Test–Retest Reproducibility of ^{18}F-MPPF PET in Healthy Humans: A Reliability Study

Nicolas Costes[1,2], Luc Zimmer[1,3,4], Anthonin Reilhac[1], Franck Lavenne[1], Philippe Ryvlin[2,3], and Didier Le Bars[1]

[1]CERMEP–Imagerie du vivant, PET Department, Lyon, France; [2]INSERM, U821 University of Lyon 1, Lyon, France; [3]IFNL, Federative Institute of Neuroscience, Lyon, France; and [4]FRE CNRS 3006, University of Lyon 1, Lyon, France

The aim of this study was to assess the reliability of 2′-methoxy-phenyl-(N-2′-pyridinyl)-p-^{18}F-fluoro-benzamidoethylpiperazine (^{18}F-MPPF) PET binding parameter's quantification via a test–retest study over a long-term period. **Methods:** Ten healthy volunteers underwent 2 dynamic ^{18}F-MPPF PET scans in an interval of 6 mo. As a methodologic control, 10 simulated datasets, including interindividual functional and anatomic variabilities, were also used to assess the measurement variations in the absence of intraindividual variability. Indices of tracer binding were computed using 2 different models: (a) the simplified reference tissue model (SRTM) and (b) the Logan graphical model. The SRTM allows computing the binding potential (BP) index and plasma-to-brain transport constants (R_1, k_2). The Logan model evaluates the distribution volume (DV). For both methods, cerebellum was taken as the reference region. From both models, binding indices were calculated with time–activity curves extracted from regions of interest, on one hand, and for each voxel to perform parametric images on the other hand. **Results:** Reliability indices—that is, bias, variability, and intraclass correlation (ICC)—indicated a good reproducibility: the BP percentage change in mean between test and retest is close to 1% in rich regions and 2% in poor regions. The typical error is around 7%. Mean ICC is over 0.70. The DV percentage change in the mean is ±2.5%, with a typical error close to 6% and an ICC over 0.60. **Conclusion:** Our results show a good reliability, with a reasonable level of intraindividual biologic variability that allows crossover studies with ^{18}F-MPPF in which small percentage changes are expected between test and retest measurements, in group studies and for single subject assessment.

Key Words: PET; serotonin 5-HT$_{1A}$ receptors; test–retest study; reliability; ^{18}F-MPPF

Serotonin (5-HT) mediates a large variety of physiologic responses (development, pain, sleep, mood, eating, memory and attention), behaviors (stress, aggression, panic, sexual behavior), or neuropsychiatric problems (depression, sleep disturbance, eating disorders, anxiety, suicidal behavior, schizophrenia, obsessive–compulsive disorder) through one of the widest range of receptors known for any neurotransmitter (1). 5-HT$_{1A}$, the best-characterized subtype of currently known 5-HT receptors, is tightly involved in the pathogenesis of the previous problems and, thus, represents an important target for drug therapy (2). Recently, the selective 5-HT$_{1A}$ receptor antagonist 2′-methoxyphenyl-(N-2′-pyridinyl)-p-^{18}F-fluoro-benzamidoethylpiperazine (^{18}F-MPPF) has been successfully labeled with ^{18}F-fluorine, and an increasing number of PET studies with ^{18}F-MPPF have been performed (3). The in vivo exploration of the 5-HT$_{1A}$ receptor subtype of the 5-HT neurotransmission system with the ^{18}F-MPPF PET radiotracer has revealed significant modulations of tracer binding due to pathologic (4–6) or pharmacologic actions in humans (7), and in animals (8,9).

The ^{18}F-MPPF radiotracer has been characterized in humans in terms of selectivity (10). In addition, the modeling, studied using a 3-compartment model, confirmed that binding potential (BP) values were linearly correlated with the binding site density (B$_{max}$) and, therefore, could be considered as a reliable index of local 5-HT$_{1A}$ receptor concentrations (11). Characterization in the healthy population has been evaluated for age and sex (12). However, reproducibility and control of variability of the measurement have not been assessed yet for this radiotracer, contrary to other serotoninergic radiotracers (13–16). In several ^{18}F-MPPF studies, ^{18}F-MPPF binding was measured at consecutive sessions and compared within each subject. In this way, issues associated with variations between subjects due to individual differences (intersubject variability) can be limited. However, resting physiology may vary within an individual patient (intrasubject variability) and limits the ability to detect significant changes between baseline and "postintervention" conditions. Moreover, bias and noise introduced by the data acquisition, the reconstruction, and the correction processes, and also the simplified model used for binding parameter estimation (for the 5-HT$_{1A}$ radiotracer ^{11}C-WAY-100635 (14,15)), must be accounted for.

