diffuse amyloid could lead to underestimation of regional specific binding.

Only limited information is available describing binding of [11C]PIB in the cerebellum in vivo. In two studies [5, 6] published to date, the volume of distribution of [11C]PIB in the cerebellum (V_T) was similar for AD patients and controls. It should be mentioned, however, that the first of these studies only had limited power.

The aim of the present study was twofold. The first aim was to further assess test-retest variability of [11C]PIB studies using several modelling approaches. The second aim was to investigate binding of [11C]PIB in the cerebellum in order to evaluate whether it can be used as reference tissue for [11C]PIB studies in AD.

Materials and methods

Participants

Eight AD patients [three women, five men, Mini-Mental State Examination (MMSE) 21±2, age 61±3, mean ± SD] and eight healthy controls (three women, five men, MMSE 30±1, age 66±6) were included in this study. All participants received a standard dementia screening that included medical history, physical and neurological examinations, screening laboratory tests, brain magnetic resonance imaging (MRI), and extensive neuropsychological testing. Final diagnosis was established by a multidisciplinary team, without knowledge of the PET results. All AD patients met NINCDS-ADRDA [12] criteria for "probable AD". Control subjects were recruited through advertisements in newspapers and underwent the same diagnostic procedures. Exclusion criteria for controls were subjective memory complaints, a history of major neurological or psychiatric illness, or clinically significant abnormalities on the MRI (as determined by a neuroradiologist). Written informed consent was obtained from all subjects after a complete written and verbal description of the study. The study was approved by the Medical Ethics Review Committee of the VU University Medical Centre.

Radiochemical synthesis

Radiolabelled [11C]PIB was synthesized according to a modification of the procedure described by Wilson et al. [13]. The starting material 2-(4'-aminophenyl)-6-hydroxybenzothiazole (4.1 μmol, 1.0 mg, purchased from ABX, Radeberg, Germany) was dissolved in 250 μl of acetone. To this solution [11C]methyl triflate [14] was added at room temperature. After reaction at 60°C for 1

min, the reaction mixture was diluted with 750 μl of a solution of trifluoroacetic acid (pH = 2) in water. This mixture was injected onto a Chromolith RP18e 100*25 mm HPLC column (Merck, Darmstadt, Germany) which was eluted with a 50/50 mixture of trifluoroacetic acid (pH = 1.5) in water/methanol. The product, [11C]PIB, eluted at 8 min. This fraction was collected in 60 ml of water for injection. The total solution was passed over a preconditioned (10 ml of 96% sterile ethanol and subsequently 10 ml of water for injection) tC$_{18}$ Sep-Pak (Waters, Milford, MA, USA). The Sep-Pak was washed with a mixture of 5 ml of sterile ethanol in 15 ml of water for injection to elute residual methanol from the HPLC eluent and to elute an unidentified impurity. Finally, the product was eluted from the Sep-Pak with 1.0 ml of sterile ethanol (96%) and diluted with 14.0 ml of a sterile and pyrogen-free solution of 7.09 mM NaH$_2$PO$_4$ in saline. This final mixture was passed over a Millex GV 0.22-μm filter (Millipore, Billerica, MA, USA), yielding a sterile, isotonic and pyrogen-free solution of 1,491–3,997 MBq of [11C]PIB with a (radio)chemical purity >98% and a specific activity (SA) of 40–240 GBq/μmol at the end of synthesis.

PET

PET scans were performed on an ECAT EXACT HR + scanner (Siemens/CTI, Knoxville, TN, USA), equipped with a neuro-insert to reduce the contribution from outside field of view (FOV) activity. This scanner enables the acquisition of 63 transaxial planes over a 15.5-cm axial FOV, thus allowing the whole brain to be imaged in a single bed position. The characteristics of this scanner have been described elsewhere [15]. [11C]PIB studies were performed in 16 subjects: 8 AD patients and 8 healthy controls. Of these subjects, six AD patients and six healthy controls underwent repeat [11C]PIB studies. All repeat studies, except for one, were performed on the same day. All subjects received an indwelling radial artery cannula for arterial blood sampling and a venous cannula for tracer injection. Patient motion was restricted by the use of a head holder and monitored by checking the position of the head using laser beams.

First, a 10-min transmission scan was performed in 2-D acquisition mode using three retractable rotating line sources. This scan was used to correct the subsequent emission scan for photon attenuation. Next, a dynamic emission scan in 3-D acquisition mode was started simultaneously with the intravenous injection of [11C]PIB using an infusion pump (Med-Rad, Beek, The Netherlands) in approximately 4 s followed by a flush of 40 ml saline at 2.0 ml/s. The average injected [11C]PIB dose for AD subjects was 348±58 MBq (SA = 52±22 GBq/μmol) and 394±47 MBq [11C]PIB (SA = 53±22 GBq/μmol) for the retest scans. For healthy controls, the average injected [11C]

PIB dose for test scans was 347±25 MBq (SA = 53± 30 GBq/μmol) and 369±48 (SA = 50±18 GBq/μmol) for the retest scans. Each dynamic emission scan consisted of 23 frames with progressive increase in frame length (1×15, 3×5, 3×10, 2×30, 3×60, 2×150, 2×300, 7×600 s) and a total duration of 90 min. Arterial blood was withdrawn continuously at a rate of 5 ml/min for the first 10 min and 2.5 ml/min thereafter until 60 min post-injection, using an online detection system [16]. In addition, at 5, 10, 20, 30, 40, 60, 75 and 90 min post-injection, arterial blood samples were collected manually. These were used to calibrate the online blood curve, to determine plasma to whole blood ratios of radioactivity and to measure plasma parent tracer and metabolite fractions.

