



Parametric imaging and quantitative analysis of the PET amyloid ligand [^{18}F]flutemetamol



Kerstin Heurling^{a,*}, Chris Buckley^b, Koen Van Laere^c, Rik Vandenberghe^d, Mark Lubberink^{a,e}

^a Department of Surgical Sciences, Uppsala University, Uppsala, Sweden

^b GE Healthcare, Amersham, UK

^c Department of Imaging and Pathology, KU Leuven, University Hospitals Leuven, Belgium

^d Department of Neurosciences, KU Leuven, University Hospitals Leuven, Belgium

^e Medical Physics, Uppsala University Hospital, Uppsala, Sweden

ARTICLE INFO

Article history:

Received 13 April 2015

Accepted 14 July 2015

Available online 22 July 2015

Keywords:

PET

Amyloid imaging

[^{18}F]flutemetamol

Parametric imaging

ABSTRACT

Objectives: The amyloid imaging PET tracer [^{18}F]flutemetamol was recently approved by regulatory authorities in the US and EU for estimation of β -amyloid neuritic plaque density in cognitively impaired patients. While the clinical assessment in line with the label is a qualitative visual assessment of 20 min summation images, the aim of this work was to assess the performance of various parametric analysis methods and standardized uptake value ratio (SUVR), in comparison with arterial input based compartment modeling.

Methods: The cerebellar cortex was used as reference region in the generation of parametric images of binding potential (BP_{ND}) using multilinear reference tissue methods (MRTMo, MRTM, MRTM2), basis function implementations of the simplified reference tissue model (here called RPM) and the two-parameter version of SRTM (here called RPM2) and reference region based Logan graphical analysis. Regionally averaged values of parametric results were compared with the BP_{ND} of corresponding regions from arterial input compartment modeling. Dynamic PET data were also pre-filtered using a 3D Gaussian smoothing of 5 mm FWHM and the effect of the filtering on the correlation was investigated. In addition, the use of SUVR images was evaluated. The accuracy of several kinetic models were also assessed through simulations of time-activity curves based on clinical data for low and high binding adding different levels of statistical noise representing regions and individual voxels.

Results: The highest correlation was observed for pre-filtered reference Logan, with correction for individual reference region efflux rate constant k_2' ($R^2 = 0.98$), or using a cohort mean k_2' ($R^2 = 0.97$). Pre-processing filtered MRTM2, unfiltered SUVR over the scanning window 70–90 min and unfiltered RPM also demonstrated high correlations with arterial input compartment modeling (MRTM2 $R^2 = 0.97$, RPM $R^2 = 0.96$ and SUVR $R^2 = 0.95$). Poorest agreement was seen with MRTM without pre-filtering ($R^2 = 0.68$).

Conclusions: Parametric imaging allows for quantification without introducing bias due to selection of anatomical regions, and thus enables objective statistical voxel-based comparisons of tracer binding. Several parametric modeling approaches perform well, especially after Gaussian pre-filtering of the dynamic data. However, the semi-quantitative use of SUVR between 70 and 90 min has comparable agreement with full kinetic modeling, thus supporting its use as a simplified method for quantitative assessment of tracer uptake.

© 2015 Elsevier Inc. All rights reserved.

Introduction

Alzheimer's Disease (AD) is typically described as a condition where presence of progressive dementia is accompanied by neuropathological changes that include neurofibrillary tangles (Braak and Braak, 1991) and β -amyloid plaques (Thal et al., 2002). With a steadily increasing elderly population, the prevalence of dementia and in particular AD is

increasing and more specific techniques to assist in the accurate diagnosis are demanded.

Positron emission tomography (PET) is a non-invasive molecular imaging method for diagnosis and monitoring of treatment effects, where it visualizes physiological processes *in vivo*. Due to its high spatial and temporal resolution, it is well suited for neurological applications including imaging of β -amyloid, initially introduced by Klunk et al. (2004) using the ^{11}C -labelled tracer Pittsburgh Compound B (^{11}C]PiB) (Klunk et al., 2004). The β -amyloid imaging PET tracer [^{18}F]flutemetamol, an ^{18}F -labelled analogue of [^{11}C]PiB, was recently approved by the US Food and Drug Administration and the European Medicines Agency for estimation of β -amyloid neuritic

* Corresponding author at: Uppsala University, Akademiska sjukhuset, 751 85 Uppsala, Sweden.

E-mail address: Kerstin.Heurling@radiol.uu.se (K. Heurling).

plaque density in cognitively impaired patients under the trademark name VizamyTM (GE Healthcare Ltd, UK).

Data included in this study are part of a previously performed Phase I study of [¹⁸F]flutemetamol sponsored by GE Healthcare Ltd. The original study included eight subjects with probable AD (pAD) and 14 healthy volunteers (HV), in total 22 subjects, of which six HV subjects underwent whole body scanning for investigation of biodistribution and dosimetry (Koole et al., 2009). Six subjects underwent dynamic brain PET between time of injection and 250 min post injection divided into three scanning windows, including arterial blood sampling. Ten subjects underwent a simplified scanning procedure, following the interim analysis of the six dynamically scanned subjects, over a limited time window of 80–170 min post injection of [¹⁸F]flutemetamol with no blood sampling (Nelissen et al., 2009). Based on the six dynamically scanned subjects, it was concluded that the volume of interest (VOI) based tracer kinetics were best described by a reversible 2-tissue compartment (2-TC) model (Nelissen et al., 2009). It was also shown that the tracer kinetics in the proposed reference region, the cerebellar cortex—a region known to be largely devoid of β -amyloid accumulation—could be described using a 1-tissue compartment model, but slightly better by a 2-TC model, indicating a small contribution of binding with different kinetic properties than the influx/efflux of the tracer to the uptake. The use of cerebellar cortex as a reference region was validated as having high correlation with plasma input based quantification, along with the use of standardized uptake value ratio (SUVR) over a shorter time window. In the ten subjects scanned between 80 and 170 min, the optimal time for SUVR estimation was investigated. Scanning length and scanning start time was determined to have little effect on the discriminatory abilities of [¹⁸F]flutemetamol beyond 80 min post injection (Nelissen et al., 2009). For clinical use the recommended scanning window is to start acquisition at approximately 90 min post injection and scan over 20 min.

Following this, the efficacy of blinded visual assessments of [¹⁸F]flutemetamol as well as of SUVR for discrimination of pAD subjects from HV was determined in a Phase II study (Vandenberghe et al., 2010) using the clinical diagnosis as standard of truth. Forty subjects (pAD and amnesic mild cognitive impairment (aMCI)) also received a [¹¹C]PiB scan and the correlation of SUVR between the two tracers was found to be excellent in the grey matter cortical regions, but less so in white matter rich regions of the brain. There was a total agreement in the visual assessment of the two tracers. In a cohort of end-of-life patients, the sensitivity and specificity of visual assessments of [¹⁸F]flutemetamol images compared to post mortem diagnosis were determined (Curtis et al., 2015). The sensitivity ranged from 81% to 93% (median 88%) among five blinded readers and the median specificity was 88%.