Therefore, without the estimation of the test–retest reproducibility; it is difficult to accurately determine the clinical significance of pharmacologic or pathophysiologic changes of the 5-HT$_{1A}$ receptor status. In this context,

reliability studies are crucial when one wants to account for the error term in measurements. Because the test–retest reliability study has never been performed for ^{18}F-MPPF, to our knowledge, the objective of our work was to perform a reliability study to evaluate reproducibility and measurement error of ^{18}F-MPPF PET binding measures over a 6-mo period.

MATERIALS AND METHODS

Subjects

Ten healthy volunteers (5 females, 5 males; mean age ± SD, 30 ± 5 y; age range, 23–38 y) were selected to participate in the study. Subjects gave their written consent to participate in the study, which was approved by the local ethical committee (Centre Léon Bérard, Lyon, France) in accordance with the Declaration of Helsinki. According to a screening assessment of history and physical examination, all subjects were free of neurologic, psychiatric, cardiovascular, pleuropulmonary, or hematologic disease and did not meet any exclusion criteria: (a) neuroleptic, antiparkinsonian α-methyl-dopa, β-blocker, monoamine oxidase A or B inhibitor, tricyclic antidepressant, or thymoregulator treatment; (b) pregnancy; (c) hormone replacement therapy; (d) consumption of recreational drugs (cannabis, ecstasy); (e) contraindication to MRI; (f) MRI detection of brain lesion. Before the PET scan, subjects were evaluated for depression using the General Health Questionnaire (GHQ-28) of Goldberg and Hillier (17). No subject with a score above the threshold for depression (7) was included in the study. Scores ranged between 0 and 2 (mean ± SD, 0.3 ± 0.7). No significant difference in the GHQ was found between both scans (Table 1).

MRI

A 3-dimensional multiplanar reconstruction anatomic MRI scan was performed on each subject, yielding a volume containing 130–170 transverse planes of $256 \times 256 \times 1$ mm^3 voxels.

PET Scan with ^{18}F-MPPF

Tracer Synthesis. The ^{18}F-MPPF was obtained by nucleophilic fluoration on a nitro precursor with a radiochemical yield of 20%–25% at the end of synthesis and a specific activity of 37–111 GBq/μmol (18,19).

Scanning Procedure. Subjects underwent 2 ^{18}F-MPPF PET scans (test and retest) separated by a 6-mo period (mean delay ± SD, 27 ± 2 wk), randomly distributed along the year. Each PET session began at 1 PM. The PET scan acquisition, correction, and reconstruction procedures followed those described in (12). PET scans were obtained with a CTI Exact HR+ camera for 60 min after the injection of 2.7 MBq/kg (mean total dose ± SD, 169 ± 30 MBq) of ^{18}F-MPPF. There was no significant difference ($P = 0.93$) in the paired t test between the injected dose of the test and retest scans.

Simulated Data. Following the methodology for the simulation of realistic PET described in (20), we performed a joint simulation of a test–retest study. Ten individual different numeric brains associated with 10 different sets of regional time–activity curves were used to simulate 10 realizations of ^{18}F-MPPF PET dynamic acquisition. The simulations were repeated to obtain the retest simulated data. The input time–activity curves used for the test and retest simulated acquisition set were identical for each subject. Therefore, the only difference between the simulated test and retest datasets was due to the degradation induced by the physical acquisition processes. Thus, physical variability can be compared with the measured one estimated from the actual data that include all sources of variability.

Image Processing

Modeling. The binding parameters of the tracer were estimated according to the 3-compartment simplified reference tissue model (SRTM) (21) and with the Logan graphical method (22). The SRTM method lies on an analytic solution of the compartment model (Eq. 1) and allows estimating 3 indices—R_1, k_2, and BP—without requirement of an arterial sampling input function. The SRTM works under several assumptions: (a) the existence of a reference tissue region with negligible concentration of specific binding sites, (b) the magnitude of nonspecific binding is the same in the reference and in the target regions, (c) the distribution volumes (DVs) in the free and nonspecific compartments are the same in the reference and in the target regions, and (d) the exchanges between the free and unspecific binding compartments are rapid.

The analytic solution of the partial derivatives system is of the following form:

$$C_{roi}(t) = R_1 \cdot C_{ref}(t) + \left[k_2 - \frac{R_1 k_2}{1 + BP}\right] \cdot C_{ref}(t) \otimes e^{\frac{k_2}{1+BP}t}, \quad \text{Eq. 1}$$

where C_{ref} and C_{roi} are the PET time–activity curves of the chosen reference region and the target region of interest (ROI), R_1 is the ratio of the plasma-to-brain transport constant in the target region and in the reference region ($R_1 = k_{1roi}/k_{1ref}$), k_2 is the tracer's efflux in the vascular system, and BP is the binding potential of the tracer, defined as the ratio of available receptor density to the receptor affinity ($BP = B_{max}/K_d$). For the ^{18}F-MPPF, the cerebellum was taken as the reference region as it is considered devoid of binding sites (11). The Logan model is based on the parametric plot involving the PET time–activity curve of the reference region and the target region. As for the SRTM, the cerebellum was taken as the reference region. The Logan model provides the DV of the target ROI, given by the slope of the linear regression after equilibrium.