Finally, after a resting period of at least 1 h, on average 4 h and 32 min±8 min after the first administration to allow for decay of [11C], exactly the same procedure was repeated.

MRI

All subjects underwent a structural MRI scan using a 1.5 T Sonata scanner (Siemens Medical Solutions, Erlangen, Germany). The scan protocol included a coronal T1-weighted 3-D MPRAGE (magnetization prepared rapid acquisition gradient echo; slice thickness 1.5 mm, 160 slices; matrix size 256× 256; voxel size 1×1×1.5 mm; echo time = 3.97 ms; repetition time = 2,700 ms; inversion time = 950 ms; flip angle 8°).

Metabolite analysis

Metabolite fractions were determined using online solid phase extraction (SPE) combined with HPLC and radioactivity detection [17]. A tC_{18} SPE online column was used in combination with a 5-μm Gemini HPLC 5-μm C18 column (250*10 mm, Phenomenex, Torrance, CA, USA). MeCN/ NH_4OAc 10 mM [50/50 (v/v %)] was used as mobile phase at a flow rate of 2.5 ml·min^{-1}. The total recovery of [11C] PIB from plasma using this method was >90% for all samples. There was no time dependency with respect to recovery of radioactivity. Retention times of 11C-labelled metabolites and [11C]PIB were 10 and 15 min, respectively.

Image analysis

All PET sinograms were corrected for dead time, tissue attenuation using the transmission scan, decay, scatter and randoms and were reconstructed using a standard filtered back projection algorithm and a Hanning filter with cut-off at 0.5 times the Nyquist frequency. A zoom factor of 2 and a matrix size of 256×256×63 were used, resulting in a voxel size of 1.2×1.2×2.4 mm and a spatial resolution of approximately 7 mm full-width at half-maximum at the centre of the FOV.

ROI definition

MR images were aligned to corresponding PET images using a mutual information algorithm [18]. Segmentation of MRI and further analysis of the PET data was performed using PVE-lab, a software programme that uses a probability map of 35 delineated (grey matter) regions of interest (ROI) that has been validated previously [19]. For the present study, grey matter ROIs of frontal (volume-weighted average of orbital frontal, medial inferior frontal and superior frontal), parietal and temporal (volume-weighted average of superior temporal and medial inferior temporal) cortex and medial temporal lobe (MTL) (volume-weighted average of enthorinal cortex and hippocampus), posterior cingulate and cerebellum were used. In addition, a "global cortical" ROI was defined, based on the volume-weighted average of all the aforementioned regions except the cerebellum. As the purpose of this study was to assess intra-subject variability, no correction for partial volume effects was applied.

Kinetic analyses

Kinetic analyses were performed using dedicated software developed within Matlab 5.3 (The Mathworks, Natick, MA, USA). The online blood curve was calibrated using the manual whole blood samples. To this end, the interval of the curve between first and last manual samples was fitted to a multi-exponential function, adding exponentials until both Akaike [20] and Schwarz [21] criteria indicated no further improvement. The (original) online sampler curve was then multiplied with the average of the ratios of (whole blood) sample values and corresponding interpolated values of the fitted exponential function. Next, the plasma to whole blood ratios of the manual samples were also fitted to a multi-exponential function, again using Akaike and Schwarz criteria to determine the number of exponentials. The (calibrated) whole blood curve was then multiplied with this multi-exponential function to generate the corresponding (total) plasma curve. Finally, the parent fraction was fitted to a Hill function and the total plasma curve was multiplied with this fitted function to obtain the metabolite-corrected plasma input function

Plasma input method

ROI data were analysed using a reversible two-tissue compartment (2T4k) plasma input model, including a parameter for blood volume fraction [22] and providing both binding potential (BP_{2T4k}=k3/k4, which can contain a non-specific component) and volume of distribution (V_T). In addition, the distribution volume ratio (DVR) with cerebellar grey matter was calculated [DVR = V_T(target)/

V_T(cerebellum)] [23]. Data were fitted using parameter boundary constraints. Furthermore, an optimized non-linear least-squares procedure was used [24]. According to this procedure, parameter boundaries were estimated using several runs, each time activity curve was fitted multiple times using different starting parameters and the fit was visually inspected afterwards. Note that, in order to distinguish BP (k_3/k_4) derived from plasma input and reference tissue models, BP_{2T4k} will be used for the 2T4k plasma input model, as it does not separate specific from non-displaceable binding.

Reference tissue methods

In addition to the plasma input method, data were also analysed using both reference tissue models and simple ratios. First, BP_{ND} was derived for ROI data using the simplified reference tissue method (SRTM) [25] with cerebellar grey matter as reference tissue. With SRTM, for k_2 a lower bound of 0.006 was used, but no upper bound was set. Next, parametric images of BP_{ND} were generated using a basis function implementation of the "two-step" SRTM [26]: parametric SRTM2 (also referred to as RPM2 in previous reports [27, 28]). In the first run, parametric SRTM [29] was used to obtain an estimate of the reference tissue efflux rate constant, k_2'. To reduce noise, in the second run (parametric SRTM2) k_2' was fixed to the median (parametric SRTM) value for all pixels with $BP_{ND}>0$. For both parametric SRTM and parametric SRTM2, 40 basis functions were used. $\theta_3^{min} = \frac{k_2^{min}}{(1+BP^{max})}$ and $\theta_3^{max} = k_2^{max}$ were set to 0.01 and 0.3 min^{-1}, respectively. This method was chosen, as it has been identified as the parametric method of choice [28].

Finally, data were analysed using SUV ratios (SUV$_r$), which is the target to grey matter cerebellum SUV ratio over the interval 60–90 min after injection.