To date, [¹⁸F]flutemetamol PET scans have mainly been evaluated using visual assessment, or by regional semi-quantitative analysis of the tracer uptake by the means of SUVR using cerebellar cortex or pons as a reference region. Regional quantitative analysis removes the detailed spatial information of the PET images. The loss of such detail can however be circumvented by applying the kinetic models on the finest spatial level i.e. at the voxel level. This enables voxel-based statistical comparisons without any bias from pre-definition of VOIs to test the spatial difference in parameter distribution between groups or in a longitudinal study setting. In addition to the mere quantitative benefit of using parametric imaging, the application of these images in the context of visual assessment has recently been investigated for [¹¹C]PiB images demonstrating a higher interreader and intermethod agreement for BP_{ND} images than SUVR (Zwan et al., 2014). As of today, there have been no reports of the application of parametric imaging to [¹⁸F]flutemetamol PET scanning.

The objective of the current study was to investigate the performance of various methods for creating [¹⁸F]flutemetamol parametric images representing the specific uptake of the tracer, using clinical data as well as testing the applicability of different kinetic models through simulation studies.

Methods

Participants

This post hoc study was limited to only include the image data from the first scanning window in the six subjects with dynamic data of the Phase I subjects, as the remaining subjects did not have data necessary for the analyses performed. Three were healthy volunteers (HV), one male and two females (mean age, range: 62.3, 56–71 years) with a Mini-Mental State Examination (MMSE) scoring above 27 of 30 and a Clinical Dementia Rating of 0. Three subjects had a diagnosis of early-stage clinically probable Alzheimer's Disease, all were males (mean age, range: 62, 55–68 years). The pAD subjects had to fulfil the criteria of the National Institute of Neurological and Communicative Disorders and Stroke and of the Alzheimer's Disease and Related Disorders Association for clinically probable AD as well as the DSM-IV criteria for dementia of Alzheimer's type. The inclusion criteria included a MMSE score between 18 and 26 and a Clinical Dementia Rating between 0.5 and 2. At the time of the PET scanning, all pAD patients were receiving regular treatment with a cholinesterase inhibitor and selective serotonin-reuptake inhibitors, and one was receiving quetiapine. For more detailed information on the neuropsychology testing, see Nelissen et al., 2009. All participants provided their written informed consent in accordance with the Declaration of Helsinki and the study protocol was approved by the Ethical Committee of the University Hospitals Leuven.

Data acquisition

As part of the screening, each subject underwent an MRI scan within 30 days before the subject's PET scan. The MRI was performed on a 1.5 T Intera scanner (Philips, Guildford, Surrey, UK), and the protocol included a high-resolution T1-weighted image. The scan was used to exclude participants with structural abnormalities and for co-registration and segmentation during the PET data processing. The PET studies were performed on a HiRez Biograph 16 PET/CT (Siemens, Knoxville, TN, USA) in 3D list mode, with the acquisition starting at the time of intravenous administration of Flutemetamol (F 18) Injection in the antecubital vein with a target dose of 185 MBq. The subject's head was fixed using a vacuum pillow to avoid excessive head movement. At the beginning of each PET scan, a low dose CT scan (80 kV tube potential, 11 mAs) was conducted for attenuation correction. The data were reconstructed into 25 frames over 90 min (4 × 30, 6 × 60, 4 × 180, 8 × 300 s) using Fourier re-binning and 2D ordered subsets expectation maximization (81 axial slices, 5 iterations and 8 subsets, 128 × 128 matrix, 2.5 × zoom, 2.13 × 2.13 × 2 mm pixels, 2 mm slice separation and a Gaussian filter with a FWHM of 5 mm, no smoothing in the axial direction). The reconstruction included all the appropriate corrections such as for attenuation, randoms and scatter.

Arterial blood samples were collected every 10 s during the first minute, every 15 s up till 3 min after injection and at 4, 8, 15, 30, 45, 60 and 90 min for radioactivity measurements in whole blood and plasma. In addition arterial samples were collected at 2, 5, 20 and 60, 180 and 240 min for determination of percentage of radioactive parent compound and metabolites. Based on this, a metabolite-corrected plasma input curve was created and used for kinetic modeling of the tracer uptake.

Image processing

The dynamic PET data were realigned frame-by-frame to correct for motion (Andersson, 1995) during the scan and each MRI was realigned to a plane parallel to the line connecting the anterior and posterior commissure using VOIager4 (GE Healthcare, Uppsala, Sweden). The PET images were co-registered to the reformatted MRI by means of a summation image of frames 2–10 (30 s–8 min p.i.) using Statistical Parametric Mapping (SPM5; Wellcome Trust Center for Neuroimaging,

University College London, UK). Grey matter VOIs were defined using a probabilistic VOI template in the PVElab software (Svarer et al., 2005). The VOIs were then applied to the dynamic PET data generating mean time–activity curves (TACs) for 16 grey matter regions averaged over the left and right hemispheres (superior frontal cortex, medial inferior frontal cortex, orbitofrontal cortex, dorsolateral prefrontal cortex, ventrolateral prefrontal cortex, superior temporal cortex, medial inferior temporal cortex, insula, parietal cortex, thalamus, sensory motor cortex, occipital cortex, putamen, caudate, anterior cingulate cortex and posterior cingulate cortex).

VOI based kinetic analysis

A 2-tissue compartment (2-TC) model (Gunn et al., 2001) was used for estimation of the tracer binding through non-linear least squares curve-fitting, using the metabolite-corrected plasma input function. The appropriateness of this model in describing the tracer uptake had previously been demonstrated (Nelissen et al., 2009). The model assumes that the equilibrium between free tracer and non-specific binding is too fast to separate them, and they are thus combined into one compartment of non-displaceable binding. The binding potential (BP_{ND}) for each region was estimated both directly and indirectly as the ratio of volumes of distribution V_T in each region relative to cerebellar grey cortex as reference region, resulting in the distribution volume ratio DVR-1 (Lammertsma and Hume, 1996). The compartment model-based binding estimates were considered the gold standard and the other models were compared against them. Clearance rates estimated from 2-TC were used to confirm the validity of the basis functions used in the parametric models. Region-based analysis was also performed on the TACs using the simplified reference tissue model (SRTM, (Lammertsma and Hume, 1996)). Use of the graphical Logan analysis methods was already previously presented on this data as being successful in quantifying the uptake of [^{18}F]flutemetamol (Nelissen et al., 2009).

Parametric images

Parametric maps were generated using an in-house-developed MATLAB based implementation of the following models: multilinear reference tissue models MRTMo, MRTM and MRTM2 (Ichise et al., 1996, 2003), the voxel-based basis function implementations of SRTM (Gunn et al., 1997; Lammertsma and Hume, 1996) and SRTM2 (Wu and Carson, 2002), commonly and hereafter in this paper called receptor parametric mapping (RPM and RPM2), as well as voxel-based reference Logan (Logan et al., 1996). Cerebellar grey matter was used as reference region in all models. A voxel-based representation of SUVR was also created for the summation image over 70–90 min using cerebellar grey matter as reference region, chosen because it was the time window closest to the recommended time window for clinical use. For RPM and RPM2, 50 basis functions with logarithmically spaced exponential clearance rates ranging between 0.012 and 0.6 min^{-1} were used. These boundaries were compared to the clearance rates estimated during the region-based 2-TC analysis of all subjects. The number of parameters estimated in RPM2 is reduced by fixing the reference tissue efflux rate constant (k_2') based on RPM estimates in order to decrease the introduction of quantification errors due to statistical noise. The multilinear reference tissue models (MRTMo, MRTM and MRTM2) estimate the tracer binding by means of multilinear regression. Similar to the RPM2, the number of estimated parameters is reduced for MRTM2, by using a fixed value for k_2' based on the average k_2' from the MRTM model.