Two approaches for binding index reliability analysis were used: The first is based on mean ROI activity curve measurements (ROI Analysis) and the second is based on a voxelwise computation leading to statistical parametric maps (Parametric Image Analysis).

TABLE 1
Demographic Data

Subject no.	Age (y)	Sex	Test dose (MBq)	Retest dose (MBq)	Delay (wk)
1	30	F	167	136	26
2	26	F	149	143	30
3	24	F	146	139	27
4	38	F	148	138	28
5	31	F	151	164	23
6	23	M	194	197	28
7	27	M	161	178	27
8	32	M	180	176	29
9	35	M	190	190	25
10	38	M	204	225	25
Mean ± SD	30 ± 5		169 ± 22	169 ± 30	27 ± 2

For reorientation and registration purposes, mean ^{18}F-MPPF static images were created by summing the individual dynamic frames from 0 to 60 min after injection. The MR image was coregistered with the static ^{18}F-MPPF image by an automated method using Mutual Information criteria (Statistical Parametric Mapping [SPM], Welcome Department of Cognitive Neurology, London, U.K.). On the coregistered MRI, a large ROI was outlined on the cerebellum and used as a unique reference region for the simplified models.

ROI Analysis. The anatomic target ROIs were drawn manually on the coregistered MRI. Four hundred ROIs were drawn and regrouped into anatomic volumes of interest (VOIs) to describe a group of regions from the limbic system, known to be rich in 5-HT$_{1A}$ receptors—that is, left and right hippocampi, amygdala, enthorinal cortex, parahippocampal gyri, anterior and posterior cingulate gyri, insula, temporal poles, and temporal cortex—and a second group of other cortical regions: left and right temporal neocortex, lateral occipitotemporal gyrus, frontal gyrus, prefrontal cortex, inferior and superior parietal cortices, occipital cortex, pole and gyrus, cerebellum. As raphe nuclei cannot be identified on MRI, this region was outlined directly on the static ^{18}F-MPPF image by thresholding the activity at 80% of the local maximum in the brain stem. This region was visualized a posteriori on the MRI to check for its proper location in the periacqueductal gray matter. Time–activity curves were measured from the dynamic PET using the set of ROIs. The measured time–activity curves were used to derive regional values of R_1, k_2, and BP and Logan DV for each ROI.

Parametric Image Analysis. From individual voxel time–activity curves, a parametric image of R$_1$, k$_2$, and BP was computed for the SRTM, and a parametric image of DV was computed for the Logan model. Individual parametric images were then transformed into a standard space using the nonlinear transformation matrix derived from the spatial normalization of the individual's MR image to the T1 MRI default template (Montreal Neurological Institute template of the International Consortium for Brain Mapping Project) with SPM. The visual inspection of the spatially registered images, particularly for subcortical structures, confirmed the accuracy of the spatial normalization. Normalized parametric images were smoothed using an $8 \times 8 \times 8$ mm^3 full width at half maximum isotropic gaussian kernel to account for the interindividual anatomy variability and to improve the sensitivity of the statistical analysis.

Reproducibility Indices. Reliability of the R$_1$, k$_2$, and BP indices issued from the SRTM, and the DV index issued from Logan model, were assessed by computation of 3 characteristic parameters from the test–retest measurements:

- The percentage change in mean (bias): the percentage change calculated as the difference between test and retest values divided by the test value. This index includes random changes and systematic biologic error;
- The typical error or within-subject SD of the bias: We expressed the typical error as the percentage of the mean;
- The intraclass correlation coefficient (ICC) estimates the respect of the rank in a test–retest study: It depends on the size and the quality of the sample in the population. ICC = (MSBS − MSWS)/(MSBS + MSWS), where MSBS is the mean sum of squares between subjects, and MSWS is the mean sum of squares within subjects.

Because these variables have residuals that may be proportional to their respective mean, their computation is performed from the logarithmic transform of the variables as suggested in (*23*).

Statistical Inference. For the ROI analysis, mean regional binding index values were considered as independent measures.

For the voxel-based analysis, SPM99 was used on the normalized smoothed parametric images of the 10 subjects. Statistical parametric maps of the *t* statistic (SPM{*t*}) were computed with a threshold of $P = 0.001$ uncorrected at the voxel level. Significant clusters were selected at a corrected cluster level of $P < 0.05$ determined from a joint probability of peak height and cluster size ("Family-Wise Error" (*24*)).