For all methods cerebellar grey matter was chosen as reference tissue because of its (histopathological) lack of Congo red and thioflavin S-positive plaques [9, 10].

Evaluation of methods and statistics

The p values for group differences were obtained using Student's t test.

Mean parameter values, test-retest variability and percentage difference between test and retest of [11C]PIB studies in patients with AD and healthy controls was assessed for 2T4k derived DVR, 2T4k $BP_{2T4k}+1$ and V_T, and for SRTM derived $BP_{ND+}+1$, parametric SRTM2 derived $BP_{ND}+1$ and SUV$_r$. Test-retest variability was assessed using $BP_{ND}+1$, similar to Price et al. [5], which corresponds to DVR. This was done because BP_{ND} in healthy controls approaches zero; therefore, small devia-

tions in values will lead to large relative variability. Intra-subject test-retest variability was calculated as the absolute difference between retest (R) and test (T): test/retest (%) = $100 \times$ absolute$\{(R-T)/(0.5(R + T))\}$. Percentage difference between test and retest was calculated using: (%)=$100 \times \{(R-T)/(T)\}$. Regional variability was further assessed for V_T, parametric SRTM2 derived $BP_{ND}+1$ and SUV$_r$.

The relationship between 2T4k derived DVR and both SRTM and parametric SRTM2 derived $BP_{ND}+1$, and between SRTM derived $BP_{ND}+1$ and parametric SRTM2 derived $BP_{ND}+1$ was assessed using the square of the Pearson's product moment correlation coefficient (r^2).

Average parietal and cerebellar 2T4k derived BP_{2T4k} and V_T values were compared between AD patients and controls.

Results

Arterial input function

Eight AD patients and eight healthy controls were included of which six AD patients and six controls underwent repeat scans. AD patients and healthy controls were similar with respect to age (p value>0.05). All test and retest scans were performed on a single day, except for one control subject, who was scanned with an interval of 14 days due to technical difficulties (radiosynthesis failure). Although arterial blood sampling was performed in all subjects, due to technical reasons, plasma input data were not available for one control (retest) and two AD patients (both test and retest) scans. Metabolite-corrected plasma input data were acquired in six AD patients and eight controls. Complete test-retest plasma input data were available for four of six AD patients and five of six controls.

Test-retest data

Table 1 summarizes mean (±SD) parameter values of the various methods for cortical areas known to be associated with high (parietal cortex) and slightly less (sensory motor cortex) [11C]PIB uptake. For the parietal cortex, differences between AD and control values were assessed. Best group differentiation was obtained using parametric SRTM2 ($p<$ 0.01) and SUV$_r$ ($p=0.01$). Table 2 summarizes absolute variability. Using absolute values, 2T4k (plasma input) derived DVR and V_T showed variability around ~10%. Using 2T4k $BP_{2T4k}+1$ values in AD were ~10% with somewhat higher values in controls. SRTM derived $BP_{ND}+$ 1 showed low variability in AD patients, but higher values for controls (AD ~4.5%; controls ~12.5%). In contrast, SUV$_r$ showed lower values for controls (~5%) and slightly higher values for AD (7–9%). Parametric SRTM2 derived

Table 1 Mean (± SD) test and retest values of plasma input and reference tissue methods together with a simple ratio

ROI	2T4k (DVR)		2T4k ($BP_{2T4k}+1$)		2T4k (V_T)		SRTM ($BP_{ND}+1$)		Parametric SRTM2 ($BP_{ND}+1$)		SUV_r	
	Test	Retest	Test	Retest	Test	Retest	Test	Retest	Test	Retest	Test	Retest
AD parietal	1.76 (0.48)	1.87 (0.39)	3.43 (1.05)	3.69 (0.88)	6.67 (2.05)	6.93 (1.69)	1.67 (0.35)	1.71 (0.32)	1.75 (0.36)	1.80 (0.33)	1.93 (0.49)	2.04 (0.40)
AD sensory motor Cx	1.50 (0.44)	1.57 (0.33)	3.16 (1.01)	3.29 (0.78)	5.69 (1.88)	5.83 (1.49)	1.51 (0.30)	1.55 (0.35)	1.41 (0.25)	1.44 (0.24)	1.63 (0.41)	1.71 (0.32)
Control parietal	1.19[a] (0.10)	1.14 (0.11)	2.77[b] (0.96)	2.66 (1.12)	5.09[c] (2.16)	4.94 (2.60)	1.32[d] (0.22)	1.27 (0.23)	1.12[e] (0.22)	1.09 (0.21)	1.25[f] (0.18)	1.22 (0.18)
Control sensory motor Cx	1.17 (0.06)	1.13 (0.08)	2.75 (0.99)	2.65 (1.12)	4.98 (2.02)	4.89 (2.52)	1.39 (0.25)	1.33 (0.18)	1.10 (0.13)	1.07 (0.13)	1.24 (0.12)	1.21 (0.12)

Compared to AD parietal: [a] $p<0.05$, [b] $p=0.36$, [c] $p=0.30$, [d] $p<0.10$, [e] $p<0.01$, [f] $p=0.01$

In the case of 2T4k, $n=4$ and 5 for Alzheimer patients and healthy controls, respectively; in the case of SRTM, parametric SRTM2 and SUV_r $n=6$ for both patient groups

ROI region of interest, *AD* Alzheimer's disease, *Cx* cortex

$BP_{ND}+1$ had lowest variability for both controls and AD (2–4%). Table 3 summarizes percent differences. The smallest percent difference was seen using the reference tissue approaches SRTM derived $BP_{ND}+1$ and parametric SRTM2 derived $BP_{ND}+1$. There was a trend towards higher values for retest scans in the AD patients for all methods used, but this was not significant. For controls the opposite was seen, but again this was not significant.