The reference Logan model provides parametric maps of DVR, using a linearization of the uptake in each voxel and the reference region over the dynamic window specified, here 50–90 min. As the ratio of target to reference region uptake could not be assumed to be constant during the time window studied, the k_2' for the cerebellar cortex VOI was

determined graphically using plasma input Logan (Logan et al., 1996). Parametric images were created with and without inclusion of the k_2' to assess the impact of this correction, and k_2' was either individually determined or based on the cohort mean. Since some parametric methods, especially graphical analyses, are sensitive to noise (Slifstein and Laruelle, 2000), the dynamic reconstructed data were smoothed by applying an isotropic Gaussian filter of 5 mm FWHM (Klumpers et al., 2012) before estimation of parametric images and compared to no additional smoothing. SUVR images were excluded from this analysis, as they do not suffer from noise impact on model fitting in the same way as the parametric models.

Evaluation of parametric methods

The grey matter masked VOIs, generated using PVElab, were applied to all parametric images for estimation of regional average values of the voxels in each VOI, for either BP_{ND} , DVR-1 or SUVR-1. The square of the correlation coefficient (R^2) and slope and intercept for the linear regression of the 2-TC based BP_{ND} and the parametric image estimate were estimated for each model. The regional estimate of orbitofrontal cortex for each subject and each model was used to visualize the discriminatory ability of the models, however due to the limited sample size no formal statistical analysis was made.

Simulations

The validity of different kinetic models in the estimation of [^{18}F]flutemetamol was tested through simulations, where TACs were generated for target and reference regions representing typical [^{18}F]flutemetamol kinetics using a plasma input curve and the plasma input reversible 2-TC model (Gunn et al., 2001). Parameter values were based on the clinical data. A noise level of 2% or 10%, corresponding to noise levels present in VOI TACs or single voxel TACs, was added to the simulated target signal, whereas a noise level of 2% was added to the simulated reference region TAC. The models tested were the 2-TC model estimation of BP_{ND} and DVR-1, SRTM, reference Logan and RPM. One hundred TACs with added noise were generated, and the evaluation was based on the bias and coefficient of variation (CoV) of both BP_{ND} and DVR-1. In addition to this, SUVR was investigated for the time interval 70–90 min. To test the correlation between the true BP_{ND} or DVR-1 and the different simplifications, 100 different BP_{ND} and DVR-1 levels randomly chosen between 0.38 and 1.14 were simulated and the square of correlation coefficient (R^2) and slope between the different binding estimates were computed.

Results

Plasma input functions and time activity curves

Examples of TACs for orbitofrontal cortex, a region where β -amyloid accumulation is often observed, and cerebellar cortex, known to have none to very little accumulation of β -amyloid, are shown for a representative HV subject and pAD in Fig. 1a and b, together with the average metabolite-corrected plasma input function in 1c. In Fig. 1d the mean and standard deviation across all subjects for the fraction of remaining unmetabolized [^{18}F]flutemetamol at each time point are shown.

VOI based kinetic analysis

The kinetics of [^{18}F]flutemetamol were well described by the 2-TC model. However, the indirect estimates of BP_{ND} as DVR-1, through estimation of V_T and using cerebellar cortex as a reference region proved more robust than the direct estimations of BP_{ND} based on the variance. While SRTM resulted in the best fits to experimental data according to the Akaike criterion, the correlation with the 2-TC model estimates was poor. The estimated kinetic parameters from the direct 2-TC

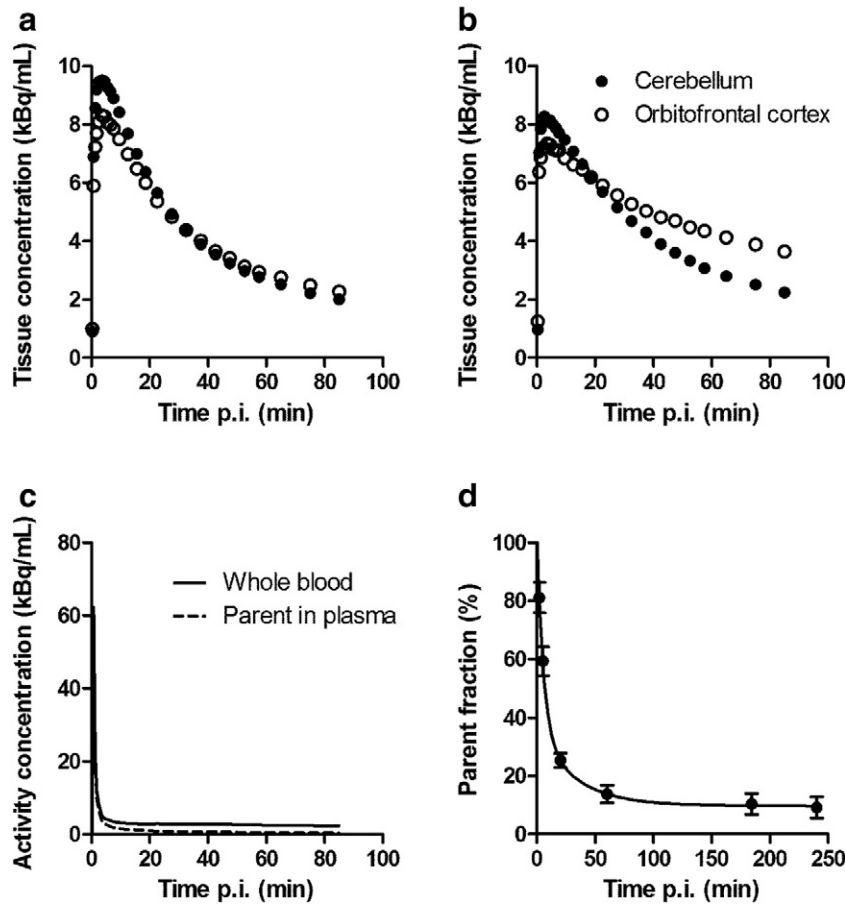


Fig. 1. Time activity curves for a representative a) HV subject and b) pAD subject in the orbitofrontal cortex and cerebellum, c) the mean metabolite corrected plasma input function and whole blood activity concentration and d) the mean fraction of unmetabolized tracer at different time points and standard deviation of all 6 subjects.

model in cerebellar cortex and a global cortical region (including frontal cortex, parietal cortex, anterior cingulate, posterior cingulate and temporal cortex) for pAD and HV subjects are listed in Table 1. The estimated parameters were used as input for the simulation studies. Based on the conclusion that the indirectly estimated BP_{ND} (i.e. DVR-1) values were a more robust estimate, the performance of the parametric models was assessed using that as gold standard.

Parametric images

The grey matter VOI averages from each parametric image were compared to the 2-TC estimate of DVR-1. The cohort mean k_2' in cerebellar cortex, as determined graphically using region-based arterial input Logan, was 0.049 ± 0.010 (range 0.036–0.074 min^{-1}), hence reference Logan parametric images were generated using 0.049 as k_2' for the reference region, each subject's individually determined k_2' or none.