RESULTS

SRTM

Actual BP. As shown in Table 2 (top) test and retest, mean regional BP values range from 0.28 ± 0.08 in the raphe to 1.47 ± 0.16 in the hippocampus. The percentage changes between test and retest values ranges from -1.15% (anterior cingulum) to 4.80% (enthorinal cortex), with a mean typical error of 7.75% in the limbic area and 7.71% in other regions. The maximal typical error is in the raphe nucleus (14.97%) and the minimal error is in the parahippocampal gryus (4.67%). The ICC goes from 0.50 (anterior cingulum) to 0.93 (inferior parietal cortex). Mean ICC values are 0.69 in the limbic area and 0.84 in other cortical regions. None of the regional BP differences was found to be statistically significant with a paired *t* test comparison.

Actual R$_1$ (Table 2, Middle). Reproducibility of the relative perfusion parameter (R$_1$) in the ROIs is also excellent, with an average percentage changes of 1.63% in the limbic area (from -5.03% in the temporal pole to 5.21% in the hippocampus), and a mean typical error of 5.77%. The typical error of the R$_1$ parameter is generally inferior to the error of the BP parameter, going from 4.60% (raphe) to 7.86% (hippocampus). The average change is -3.43% in the other cortical regions, with a mean typical error of 5.98%. The ICC revealed very puzzling values, from -0.17 in the prefrontal cortex to 0.67 in the temporal cortex. The mean ICC is 0.37 in the limbic areas and 0.25 in the other cortical regions.

Actual k$_2$ (Table 2, Bottom). The k$_2$ values go from 0.11 ± 0.01 min^{-1} in the raphe, to 0.35 ± 0.05 min^{-1} in the occipital gyrus. The percentage changes in the mean between test and retest series go from -4.42% in the enthorinal cortex to $+2.25\%$ in the prefrontal cortex. In the limbic areas, the average percentage change in mean was -0.50%, very similar to that of the other regions (-0.51%). The mean typical error is 9.50% in the limbic regions and 10.53% in the other cortical regions. The mean ICC value is 0.40, with a large range from -0.03 in the anterior cingulated cortex to 0.77 in the parahippocampal gyrus. In other regions, the mean ICC value is 0.31, with a range from 0.15 in the occipital cortex to 0.42 in the occipital gyrus.

Simulated Data. Simulated data have an excellent ICC (>0.95), a mean percentage change in BP values around

TABLE 2
Test–Retest Characteristics of BP, Delivery in ROI Relative to Cerebellum (R_1), and Tracer Efflux Constant from Tissue to Plasma (k_2) Computed with SRTM with Actual ^{18}F-MPPF Data of 10 Subjects

Actual SRTM BP Region	Test (mean ± SD)	Retest (mean ± SD)	Change in mean (%)	Typical error as CV (%)	ICC	Paired t test P
Hippocampus	1.43 ± 0.17	1.47 ± 0.16	2.48	7.30	0.57	0.426
Amygdala	1.16 ± 0.15	1.21 ± 0.16	4.04	8.68	0.57	0.270
Enthorinal cortex	1.25 ± 0.19	1.30 ± 0.16	4.80	7.43	0.73	0.188
Parhippocampal gyrus	0.90 ± 0.11	0.91 ± 0.10	1.67	4.67	0.84	0.474
Anterior cingulum	0.75 ± 0.09	0.74 ± 0.09	−1.15	9.08	0.50	0.771
Posterior cingulum	0.65 ± 0.07	0.65 ± 0.07	−0.47	6.12	0.67	0.876
Insula	0.96 ± 0.10	0.97 ± 0.11	−0.68	6.86	0.67	0.818
Temporal pole	0.91 ± 0.14	0.91 ± 0.12	−0.46	7.10	0.77	0.994
Temporal cortex	0.90 ± 0.11	0.90 ± 0.11	−0.02	5.32	0.81	0.986
Raphe	0.28 ± 0.08	0.28 ± 0.09	−0.50	14.97	0.78	0.944
Mean limbic area	**0.92 ± 0.12**	**0.93 ± 0.12**	**0.97**	**7.75**	**0.69**	
Temporal neocortex	0.85 ± 0.10	0.84 ± 0.11	1.54	5.43	0.80	0.565
Lateral occipitotemporal gyrus	0.80 ± 0.13	0.77 ± 0.09	3.90	6.47	0.79	0.155
Frontal gyrus	0.61 ± 0.11	0.60 ± 0.11	1.85	7.36	0.83	0.564
Prefrontal cortex	0.53 ± 0.09	0.52 ± 0.11	1.68	7.14	0.84	0.703
Inferior parietal cortex	0.61 ± 0.13	0.59 ± 0.11	2.65	5.47	0.93	0.179
Superior parietal cortex	0.57 ± 0.10	0.55 ± 0.08	2.09	7.80	0.78	0.451
Occipital cortex	0.61 ± 0.12	0.60 ± 0.11	1.94	6.31	0.88	0.485
Occipital pole	0.40 ± 0.15	0.39 ± 0.16	1.56	14.30	0.87	0.900
Occipital gyrus	0.55 ± 0.11	0.53 ± 0.11	2.20	9.11	0.82	0.548
Mean other area	**0.61 ± 0.12**	**0.60 ± 0.11**	**2.16**	**7.71**	**0.84**	