Next, regional variabilities of a plasma input method (V_T), a reference tissue method (parametric SRTM2) and a simple ratio method (SUV_r) were assessed. Figure 1 shows test-retest variability of V_T, parametric SRTM2 derived $BP_{ND}+1$ and SUV_r for several cortical regions, i.e. frontal, MTL, temporal and parietal cortex, posterior cingulate, together with values for the global cortical ROI. As mentioned above, V_T test-retest data were obtained for five controls and four AD patients, and parametric SRTM2 and SUV_r data for six controls and six AD patients. For the global cortical ROI, the highest variability was seen in V_T (AD: 9.8±8.2%; controls: 11.8±13.7%). Test-retest variability was lowest for parametric SRTM2 derived $BP_{ND}+1$ (AD: 3.0±2.9%; controls: 2.5±1.8%). SUV_r yielded acceptable variability (AD: 8.0±7.0%; controls: 4.4±4.2%).

The same patterns were observed for other regions. All methods showed the highest intra-test variability in MTL, the ROI with the smallest size, but again variability of parametric SRTM2 was better than that of the other models. For none of the regions and none of the binding measures were significant differences between AD patients and controls found.

Overall, plasma input methods showed markedly higher variability in both AD and controls compared with reference tissue based methods. For all regions parametric SRTM2 derived $BP_{ND}+1$ showed the best variability.

Correlation between plasma input and reference tissue models

Good agreement (r^2 0.85–0.87) was found between 2T4k derived DVR and both SRTM and parametric SRTM2 derived $BP_{ND}+1$ (Fig. 2a and b, respectively). Data were pooled over all subjects and all regions. Figure 2c illustrates the expected good agreement ($r^2=0.93$) between $BP_{ND}+1$ derived with SRTM and parametric SRTM2. It compares BP_{ND} derived using parametric SRTM2 and SRTM, i.e. with and without fixing the reference tissue efflux rate k_2'.

Table 2 Mean (± SD) absolute test-retest variability (%)

ROI	2T4k (DVR)	2T4k ($BP_{2T4k}+1$)	2T4k (V_T)	SRTM ($BP_{ND}+1$)	Parametric SRTM2 ($BP_{ND}+1$)	SUV_r
AD parietal	10.5 (10.0)	9.5 (11.0)	10.3 (8.9)	3.5 (3.3)	4.0 (3.7)	8.9 (8.0)
AD sensory motor Cx	11.1 (11.3)	9.2 (11.1)	12.5 (10.8)	5.5 (5.7)	2.3 (2.4)	8.9 (8.9)
Control parietal	9.9 (9.2)	12.7 (12.9)	11.4 (15.1)	11.5 (13.3)	2.4 (1.9)	4.7 (5.6)
Control sensory motor Cx	8.3 (8.2)	11.7 (7.4)	11.3 (9.3)	13.4 (18.6)	2.5 (2.8)	4.5 (7.5)

Number of subjects are identical to those given in Table 1

ROI region of interest, *AD* Alzheimer's disease, *Cx* cortex

Table 3 Mean (± SD) difference between test and retest (%)

ROI	2T4k (DVR)	2T4k (BP$_{2T4k}$+1)	2T4k (V$_T$)	SRTM (BP$_{ND}$+1)	Parametric SRTM2 (BP$_{ND}$+1)	SUV$_r$
AD parietal	9.1 (16.9)	10.5 (12.8)	6.8 (14.4)	2.7 (4.3)	2.2 (4.6)	7.5 (11.0)
AD sensory motor Cx	8.6 (14.1)	7.6 (15.2)	6.5 (18.3)	3.4 (8.0)	3.4 (2.6)	7.3 (12.4)
Control parietal	−4.2 (11.3)	−5.1 (13.9)	−7.8 (11.9)	−2.9 (17.0)	−2.2 (2.1)	−2.2 (6.7)
Control sensory motor Cx	−3.3 (10.3)	−4.9 (11.2)	−7.2 (14.7)	−1.5 (19.6)	−2.1 (3.1)	−2.4 (7.8)

Number of subjects are identical to those given in Table 1

ROI region of interest, *AD* Alzheimer's disease, *Cx* cortex

Note, however, that SRTM was applied to ROI time activity curves with low levels of noise, whilst in parametric SRTM2 fitting was performed at the voxel level. The good correlation seen in this figure indicates that parametric SRTM2 can be used without loss of accuracy. Fixation of k_2' is required for generating accurate parametric BP$_{ND}$ images.

Cerebellar data

In Fig. 3 2T4k derived BP$_{2T4k}$ and V$_T$ values of cerebellum grey matter and parietal cortex are given for AD patients ($n=$ 6) and healthy controls ($n=7$). Data of one healthy control was excluded because of a high value in the cerebellum. Although average cerebellum BP$_{2T4k}$ and V$_T$ values for AD subjects (0.96±0.24 and 3.85±0.70, respectively) were somewhat higher than those for controls (0.79±0.21 and 3.24±0.55), there were no significant differences between the groups ($p=0.19$ for BP$_{2T4k}$ and $p=0.11$ for V$_T$). In contrast, compared with controls, a twofold increase was found in the parietal cortex of AD patients. V$_T$ of the parietal cortex in controls was 3.79±0.63 and in AD patients 7.28±1.48 ($p<$ 0.001). BP$_{2T4k}$ of the parietal cortex was 1.30±0.32 and 2.64± 0.66 in controls and AD patients, respectively ($p=0.001$).