The definition of basis functions in the RPM methods was compared to the clearance rates estimated from the regional 2-TC, and confirmed

that the clearance rates observed were well within the defined window used for the basis functions (2-TC clearance rates: 0.033 to 0.13 min^{-1} , basis function definition clearance rates: 0.012 to 0.6 min^{-1}).

R^2 , slope and intercept of the linear regression between the 2-TC model estimates and corresponding based on each parametric method for all subjects and all regions are listed in Table 2. The inclusion of k_2' in the generating of reference Logan parametric images had no major effect on slope and intercept, although the highest R^2 was achieved when including each subject's individual k_2' , and decreased somewhat when approximating the k_2' with the cohort mean or ignoring k_2' altogether. Pre-filtering of the dynamic data using a 3D Gaussian filter of 5 mm FWHM improved the correlation with 2-TC in all multilinear reference models, as well as in the reference Logan where k_2' was based on the cohort mean. In the remaining two reference Logan models, the correlation coefficient was largely unaffected by the applied filter, while the correlation of basis function methods RPM and RPM2 decreased slightly.

The highest correlations between 2-TC and the parametric methods were found for MRTM2 with pre-filter, reference Logan with individual

Table 1
Average estimated values (\pm SD) of kinetic parameters, as obtained by the 2-TC.

	V_T	BP_{ND}	K_1	k_2	k_3	k_4	V_b
HV, Cerebellar cortex (mean \pm SD, n = 3)	4.57 ± 0.49	0.46 ± 0.03	0.28 ± 0.02	0.09 ± 0.01	0.02 ± 0.005	0.03 ± 0.01	0.06 ± 0.004
pAD, Cerebellar cortex (mean \pm SD, n = 3)	3.88 ± 0.57	0.55 ± 0.37	0.30 ± 0.23	0.18 ± 0.02	0.05 ± 0.05	0.09 ± 0.03	0.05 ± 0.001
HV, Global cortical (mean \pm SD, n = 3)	5.76 ± 0.72	0.87 ± 0.28	0.25 ± 0.03	0.08 ± 0.004	0.02 ± 0.005	0.02 ± 0.003	0.06 ± 0.01
pAD, Global cortical (mean \pm SD, n = 3)	5.86 ± 0.93	0.90 ± 0.26	0.23 ± 0.03	0.08 ± 0.01	0.02 ± 0.01	0.02 ± 0.002	0.050 ± 0.01

Table 2Regression parameters (Pearson's correlation coefficient R^2 , slope and intercept) of parametric image estimates with DVR-1 from arterial input 2-TC.

Parametric outcome	Unfiltered			Pre-filtered 5 mm FWHM 3D Gaussian		
	R^2	Slope	Intercept	R^2	Slope	Intercept
MRTMo (BP_{ND})	0.87	0.73	-0.07	0.89	0.76	-0.09
MRTM (BP_{ND})	0.68	0.79	-0.04	0.84	0.99	-0.10
MRTM2 (BP_{ND})	0.81	0.97	0.08	0.97	1.08	-0.01
Ref Logan with individual k_2' (DVR-1)	0.97	0.90	-0.03	0.98	0.89	-0.03
Ref Logan with cohort mean k_2' (DVR-1)	0.85	0.84	0.01	0.97	0.92	-0.04
Ref Logan without k_2' (DVR-1)	0.94	0.83	-0.05	0.94	0.82	-0.07
RPM (BP_{ND})	0.95	1.08	-0.04	0.94	1.08	-0.06
RPM2 (BP_{ND})	0.89	1.09	-0.10	0.86	1.05	-0.12
SUVR ₇₀₋₉₀ (SUVR-1)	0.96	1.22	0.04	-	-	-

k_2' (both with and without pre-filter), reference Logan with cohort mean k_2' if pre-filter was applied, followed by RPM and SUVR without pre-filtering. Correlation plots of MRTM, RPM, reference Logan and SUVR₇₀₋₉₀₋₁ are shown in Fig. 2.

The parametric images of the different models from a subject with pAD are shown in Fig. 3, selecting either the unfiltered or pre-filtered version of each model that had the highest correlation to 2-TC quantification. The multilinear reference models and the basis function methods had a similar visual appearance while the reference Logan methods and the SUVR image looked alike. MRTM suffered from extreme value artifacts in spite of the applied pre-filter, although this decreased through the application of a pre-processing filter. Examples of the improvement in image quality of the multilinear reference models can be seen in Fig. 4.

The estimated binding or uptake in orbitofrontal cortex was plotted for all subjects to visualize the separation of the data for patient group in

each model. All models demonstrate a good separation between pAD and HV subjects, but also identify one of the HV subjects as having raised [¹⁸F]flutemetamol uptake. The estimated uptake in orbitofrontal cortex of each subject for all models, with or without pre-processing filter applied as determined above, is shown in Fig. 5.

Simulations

Strong correlations were seen for RPM between the true values and the model estimates ($R^2 = 0.99$) when noise corresponding to the noise in VOI TACs was added. SUVR₇₀₋₉₀₋₁ and DVR-1 based on both reference Logan and 2-TC also showed good correlation ($R^2 = 0.97$ for SUVR₇₀₋₉₀₋₁, $R^2 = 0.96$ for reference Logan and $R^2 = 0.94$ for 2-TC), whereas SRTM performed poorly ($R^2 = 0.18$) failing to find an appropriate fit of the model to the data in approximately 50% of the simulations indicating that SRTM is not appropriate for quantification of [¹⁸F]flutemetamol

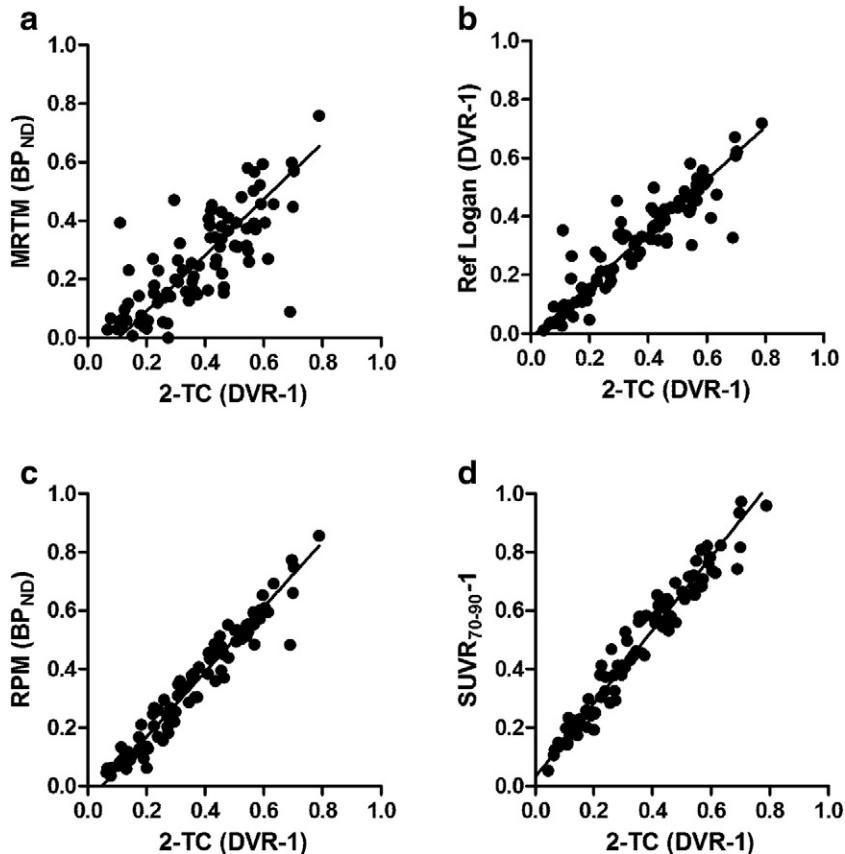


Fig. 2. Scatter plots of estimates of uptake in VOIs of HV and pAD subjects ($n = 6$ total) with 2-TC DVR-1 versus a) MRTM BP_{ND} , b) reference Logan (with cohort mean k_2') DVR-1, c) RPM DVR-1 and d) SUVR₇₀₋₉₀₋₁.