Actual SRTM R_1 Region	Test (mean ± SD)	Retest (mean ± SD)	Change in mean (%)	Typical error as CV (%)	ICC	Paired t test P
Hippocampus	0.79 ± 0.05	0.83 ± 0.09	5.21	7.86	0.13	0.175
Amygdala	0.73 ± 0.04	0.75 ± 0.06	3.00	6.36	0.23	0.309
Enthorinal cortex	0.76 ± 0.05	0.80 ± 0.06	4.18	4.96	0.35	0.090
Parhippocampal gyrus	0.82 ± 0.06	0.86 ± 0.09	4.87	5.96	0.54	0.097
Anterior cingulum	0.92 ± 0.07	0.96 ± 0.07	4.35	6.88	0.24	0.204
Posterior cingulum	0.94 ± 0.06	0.96 ± 0.07	2.83	4.87	0.53	0.228
Insula	0.97 ± 0.04	1.00 ± 0.09	−2.59	5.31	0.36	0.274
Temporal pole	0.73 ± 0.07	0.77 ± 0.07	−5.03	5.52	0.67	0.063
Temporal cortex	0.80 ± 0.07	0.83 ± 0.07	−3.74	5.42	0.59	0.161
Raphe	0.31 ± 0.01	0.32 ± 0.02	3.19	4.60	0.03	0.153
Mean limbic area	**0.78 ± 0.05**	**0.81 ± 0.07**	**1.63**	**5.77**	**0.37**	
Temporal neocortex	0.91 ± 0.06	0.93 ± 0.07	−2.07	5.03	0.48	0.377
Lateral occipitotemporal gyrus	0.91 ± 0.06	0.95 ± 0.08	−3.63	5.80	0.45	0.176
Frontal gyrus	0.91 ± 0.05	0.93 ± 0.05	−2.65	5.73	0.05	0.324
Prefrontal cortex	0.91 ± 0.05	0.93 ± 0.05	−2.75	6.19	−0.17	0.327
Inferior parietal cortex	0.91 ± 0.04	0.94 ± 0.07	−2.83	6.34	−0.01	0.321
Superior parietal cortex	0.95 ± 0.04	0.97 ± 0.06	−2.54	5.56	−0.02	0.307
Occipital cortex	0.86 ± 0.08	0.90 ± 0.07	−4.66	6.35	0.46	0.126
Occipital pole	0.79 ± 0.08	0.83 ± 0.09	−4.74	7.24	0.51	0.163
Occipital gyrus	1.02 ± 0.07	1.08 ± 0.10	−5.00	5.55	0.52	0.064
Mean other area	**0.91 ± 0.06**	**0.94 ± 0.07**	**−3.43**	**5.98**	**0.25**	

Actual SRTM k_2 Region	Test (mean ± SD)	Retest (mean ± SD)	Change in mean (%)	Typical error as CV (%)	ICC	Paired t test P
Hippocampus	0.25 ± 0.04	0.24 ± 0.02	−2.83	9.66	0.45	0.417
Amygdala	0.25 ± 0.04	0.25 ± 0.02	−1.26	11.58	0.21	0.648
Enthorinal cortex	0.23 ± 0.03	0.22 ± 0.02	−4.42	9.64	0.37	0.250
Parhippocampal gyrus	0.27 ± 0.04	0.27 ± 0.04	1.69	7.25	0.77	0.707
Anterior cingulum	0.32 ± 0.04	0.31 ± 0.03	−2.88	13.36	−0.03	0.574
Posterior cingulum	0.30 ± 0.03	0.30 ± 0.04	0.24	8.69	0.47	0.929
Insula	0.29 ± 0.04	0.28 ± 0.03	1.22	9.48	0.46	0.713
Temporal pole	0.24 ± 0.03	0.24 ± 0.02	0.14	7.67	0.55	0.923
Temporal cortex	0.27 ± 0.03	0.26 ± 0.03	1.21	10.00	0.35	0.763
Raphe	0.11 ± 0.01	0.11 ± 0.01	1.90	7.66	0.41	0.573
Mean limbic area	**0.25 ± 0.03**	**0.25 ± 0.03**	**−0.50**	**9.50**	**0.40**	

TABLE 2
(Continued)