Discussion

This study evaluated test-retest variability of [^{11}C]PIB studies using several modelling approaches. Test-retest variability was markedly higher in methods using plasma input data as compared to reference tissue methods, with best variability for parametric SRTM2. The high test-retest variability for plasma input methods was primarily due to the higher number of fit parameters when including fractional blood volume and additional uncertainty inherent in the metabolite analysis. Average cerebellum BP$_{2T4k}$ and V$_T$ values derived with the plasma input model showed no significant differences between the groups, although slightly higher values were observed in AD patients. As such, cerebellum grey matter can be used as a reference tissue for

[^{11}C]PIB studies in AD patients. In addition, parametric SRTM2 is a stable and reliable method for quantifying [^{11}C]PIB binding, even when assessing small target regions. Parametric SRTM2 with the cerebellum as reference region showed excellent test-retest variability and good group differentiation and is therefore the method of choice for quantifying [^{11}C]PIB studies, especially when addressing longitudinal follow-up or treatment response.

The uncertainty in the metabolite analysis is largely due to rapid peripheral metabolism of [^{11}C]PIB, leading to negligible amounts of parent tracer at 90 min [28] and consequently to increased uncertainty in the fits. 2T4k derived DVR performed slightly better than BP$_{2T4k}$+1 in controls. This can be due to the fact that DVR is calculated using V$_T$(target)/V$_T$(cerebellum) which could lead to more stability at low values compared to other methods using plasma input data.

When comparing absolute values obtained using SRTM, parametric SRTM2 and 2T4k derived DVR, good agreement was observed (Fig. 2). When comparing variability between these methods, the better variability obtained with parametric SRTM2 probably is due to the reduced number of fit parameters as a result of fixing k_2'. The poorer variability of 2T4k derived might be due to the fact that this parameter is derived from a ratio of two fitted parameters, which might be more sensitive to error propagation. Furthermore, the uncertainty inherent in the metabolite analysis also contributes to the higher variability. Apart from differences between methods, also differences between AD patients and controls and between regions were seen. Differences in variability (Table 2) were less distinct for the plasma input than for the reference tissue methods. This may be due to noise from the metabolite correction or result from the contribution of non-specific binding, which in contrast to reference tissue methods may be incorporated in BP$_{2T4k}$. For SRTM variability was somewhat higher in controls than in AD patients. This could be due to the lower BP$_{ND}$ in controls (and to some extent in sensory motor cortex of AD), leading to relatively higher variability. With parametric SRTM2 and SUV$_r$ the opposite pattern was seen. In the case of parametric SRTM2 this could be a
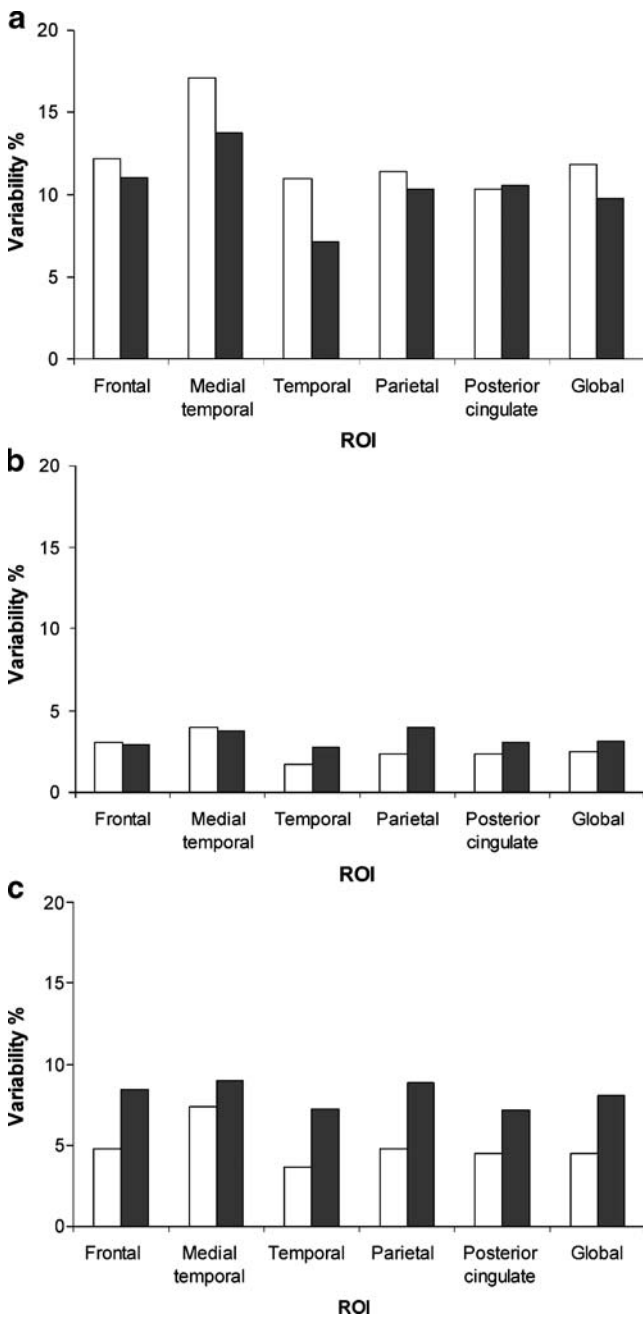
Fig. 1 Test-retest variability is shown for **a** V_T, **b** parametric SRTM2 derived $BP_{ND}+1$ and **c** SUV_r of several regions of interest (ROI): frontal, medial temporal, temporal and parietal cortex, posterior cingulate and global cortical binding. The latter is the volume-weighted average of the previously mentioned regions. In the case of V_T, $n=4$ and 5 for patients with Alzheimer's disease (AD) and healthy controls, respectively; in the case of parametric SRTM2 and SUV_r, $n=6$ for both patient groups. AD: *filled columns*; healthy controls: *open columns*

between subjects in reaching equilibrium. Differences in percent (Table 3) showed a trend towards higher values for retest scans in the AD patients, with the opposite pattern for controls. Differences between test and retest scans were not significant, possibly due to the limited power of the study and the relatively large standard deviations found. However, further research should assess if these trends are also present and possibly even significant in a larger population.