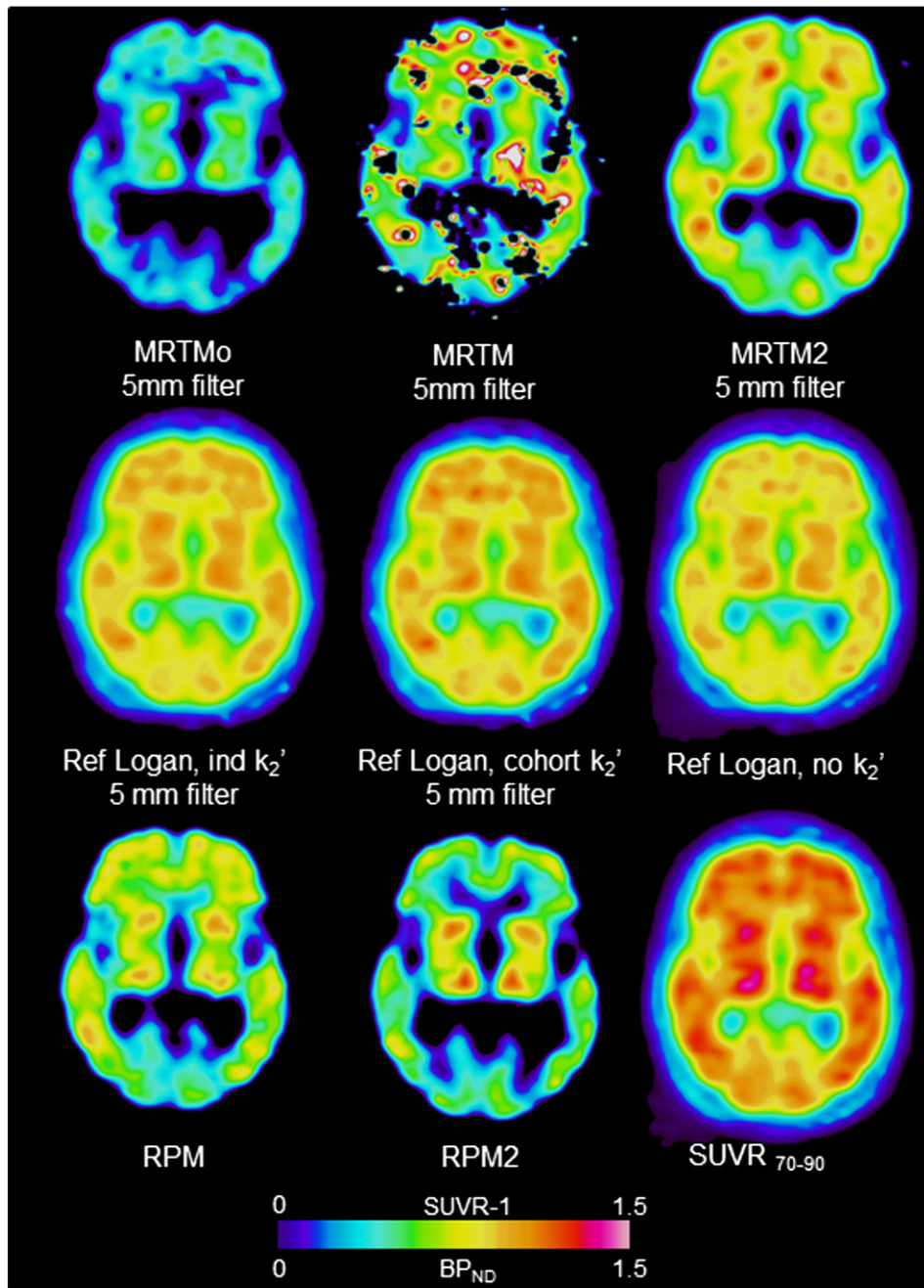


Fig. 3. [^{18}F]Flutemetamol parametric images of an axial plane through the middle of the brain from a typical pAD subject. Four different classes of parametric methods, and in total nine different variants, were used. For reference Logan the voxel unit is DVR-1, with the same maximum image intensity as BP_{ND}, and for SUVR₇₀₋₉₀ the ratio of voxel radioactivity concentration and cerebellar radioactivity concentration (SUVR-1).

uptake. The highest accuracy was found for SUVR₇₀₋₉₀₋₁ and reference Logan: the Pearson correlation coefficients, slopes and intercept of the correlations and the bias and CoV of the simulated models are listed in Table 3.

With a higher noise level added to the simulated TACs, reflecting the noise of a voxel TAC, the correspondence was poorer for all models demonstrating the main challenge with parametric imaging. However, RPM showed the highest correlation ($R^2 = 0.83$) again, followed by SUVR₇₀₋₉₀₋₁ ($R^2 = 0.76$) and reference Logan ($R^2 = 0.64$). The Pearson correlation coefficients, slopes, intercept, bias and CoV are listed in Table 3. Graphs of the correlation for the 100 simulated BP values both on VOI and voxel noise levels are shown in Fig. 6.

Discussion

To date, there have been no records of the application and performance of parametric image methods on any of the [^{18}F]-labelled compounds for amyloid imaging. However, both the technical and clinical application of parametric imaging have been investigated for the ^{11}C -labelled compound [^{11}C]PiB, demonstrating high accuracy and precision of RPM2 and MRTM2 parametric methods (Yaqub et al., 2008) as well as higher interreader and intermethod agreement for BP_{ND} images (Zwan et al., 2014) than SUVR images.