Actual SRTM k_2 Region	Test (mean ± SD)	Retest (mean ± SD)	Change in mean (%)	Typical error as CV (%)	ICC	Paired t test P
Temporal neocortex	0.27 ± 0.04	0.27 ± 0.03	0.64	10.09	0.40	0.832
Lateral occipitotemporal gyrus	0.31 ± 0.04	0.32 ± 0.03	−3.81	9.05	0.36	0.387
Frontal gyrus	0.29 ± 0.03	0.29 ± 0.03	−0.02	8.77	0.23	0.977
Prefrontal cortex	0.29 ± 0.03	0.29 ± 0.03	−0.35	9.65	0.33	0.939
Inferior parietal cortex	0.30 ± 0.04	0.29 ± 0.04	2.25	12.18	0.20	0.699
Superior parietal cortex	0.32 ± 0.04	0.32 ± 0.04	0.01	10.43	0.35	0.991
Occipital cortex	0.29 ± 0.04	0.29 ± 0.03	−0.09	11.66	0.15	0.988
Occipital pole	0.27 ± 0.04	0.27 ± 0.05	1.14	13.34	0.32	0.907
Occipital gyrus	0.34 ± 0.04	0.35 ± 0.05	−4.39	9.58	0.42	0.282
Mean other area	**0.30 ± 0.04**	**0.30 ± 0.04**	**−0.51**	**10.53**	**0.31**	

−1%, with a mean typical error of 2.5% (Table 3, top). The R_1 parameter shows a percentage change in the mean around zero and a typical error of 1.79% in the limbic areas and 0.85% in the other regions (Table 3, middle). The k_2 of simulated data has a percentage change in the mean of <1% and a typical error of <3% (Table 3, bottom).

Logan Model

The DV in the ROIs of test and retest series goes from 0.56 ± 0.07 in the raphe nucleus to 2.45 ± 0.24 in the hippocampus (Table 4). No difference was found between test and retest paired t test comparisons. The mean percentage change by region goes from −4.37% in the occipital pole to +8.05% in the amygdala, with a range of typical error from 3.45% in the posterior cingulum to >12.85% in the amygdala. The mean typical error is at 6.48% in the limbic area and 5.65% in the other cortical regions. ICC values range from 0.42 in the amygdala to 0.88 in the posterior cingulate gyrus, with a mean value of 0.66 in the limbic area and 0.74 in the other regions.

SPM

SPM results did not show any significant difference between test and retest scan series, in terms of variance and the mean difference with a paired t test model.

DISCUSSION

This test–retest reliability study of ^{18}F-MPPF binding was designed to support interpretation of clinical studies implying a long delay between the first PET scan and the second. The obtained pairs of ^{18}F-MPPF PET images were apparently identical by visual inspection, as exemplified in Figure 1. This similarity includes experimental conditions, such as reproducibility of head positioning, and injected radioactivity. Nineteen regions were studied and their binding index was calculated. This reliability of PET ^{18}F-MPPF binding index, assessed by a long-term test–retest acquisition procedure, has shown high reproducibility. This reliability study gives the precision of the measurement and the ability to test differences between measurements with PET ^{18}F-MPPF.

Methodologic Considerations

Bias and Typical Error. The mean percentage changes for the BP are inferior in the limbic areas (<1%) to that in the other cortical regions (>2%); however, the typical errors are similar (around 7%). The simulated data predicted a mean percentage change of around 1%, with a typical error of 2%. Thus, we can conclude in favor of a better stability of the test–retest measurement close to the ideal—with a biologic uncertainty equivalent in regions rich and poor in 5-HT$_{1A}$ receptors—larger than the simple measurement error due to the PET image formation process. The R_1 parameter has a variability around ±5% in rich regions, which systematically increases (values from 2% to 5%) in poor regions. For the 2 classes of regions, the typical error is close to 6%. The simulation study indicates that the reproducibility should be near zero and the typical error should be between 1% and 2%. Because R_1 is related to cerebral blood flow, we can state the hypothesis that the regional cerebral blood flow was greater during the retest scan than during the first scan. Because this finding is not observable in the BP results, we can state that modeled parameters are effectively identified independently. Finally, the k_2 parameter has a very good stability (1% in the mean) for a typical error of measurement close to 10%, whereas simulation predicted a 3% typical error in the absence of intraindividual variability. In conclusion with the SRTM, the R_1 is the estimated parameter that presents the lowest measurement error (5%), followed by the BP (7%) and by the k_2 (10%). The Logan model shows a reliability result similar to the BP reliability. Precisely, DV has a higher variability but a lower typical error than the BP.