Overall, test-retest variability was similar to that observed for other radiotracers [30]. More importantly, it was also in agreement with values reported earlier for [11C]PIB studies [5, 6, 8]. Somewhat higher values for 2T4k derived DVR were seen compared with an earlier test-retest study of five subjects (two controls, one patient with mild
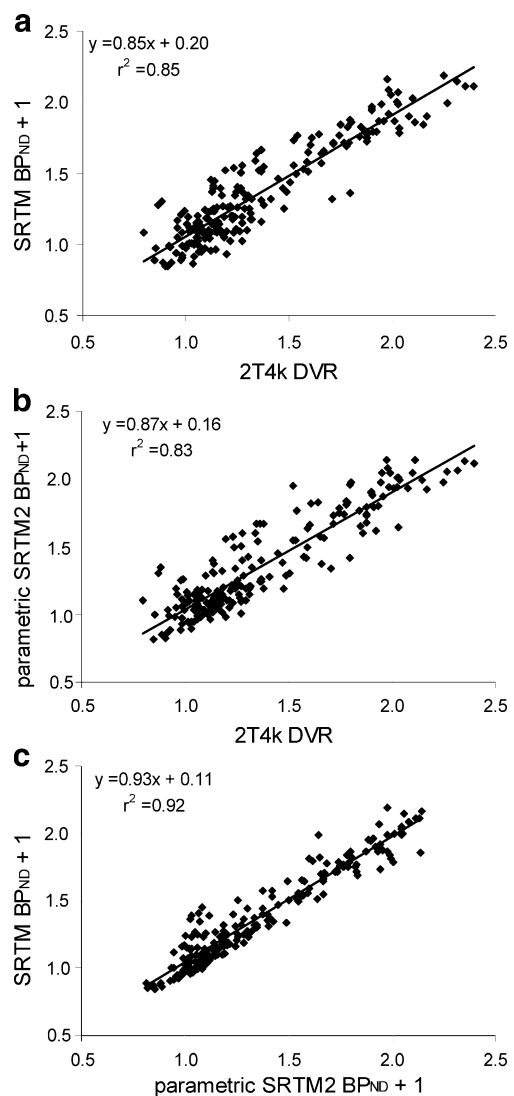

Fig. 2 Linear regression between 2T4k derived DVR-1 and **a** SRTM derived BP_{ND} and **b** parametric SRTM2 derived BP_{ND}, and **c** between SRTM and parametric SRTM2 derived BP_{ND}, all without fixed intercept. Data are pooled over all subjects and all regions
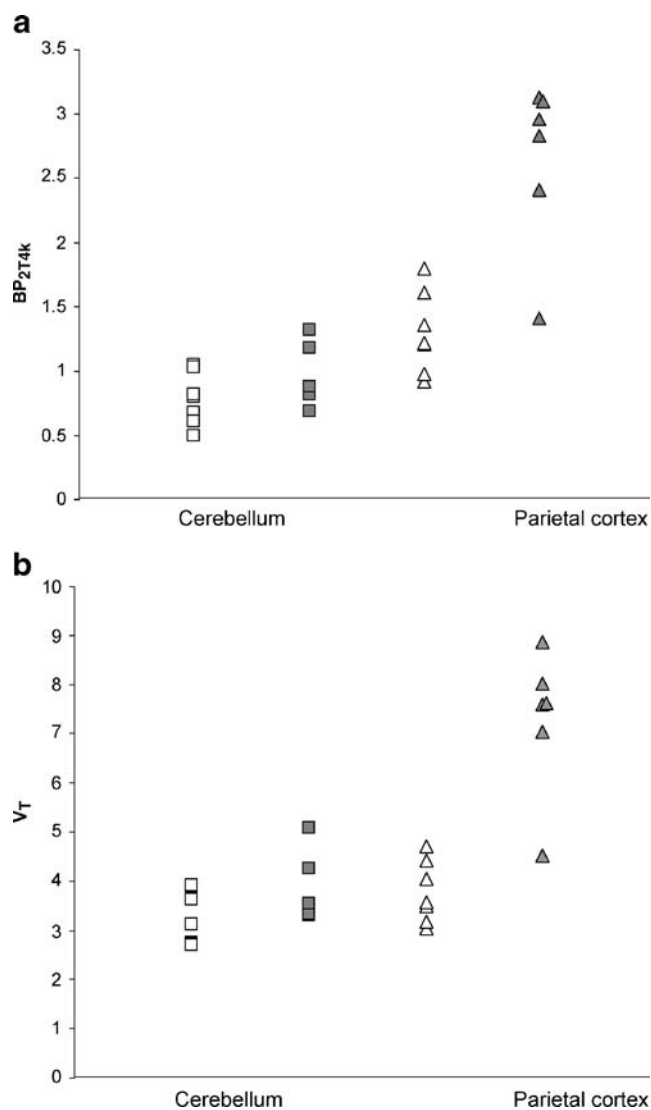
result of boundary conditions, inherent in the use of basis functions, thereby reducing the possibility of erroneous fit parameters, especially for low BP_{ND}. In the case of SUV_r this could possibly be an artefact as a result of differences