In this study the performance of eight different reference tissue based parametric imaging methods and SUVR images for [^{18}F]flutemetamol

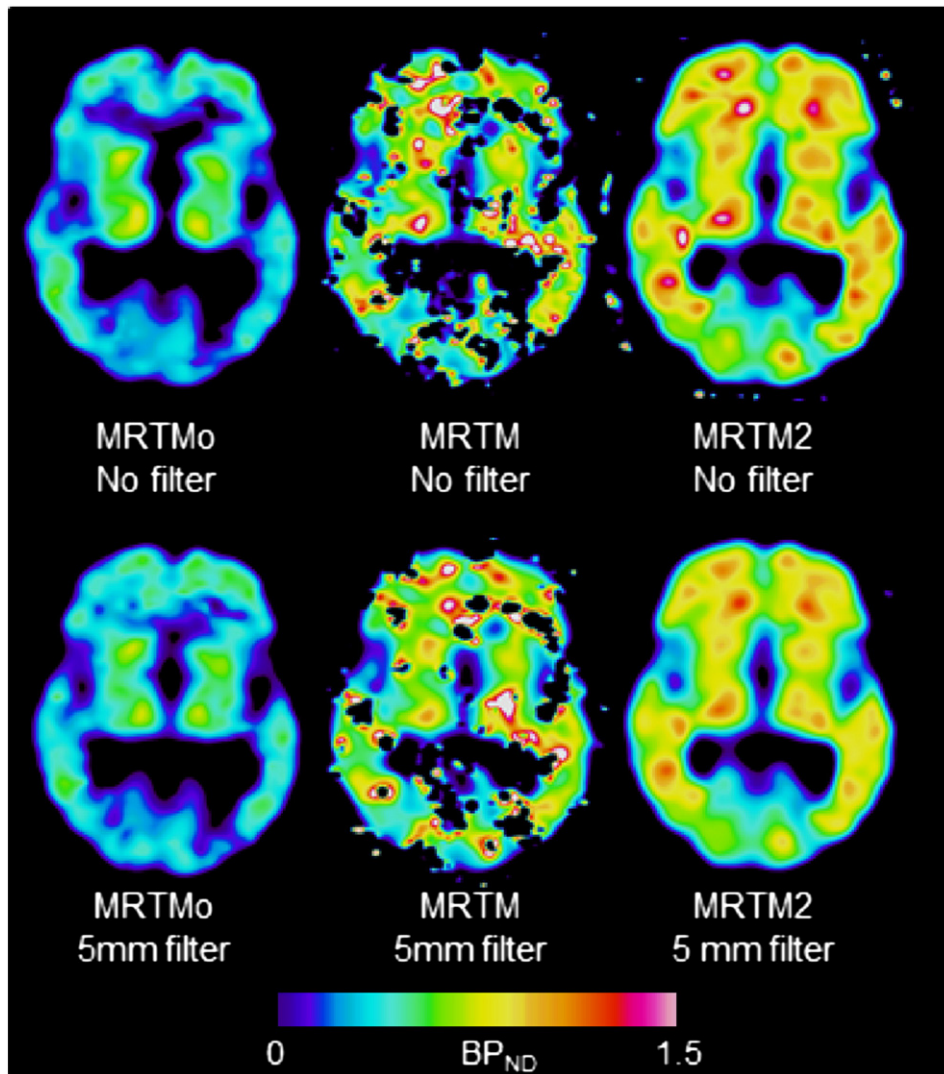


Fig. 4. Effect of pre-processing filter on multilinear reference tissue models. Top row showing MRTMo, MRTM and MRTM2 without any applied filter, bottom row showing the same models but where dynamic PET data had been smoothed using a 3D Gaussian filter with 5 mm FWHM before the kinetic models were applied on voxel-level in the same pAD subject as in Fig. 3.

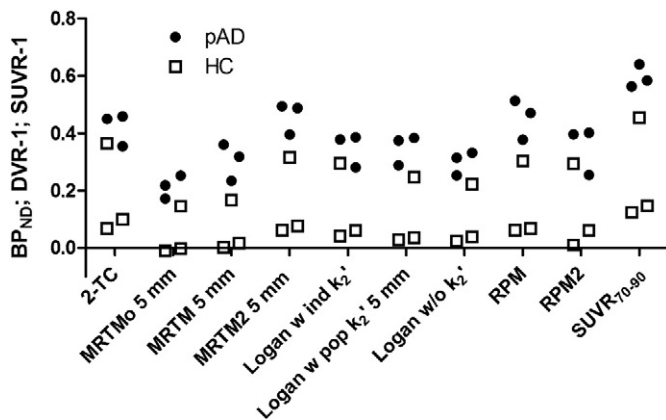


Fig. 5. Nine methods of quantification of [^{18}F]flutemetamol uptake in orbitofrontal cortex in three AD subjects (circles) and three HV subjects (squares), showing the discriminatory ability of each method. The leftmost series is the region based arterial input 2-TC quantification of DVR-1, shown as a reference. One of the HV subjects has elevated uptake, similar to the level of the pAD subjects.

was investigated by comparison with compartmental modeling using arterial plasma input, as well as the impact of pre-processing filtering on the performance of the kinetic models. The theoretical performance of different kinetic models using simulations was also investigated.

Several models performed well, demonstrating a high agreement of the regional averages of the parametric images to the regional based 2-tissue compartment (2-TC) model. Among the eight parametric methods, three had correlation coefficients $R^2 \geq 0.9$. After applied pre-filter, five of the eight parametric methods had $R^2 \geq 0.9$. The methods benefiting most of the applied filter was the linear methods (multilinear reference tissue models and reference Logan models), likely since these methods are highly sensitive to the noise decreased by the Gaussian 5 mm FWHM filter. Pre-filtered reference Logan corrected using individual k_2' and pre-filtered MRTM2 had the highest correlation with 2-TC modeling of all investigated methods, while MRTM showed the poorest agreement. SUVR over the scanning window 70–90 min also showed a high agreement with compartmental modeling. Filter type and width was chosen based on previously published findings (Klumpers et al., 2012) and no optimization was done to investigate whether a different filter size could further improve the correlations with the 2-TC quantification. While a higher FWHM may decrease noise more, it may also induce higher partial volume effects.

Table 3
Accuracy and precision of simulations. R², slope and intercept vs 2-TC.

Model	Target noise level	Reference parameter	R ²	Slope	Intercept	Bias	CoV
2-TC BP _{ND}	2%	2-TC BP _{ND}	0.94	0.98	0.03	1.5	6.1
2-TC DVR-1	2%	2-TC DVR-1	0.91	0.97	0.02	3.3	12.1
SRTM BP _{ND}	2%	2-TC DVR-1	0.18	-41.40	29.61	N/A	N/A
RPM BP _{ND}	2%	2-TC DVR-1	0.99	1.12	-0.08	-4.6	4.6
Ref Logan DVR-1	2%	2-TC DVR-1	0.96	0.89	-0.03	-17.4	8.1
SUVR ₇₀₋₉₀₋₁	2%	2-TC DVR-1	0.97	1.13	-0.02	11.7	6.6
2-TC BP _{ND}	10%	2-TC BP _{ND}	0.01	-6.77	10.52	N/A	N/A
2-TC DVR-1	10%	2-TC DVR-1	0.005	-6.34	7.13	N/A	N/A
SRTM BP _{ND}	10%	2-TC DVR-1	0.00	0.43	15.96	N/A	N/A
RPM BP _{ND}	10%	2-TC DVR-1	0.83	1.08	-0.11	-12.5	22.4
Ref Logan DVR-1	10%	2-TC DVR-1	0.64	0.77	-0.04	-28.2	38.0
SUVR ₇₀₋₉₀₋₁	10%	2-TC DVR-1	0.76	1.09	0.01	15.7	20.5

Reference Logan parametric images were created both with and without correction for k_2' , based either on a cohort mean or determined individually for each subject using arterial input Logan. The inclusion of this correction factor had no major impact on the slope or intercept of the correlation with the 2-TC model, but the correlation coefficient was highest when including the individually determined k_2' and lowest when including the cohort mean. However, applying a pre-processing filter to the dynamic data before quantification increased the correlation between 2-TC modeling and the k_2' -corrected reference Logan models substantially, resulting in the highest observed correlation coefficient among all models for reference Logan corrected using the individually determined k_2' . It should be noted however, that while the inclusion of a correction for k_2' increases the model correlation with 2-TC estimates, this requires arterial blood sampling for

determination of the individual values or the use of an approximation of a mean k_2' . Reported here was a cohort mean of 0.049 min^{-1} , but this should be used with caution as it is based on a limited sample size of only six subjects. In the light of this, it may be preferable to ignore k_2' altogether when using reference Logan for [¹⁸F]flutemetamol parametric images.