Variability. The ICC is a measure of the correlation between the values obtained with 2 methods within the same subject (*25*). It is used as an index of reliability of the test–retest measurements and combines information of the systemic difference between methods (test and retest) and of the measurement variations. In a PET study on ^{11}C-WAY 100635, ICC values above 0.50 and 0.75 were considered as acceptable and excellent, respectively (*16*). In our study, the ICC values of BP and DV were poor (<0.5) to excellent (>0.75). On average, the ICC is slightly inferior in the

TABLE 3

Test–Retest Characteristics of BP, Delivery in ROI Relative to Cerebellum (R_1), and Tracer Efflux Constant from Tissue to Plasma (k_2) Computed with SRTM with Simulated ^{18}F-MPPF Data of 10 Subjects

Simulated SRTM BP Region	Test (mean ± SD)	Retest (mean ± SD)	Change in mean (%)	Typical error as CV (%)	ICC	Paired t test P
Hippocampus	0.86 ± 0.15	0.87 ± 0.14	−1.21	1.99	0.99	0.189
Amygdala	0.71 ± 0.16	0.73 ± 0.15	−2.90	3.52	0.97	0.135
Enthorinal cortex	0.63 ± 0.15	0.63 ± 0.15	−0.26	2.34	0.99	0.938
Parhippocampal gyrus	0.53 ± 0.13	0.54 ± 0.13	−1.51	1.79	1.00	0.108
Anterior cingulum	0.48 ± 0.13	0.49 ± 0.13	−0.22	3.21	0.99	0.690
Posterior cingulum	0.45 ± 0.15	0.45 ± 0.14	−1.00	1.83	1.00	0.396
Insula	0.68 ± 0.17	0.69 ± 0.17	−1.30	1.11	1.00	0.043
Temporal pole	0.49 ± 0.15	0.50 ± 0.15	−0.42	1.99	1.00	0.898
Temporal cortex	0.57 ± 0.15	0.57 ± 0.15	−0.32	1.24	1.00	0.759
Raphe	0.09 ± 0.04	0.09 ± 0.04	−1.37	1.98	1.00	0.415
Mean limbic area	**0.55 ± 0.14**	**0.55 ± 0.13**	**−1.05**	**2.10**	**0.99**	
Temporal neocortex	0.57 ± 0.14	0.57 ± 0.13	0.06	1.13	1.00	0.810
Lateral occipitotemporal gyrus	0.53 ± 0.16	0.53 ± 0.15	−0.81	1.61	1.00	0.492
Frontal gyrus	0.30 ± 0.11	0.30 ± 0.11	−1.32	2.28	1.00	0.360
Prefrontal cortex	0.24 ± 0.13	0.24 ± 0.12	−3.12	6.26	0.99	0.542
Inferior parietal cortex	0.47 ± 0.14	0.46 ± 0.14	0.20	1.02	1.00	0.449
Superior parietal cortex	0.36 ± 0.15	0.36 ± 0.14	−0.97	2.80	1.00	0.798
Occipital cortex	0.44 ± 0.11	0.44 ± 0.10	−0.32	2.81	0.99	0.932
Occipital pole	0.24 ± 0.10	0.25 ± 0.09	−2.65	3.17	1.00	0.179
Occipital gyrus	0.42 ± 0.12	0.42 ± 0.12	−0.57	1.49	1.00	0.595
Mean other area	**0.40 ± 0.13**	**0.40 ± 0.12**	**−1.06**	**2.51**	**1.00**	

Simulated SRTM R_1 Region	Test (mean ± SD)	Retest (mean ± SD)	Change in mean (%)	Typical error as CV (%)	ICC	Paired t test P
Hippocampus	0.80 ± 0.07	0.80 ± 0.07	−0.37	2.17	0.95	0.716
Amygdala	0.73 ± 0.06	0.73 ± 0.07	−0.51	3.90	0.83	0.833
Enthorinal cortex	0.70 ± 0.07	0.68 ± 0.08	−2.58	1.93	0.97	0.022
Parhippocampal gyrus	0.75 ± 0.08	0.75 ± 0.09	−0.17	2.00	0.97	0.993
Anterior cingulum	0.80 ± 0.12	0.80 ± 0.12	−0.13	2.58	0.97	0.922
Posterior cingulum	0.72 ± 0.13	0.72 ± 0.13	0.07	1.12	1.00	0.755
Insula	0.84 ± 0.10	0.85 ± 0.11	−0.78	1.44	0.99	0.179
Temporal pole	0.63 ± 0.08	0.63 ± 0.08	−0.11	1.58	0.99	0.802
Temporal cortex	0.75 ± 0.09	0.75 ± 0.10	0.16	0.60	1.00	0.730
Raphe	0.26 ± 0.03	0.26 ± 0.03	−0.10	0.55	1.00	0.860
Mean limbic area	**0.70 ± 0.08**	**0.70 ± 0.09**	**−0.45**	**1.79**	**0.97**	
Temporal neocortex	0.74 ± 0.09	0.74 ± 0.10	0.22	0.75	1.00	0.605
Lateral occipitotemporal gyrus	0.92 ± 0.11	0.92 ± 0.11	−0.13	0.61	1.00	0.514
Frontal gyrus	0.83 ± 0.11	0.83 ± 0.11	0.18	0.54	1.00	0.581
Prefrontal cortex	0.64 ± 0.10	0.64 ± 0.10	−0.24	0.77	1.00	0.446
Inferior parietal cortex	0.74 ± 0.11	0.74 ± 0.12	0.33	1.11	0.99	0.518
Superior parietal cortex	0.87 ± 0.13	0.87 ± 0.13	0.32	0.57	1.00	0.268
Occipital cortex	0.75 ± 0.10	0.74 ± 0.10	1.00	1.38	0.99	0.161
Occipital pole	0.86 ± 0.13	0.86 ± 0.13	−0.14	1.30	0.99	0.921
Occipital gyrus	1.07 ± 0.15	1.07 ± 0.15	0.09	0.56	1.00	0.878
Mean other area	**0.82 ± 0.11**	**0.82 ± 0.09**	**0.18**	**0.85**	**1.00**	