Fig. 3 Scatter plots showing the distribution of individual subject 2T4k derived values of **a** BP_{2T4k} and **b** V_T for cerebellum (*squares*) and parietal cortex (*triangles*) using plasma input data of six patients with Alzheimer's disease (AD) (*filled symbols*) and seven healthy controls (*open symbols*). Cerebellar BP_{2T4k} and V_T did not differ significantly between AD patients and healthy controls (BP_{2T4k} $p=$ 0.19, V_T $p=0.11$). Compared with controls a twofold increase was found in AD patients in the parietal cortex (BP_{2T4k} $p<0.001$, V_T $p=$ 0.001)

cognitive impairment and two AD patients) using arterial based Logan DVR (calculated with V_T) as outcome measure [5]. These slightly higher values for variability could be due to the fact that arterial based Logan V_T is sensitive to noise [31]. In another study, the same group assessed several analytical methods and concluded that the non-invasive Logan method showed the lowest test-retest variability (5%) [6]. In the present study parametric SRTM2 showed the best variability. A direct comparison between non-invasive Logan and parametric SRTM2 was made previously [28], showing that reference Logan was slightly

more "biased" than parametric SRTM2 because of its weaker correlation with BP_{2T4k} derived from 2T4k and BP_{ND} derived from SRTM.

For clinical studies, low test-retest variability is important especially for smaller ROI. Whilst in full-blown AD amyloid load is spread throughout the brain [1], in preclinical AD early depositions are more localized. In familial AD this has been reported to be in striatum [32], while in non-demented "PIB-positive" subjects the highest values have been reported in the precuneus [33]. In the present study, parametric SRTM2 had the best test-retest variability, not only for the global cortical ROI, but also for smaller ROI. Therefore, parametric SRTM2 is a stable and reliable method for quantifying [¹¹C]PIB uptake reflecting early amyloid deposition.

To date, most clinical [¹¹C]PIB studies have used the cerebellum as reference tissue for analysing PET data. The use of the cerebellum as reference tissue has circumvented the need for a metabolite-corrected arterial plasma input function. This is important because arterial sampling in general and in AD patients in particular is challenging. Nevertheless, to be able to assess specific [¹¹C]PIB binding in the cerebellum arterial sampling is necessary. The present study is one of the few in patients with AD where plasma input data were available.

An ideal reference tissue should have a negligible concentration of specific binding sites. In the present study, average cerebellar BP_{2T4k} was close to 1 for both AD patients and healthy controls. Although this finding in itself could argue against the use of the cerebellum as reference tissue, BP_{2T4k} and V_T values between both groups did not differ significantly, suggesting non-specific binding (binding other than to fibrillar amyloid) as the major component of cerebellar BP_{2T4k}. It should be noted that the lack of power due to a limited number of subjects could have affected results. Although they did not differ significantly, average values for cerebellum BP_{2T4k} and V_T in this group of mild AD patients were somewhat higher than in controls. If this (minor) difference is indeed due to diffuse amyloid deposits in the cerebellum of AD patients, this could lead to underestimation of specific binding in the target region when using the cerebellum as reference region. In cross-sectional (e.g. diagnostic) studies, minor underestimation is unimportant, as the difference in specific binding between target (e.g. cortical) and reference regions would be high. For longitudinal studies and studies in preclinical AD patients, however, it could become important. If there is an increase in amyloid depositions in the cerebellum with progression of AD, it would lead to progressive underestimation of the amyloid load over time. Future studies should focus on evaluating V_T in the cerebellum of AD patients over time in order to validate it as a proper reference region.

Conclusion

The results of this study underscore the validity of parametric SRTM2 as the method of choice for quantifying $[^{11}C]$PIB PET data, especially in longitudinal studies. In addition, data suggest that cerebellar grey matter can be used as reference tissue in $[^{11}C]$PIB studies, although further studies are needed to assess whether there is a minor degree of specific uptake in the cerebellum of AD patients.

Conflicts of interest None.

References

1. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol 1991;82:239–59. doi:10.1007/BF00308809.
2. Ritchie CW, Bush AI, Mackinnon A, Macfarlane S, Mastwyk M, MacGregor L, et al. Metal-protein attenuation with iodochlorhydroxyquin (clioquinol) targeting Abeta amyloid deposition and toxicity in Alzheimer disease: a pilot phase 2 clinical trial. Arch Neurol 2003;60:1685–91. doi:10.1001/archneur.60.12.1685.
3. Schenk D, Hagen M, Seubert P. Current progress in beta-amyloid immunotherapy. Curr Opin Immunol 2004;16:599–606. doi:10.1016/j.coi.2004.07.012.
4. Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, et al. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. Ann Neurol 2004;55:306–19. doi:10.1002/ana.20009.
5. Price JC, Klunk WE, Lopresti BJ, Lu X, Hoge JA, Ziolko 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.
6. Lopresti BJ, Klunk WE, Mathis CA, Hoge JA, Ziolko 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.
7. Logan J, Fowler JS, Volkow ND, Wang GJ, Ding YS, Alexoff DL. Distribution volume ratios without blood sampling from graphical analysis of PET data. J Cereb Blood Flow Metab 1996;16:834–40. doi:10.1097/00004647-199609000-00008.
8. Engler H, Forsberg A, Almkvist O, Blomquist G, Larsson E, Savitcheva I, et al. Two-year follow-up of amyloid deposition in patients with Alzheimer's disease. Brain 2006;129:2856–66. doi:10.1093/brain/awl178.
9. Yamaguchi H, Hirai S, Morimatsu M, Shoji M, Nakazato Y. Diffuse type of senile plaques in the cerebellum of Alzheimer-type