The discriminating abilities of the different parametric methods were not directly investigated in this study due to the limited sample size. However, we have graphically presented the uptake in an example region, orbitofrontal cortex, quantified for each subject using all methods. One of the healthy controls showed an elevated uptake, most likely a case of prodromal β -amyloid accumulation, but the remaining two can be easily distinguished from the pAD subjects regardless of method.

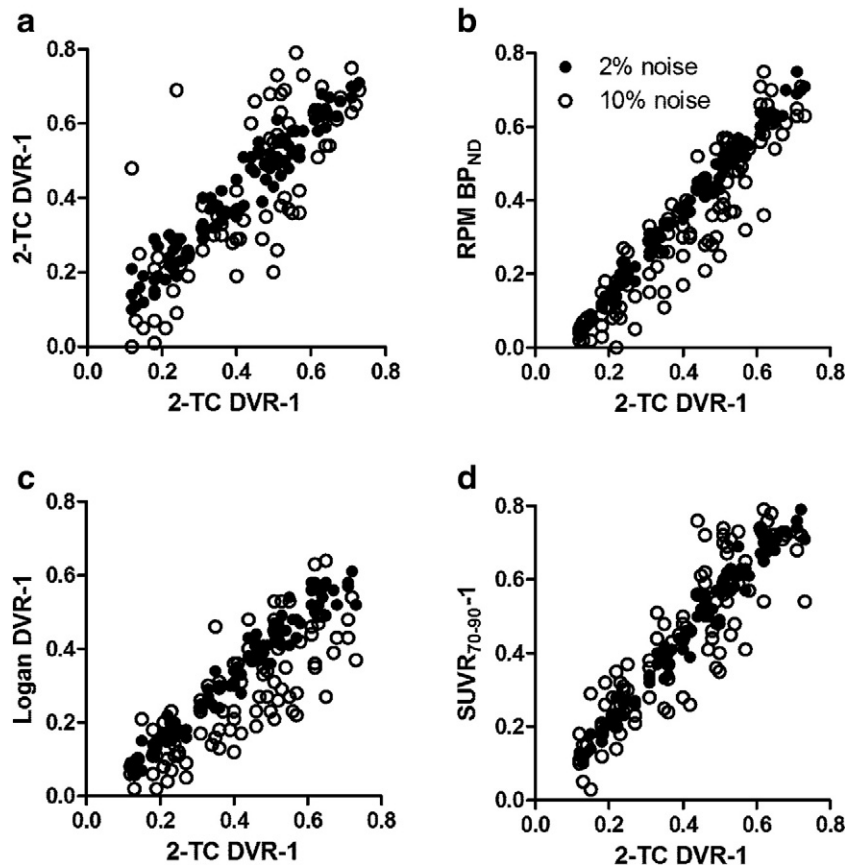


Fig. 6. Model estimation correlation of a range of simulated endpoints for a) 2-TC, b) RPM, c) reference Logan and d) SUVR₇₀₋₉₀₋₁ vs 2-TC DVR-1, for noise level 2% (filled circles) representing VOI analysis and 10% (open circles) representing voxel based analysis. Note that the y-axis maximum in a) is set to 0.8 so extreme values for the 10% estimates, causing the low R², are not shown.

The simulation study showed that 2-TC estimated BP_{ND} has a high accuracy, but a lower precision at the 2% noise level, representing the noise in a VOI based analysis. While the 2-TC model can be sensitive to noise in the time–activity data, this indicates that for the region-based quantification of [^{18}F]flutemetamol compartment modeling performs well and thus is appropriate to use as gold standard for the evaluation of the parametric methods. Of the parametric methods included in the present work, RPM and $SUVR_{70-90}$ had both high accuracy and precision.

SRTM demonstrated both a very low precision and low accuracy, due to the model failing to find a good fit to the simulated TACs in about 50% of the simulations. Therefore, relevant estimates of bias and CoV could not be achieved. A possible reason for the instability of SRTM may be the relatively small specific binding component in the [^{18}F]flutemetamol signal, as well as the fact that neither target nor reference tissue TACs can be described by a single tissue compartment model (Nelissen et al., 2009), which is one of the prerequisites for SRTM. RPM, a basis function implementation of SRTM, and RPM2 perform better probably because these models put additional constraints on the possible parameter values.

When 10% noise was added to the TACs, which was approximated to represent the noise level in a voxel-based parametric image analysis, the 2-TC estimated BP_{ND} had frequent problems finding the correct fit of the model to simulated data, so no estimates of accuracy and precision could be estimated, which is an indication of the sensitivity to noise of this model. Reference Logan and $SUVR_{70-90-1}$ had higher accuracy but a somewhat lower precision than RPM.

One challenging factor with quantification of the [^{18}F]flutemetamol signal is spill-over from white matter to grey matter regions. Since the clearance of [^{18}F]flutemetamol from white matter is slower than from grey matter, this may cause the appearance of spurious slow compartments in grey matter regions. The basis function implementation of SRTM2, RPM2, largely removes the signal from non-specific white matter and demonstrates the tracer uptake localized in cortical grey matter. Future studies will assess the effect of partial volume correction on the quantification of [^{18}F]flutemetamol data but has been shown to improve the accuracy in regional SUVR estimates (Thomas et al., 2011).

In a recent report, the use of parametric images of [^{11}C]PiB in visual assessment was investigated (Zwan et al., 2014) showing a higher interreader agreement for the assessment of images showing BP_{ND} than for the SUV and SUVR especially among non-AD dementia and amnesic MCI patients and healthy controls, but to a lesser degree in AD patients where the interreader agreement was excellent for all methods. From the results of our study, it can be hypothesized that similar effects would be seen for [^{18}F]flutemetamol. In the clinical setting however, the gain in interreader agreement has to be put in relation to the longer required scanning times.

Conclusions

Parametric imaging allows for quantification without any introduced bias due to choice of anatomical regions, thus enabling objective statistical comparisons of the uptake. For the purpose of quantification of BP_{ND} or DVR-1, reference Logan including correction for k_2' shows the highest correlation with arterial input compartment modeling. However, the semi-quantitative use of uptake ratios (SUVR) on a voxel-based level has an almost equally good agreement as the full kinetic modeling, enables shorter scanning windows and does not require arterial blood-sampling, which support its use as a simplified method for quantitative assessment of tracer uptake.

Acknowledgments

The authors would like to express their gratitude to the participating volunteers, patients and their families. We would also like to thank the staff at KU Leuven and GE Healthcare for all the work in the original

Phase I study, and for making the data available for this study. The study was designed and sponsored by GE Healthcare. Ms Heurling is a former employee of GE Healthcare, Dr Buckley is an employee of GE Healthcare. Dr Vandenberghe was principal investigator and Dr Van Laere was co-investigator in the original GE Healthcare sponsored Phase I study.