10. dementia demonstrated by beta protein immunostain. Acta Neuropathol 1989;77:314–9. doi:10.1007/BF00687584.
10. Joachim CL, Morris JH, Selkoe DJ. Diffuse senile plaques occur commonly in the cerebellum in Alzheimer's disease. Am J Pathol 1989;135:309–19.
11. Ikonomovic MD, Klunk WE, Abrahamson EE, Mathis CA, Price JC, Tsopelas ND, et al. Post-mortem correlates of in vivo PiB-PET amyloid imaging in a typical case of Alzheimer's disease. Brain 2008;131:1630–45. doi:10.1093/brain/awn016.
12. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34:939–44.
13. Wilson AA, Garcia A, Chestakova A, Kung HF, Houle S. A rapid one-step radiosynthesis of the beta-amyloid imaging radiotracer N-methyl-[C-11]2-('4-methylaminophenyl)-6-hydroxybenzothiazole ([C-11]-6-OH-BTA-1). J Labelled Compd Radiopharm 2004;47:679–82. doi:10.1002/jlcr.854.
14. Jewett DM. A simple synthesis of $[^{11}C]$methyl triflate. Int J Rad Appl Instrum [A] 1992;43:1383–5. doi:10.1016/0883-2889(92)90012-4.
15. Brix G, Zaers J, Adam LE, Bellemann ME, Ostertag H, Trojan H, et al. Performance evaluation of a whole-body PET scanner using the NEMA protocol. National Electrical Manufacturers Association. J Nucl Med 1997;38:1614–23.
16. Boellaard R, van Lingen A, van Balen SC, Hoving BG, Lammertsma AA. Characteristics of a new fully programmable blood sampling device for monitoring blood radioactivity during PET. Eur J Nucl Med 2001;28:81–9. doi:10.1007/s002590000405.
17. Takkenkamp K, Luurtsema G, Greuter HN, et al. Plasma metabolites of [11C]PIB and [18F]FDDNP [abstract]. J Labelled Compd Radiopharm 2008;50:.
18. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging 1997;16:187–98. doi:10.1109/42.563664.
19. Svarer C, Madsen K, Hasselbalch SG, Pinborg LH, Haugbøl S, Frøkjaer VG, et al. MR-based automatic delineation of volumes of interest in human brain PET images using probability maps. Neuroimage 2005;24:969–79. doi:10.1016/j.neuroimage.2004.10.017.
20. Akaike H. A new look at the statistical identification. IEEE Trans Automat Contr 1974;19:716–23. doi:10.1109/TAC.1974.1100705.
21. Schwarz G. Estimating the dimension of a model. Ann Stat 1978;6:461–4. doi:10.1214/aos/1176344136.
22. Gunn RN, Gunn SR, Cunningham VJ. Positron emission tomography compartmental models. J Cereb Blood Flow Metab 2001;21:635–52. doi:10.1097/00004647-200106000-00002.
23. Lammertsma AA, Bench CJ, Hume SP, Osman S, Gunn K, Brooks DJ, et al. Comparison of methods for analysis of clinical [11C]raclopride studies. J Cereb Blood Flow Metab 1996;16:42–52. doi:10.1097/00004647-199601000-00005.
24. Yaqub M, Boellaard R, Kropholler MA, Lammertsma AA. Optimization algorithms and weighting factors for analysis of dynamic PET studies. Phys Med Biol 2006;51:4217–32. doi:10.1088/0031-9155/51/17/007.
25. Lammertsma AA, Hume SP. Simplified reference tissue model for PET receptor studies. Neuroimage 1996;4:153–8. doi:10.1006/nimg.1996.0066.
26. Wu Y, Carson RE. Noise reduction in the simplified reference tissue model for neuroreceptor functional imaging. J Cereb Blood Flow Metab 2002;22:1440–52. doi:10.1097/00004647-200212000-00004.
27. Tolboom N, Yaqub M, van der Flier WM, Boellaard R, Luurtsema G, Windhorst AD, et al. Detection of Alzheimer pathology in vivo using both 11C-PIB and 18F-FDDNP PET. J Nucl Med 2009;50:191–7. doi:10.2967/jnumed.108.056499.

28. Yaqub M, Tolboom N, Boellaard R, van Berckel BN, van Tilburg EW, Luurtsema G, et al. Simplified parametric methods for [11C]PIB studies. Neuroimage 2008;42:76–86. doi:10.1016/j.neuroimage.2008.04.251.

29. Gunn RN, Lammertsma AA, Hume SP, Cunningham VJ. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. Neuroimage 1997;6:279–87. doi:10.1006/nimg.1997.0303.

30. Volkow ND, Fowler JS, Wang GJ, Dewey SL, Schlyer D, MacGregor R, et al. Reproducibility of repeated measures of carbon-11-raclopride binding in the human brain. J Nucl Med 1993;34:609–13.

31. Slifstein M, Laruelle M. Effects of statistical noise on graphic analysis of PET neuroreceptor studies. J Nucl Med 2000;41:2083–8.

32. Klunk WE, Price JC, Mathis CA, Tsopelas ND, Lopresti BJ, Ziolko SK, et al. Amyloid deposition begins in the striatum of presenilin-1 mutation carriers from two unrelated pedigrees. J Neurosci 2007;27:6174–84. doi:10.1523/JNEUROSCI.0730-07.2007.

33. Mintun MA, LaRossa GN, Sheline YI, Dence CS, Lee SY, Mach RH, et al. [11C]PIB in a nondemented population: potential antecedent marker of Alzheimer disease. Neurology 2006;67:446–52. doi:10.1212/01.wnl.0000228230.26044.a4.