References

- Andersson, J.L.R., 1995. A rapid and accurate method to realign PET scans utilizing image edge information. *J. Nucl. Med.* 36, 657–669.
- Braak, H., Braak, E., 1991. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* 82, 239–259.
- Curtis, C., Gamez, J.E., Singh, U., Sadowsky, C.H., Villena, T., Sabbagh, M.N., Beach, T.G., Duara, R., Fleisher, A.S., Frey, K.A., Walker, Z., Hunjan, A., Holmes, C., Escovar, Y.M., Vera, V.X., Agronin, M.E., Ross, J., Zozoki, A., Akinola, M., Shi, J., Vandenberghe, R., Ikonovic, M.D., Sherwin, P.F., Grachev, I.D., Farrar, G., Smith, A.P.L., Buckley, Christopher J., McLain, R., Salloway, S., 2015. Phase 3 trial of flutemetamol labeled with radioactive fluorine 18 imaging and neuritic plaque density. *JAMA Neurol.* 72, 287–294.
- Gunn, R.N., Lammertsma, A.A., Hume, S.P., Cunningham, V.J., 1997. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *NeuroImage* 6, 279–287.
- Gunn, R.N., Gunn, S.R., Cunningham, V.J., 2001. Positron emission tomography compartmental models. *J. Cereb. Blood Flow Metab.* 21, 635–652.
- Ichise, M., Ballinger, J.R., Golan, H., Vines, D., Luong, A., Tsai, S., Kung, H.F., 1996. Noninvasive quantification of dopamine D2 receptors with iodine-123-IBF SPECT. *J. Nucl. Med.* 37, 513–520.
- Ichise, M., Liow, J.-S., Lu, J.-Q., Takano, A., Model, K., Toyama, H., Suhara, T., Suzuki, K., Innis, R.B., Carson, R.E., 2003. Linearized reference tissue parametric imaging methods: application to [^{11}C]DASB positron emission tomography studies of the serotonin transporter in human brain. *J. Cereb. Blood Flow Metab.* 23, 1096–1112. <http://dx.doi.org/10.1097/01.WCB.0000085441.37552.CA>.
- Klumpers, U.M.H., Boellaard, R., Veltman, D.J., Kloet, R.W., Hoogendijk, W.J.G., Lammertsma, A. a., 2012. Parametric [^{11}C]flumazenil images. *Nucl. Med. Commun.* 33, 422–430. <http://dx.doi.org/10.1097/MNM.0b013e32835057b>.
- Klunk, W.E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D.P., Bergstro, M., Savitcheva, I., Debnath, M.L., Barletta, J., Price, J.C., Sandell, J., Lopresti, B.J., Wall, A., Koivisto, P., Antoni, G., Mathis, C.A., Långstrom, B., 2004. Imaging brain amyloid in Alzheimer's disease with Pittsburgh compound-B. *Ann. Neurol.* 55, 306–319.
- Koole, M., Lewis, D.M., Buckley, C., Nelissen, N., Vandenbulcke, M., Brooks, D.J., Vandenberghe, R., Van Laere, K., 2009. Whole-body biodistribution and radiation dosimetry of 18F-GE067: a radioligand for in vivo brain amyloid imaging. *J. Nucl. Med.* 50, 818–822. <http://dx.doi.org/10.2967/jnumed.108.060756>.
- Lammertsma, A.A., Hume, S.P., 1996. Simplified reference tissue model for PET receptor studies. *NeuroImage* 4, 153–158. <http://dx.doi.org/10.1006/nimg.1996.0066>.
- Logan, J., Fowler, J.S., Volkow, D., Wang, G., Ding, Y., Alexoff, D.L., 1996. Distribution volume ratios without blood sampling from graphical analysis of PET data. *J. Cereb. Blood Flow Metab.* 16, 834–840.
- Nelissen, N., Van Laere, K., Thurfjell, L., Owenius, R., Vandenbulcke, M., Koole, M., Bormans, G., Brooks, D.J., Vandenberghe, R., 2009. Phase 1 study of the Pittsburgh compound B derivative 18F-flutemetamol in healthy volunteers and patients with probable Alzheimer disease. *J. Nucl. Med.* 50, 1251–1259. <http://dx.doi.org/10.2967/jnumed.109.063305>.
- Slifstein, M., Laruelle, M., 2000. Effects of statistical noise on graphic analysis of PET neuroreceptor studies. *J. Nucl. Med.* 41, 2083–2088.
- Svarer, C., Madsen, K., Hasselbalch, S.G., Pinborg, L.H., Haugbøl, S., Frøkjær, V.G., Holm, S., Paulson, O.B., Knudsen, G.M., 2005. MR-based automatic delineation of volumes of interest in human brain PET images using probability maps. *NeuroImage* 24, 969–979. <http://dx.doi.org/10.1016/j.neuroimage.2004.10.017>.
- Thal, D.R., Rüb, U., Orantes, M., Braak, H., 2002. Phases of A β -deposition in the human brain and its relevance for the development of AD. *Neurology* 58, 1791–1800.
- Thomas, B., Erlandsson, K., Modat, M., Thurfjell, L., Vandenberghe, R., Ourselin, S., Hutton, B., 2011. The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease. *Eur. J. Nucl. Med. Mol. Imaging* 38, 1104–1119. <http://dx.doi.org/10.1007/s00259-011-1745-9>.
- Vandenberghe, R., Van Laere, K., Ivanou, A., Salmon, E., Bastin, C., Triau, E., Hasselbalch, S., Law, I., Andersen, A., Korner, A., Minthon, L., Garraux, G., Nelissen, N., Bormans, G., Buckley, C., Owenius, R., Thurfjell, L., Farrar, G., Brooks, D.J., 2010. 18F-flutemetamol amyloid imaging in Alzheimer disease and mild cognitive impairment: a Phase 2 trial. *Ann. Neurol.* 68, 319–329. <http://dx.doi.org/10.1002/ana.22068>.
- Wu, Y., Carson, R.E., 2002. Noise reduction in the simplified reference tissue model for neuroreceptor functional imaging. *J. Cereb. Blood Flow Metab.* 22, 1440–1452. <http://dx.doi.org/10.1097/01.WCB.0000033967.83623.34>.
- Yaqub, M., Tolboom, N., Boellaard, R., Van Berckel, B.N.M., Van Tilburg, E.W., Luurtsema, G., Scheltens, P., Lammertsma, A.A., 2008. Simplified parametric methods for [^{11}C] PiB studies. *NeuroImage* 42, 76–86. <http://dx.doi.org/10.1016/j.neuroimage.2008.04.251>.
- Zwan, M.D., Ossenkoppele, R., Tolboom, N., Beunders, A.J.M., Kloet, R.W., Adriaanse, S.M., Boellaard, R., Windhorst, A.D., Raijmakers, P., Adams, H., Lammertsma, A. a, Scheltens, P., van der Flier, W.M., van Berckel, B.N.M., 2014. Comparison of simplified parametric methods for visual interpretation of 11C-Pittsburgh compound-B PET images. *J. Nucl. Med.* 55, 1305–1307. <http://dx.doi.org/10.2967/jnumed.114.139121>.