Simulated SRTM k_2 Region	Test (mean ± SD)	Retest (mean ± SD)	Change in mean (%)	Typical error as CV (%)	ICC	Paired t test P
Hippocampus	0.21 ± 0.04	0.21 ± 0.04	−0.06	3.33	0.97	0.761
Amygdala	0.20 ± 0.04	0.21 ± 0.04	1.95	5.47	0.94	0.352
Enthorinal cortex	0.18 ± 0.03	0.18 ± 0.04	2.15	3.02	0.98	0.106
Parhippocampal gyrus	0.19 ± 0.04	0.18 ± 0.04	−0.95	2.55	0.99	0.289
Anterior cingulum	0.25 ± 0.04	0.24 ± 0.04	−3.06	5.38	0.90	0.199
Posterior cingulum	0.26 ± 0.05	0.26 ± 0.05	−0.53	1.66	0.99	0.493
Insula	0.29 ± 0.06	0.29 ± 0.06	0.84	1.77	0.99	0.290
Temporal pole	0.14 ± 0.03	0.14 ± 0.03	0.36	2.66	0.99	0.791
Temporal cortex	0.18 ± 0.04	0.17 ± 0.04	0.46	1.18	1.00	0.479
Raphe	0.05 ± 0.01	0.05 ± 0.01	0.20	1.17	0.99	0.668
Mean limbic area	**0.19 ± 0.04**	**0.19 ± 0.04**	**0.14**	**2.82**	**0.97**	

TABLE 3
(Continued)

Simulated SRTM k_2 Region	Test (mean ± SD)	Retest (mean ± SD)	Change in mean (%)	Typical error as CV (%)	ICC	Paired t test P
Temporal neocortex	0.25 ± 0.05	0.25 ± 0.05	0.46	1.09	1.00	0.311
Lateral occipitotemporal gyrus	0.18 ± 0.04	0.18 ± 0.04	1.45	2.48	0.99	0.224
Frontal gyrus	0.16 ± 0.03	0.16 ± 0.03	0.26	1.48	0.99	0.697
Prefrontal cortex	0.20 ± 0.04	0.20 ± 0.04	0.30	1.40	1.00	0.490
Inferior parietal cortex	0.26 ± 0.05	0.26 ± 0.05	0.07	1.84	0.99	0.985
Superior parietal cortex	0.18 ± 0.05	0.18 ± 0.04	0.49	2.10	0.99	0.690
Occipital cortex	0.23 ± 0.04	0.24 ± 0.04	−0.89	2.59	0.98	0.471
Occipital pole	0.11 ± 0.02	0.11 ± 0.02	−1.97	5.49	0.94	0.551
Occipital gyrus	0.13 ± 0.03	0.13 ± 0.02	0.35	3.73	0.96	0.711
Mean other area	**0.19 ± 0.04**	**0.19 ± 0.04**	**0.06**	**2.47**	**0.98**	

limbic area than in the other cortical regions. Because this parameter is representative of the individual stability between test and retest scans, it appeared that it was more variable in regions with high 5-HT$_{1A}$ receptor densities than in poor regions. The SDs of the value in test–retest are similar in the limbic area and in the other regions (0.12 for the BP), so the difference in the ICC is due to a higher intrasubject variability in rich regions than in poor regions.

This phenomenon must be considered when individual test–retest results are examined, but it does not affect the reproducibility of a group comparison according to Parsey et al. (2000) and Hirvonen et al. (2006) with ^{11}C-WAY 100635. Many studies reported moderate ICC values: ICC values are higher with the SRTM than with graphical or nonlinear fitting techniques with peripherical arterial blood function (14–16,26,27).

Physiologic Considerations

The 5-HT$_{1A}$ receptors are implicated in a range of behaviors and in many neuropsychiatric and neurodegenerative diseases. It explains why an increasing number of academic and industrial centers use the selective 5-HT$_{1A}$ receptor antagonist, ^{18}F-MPPF, as the radiotracer in PET clinical studies (3). It must be noted that several ^{18}F-MPPF PET studies have been designed either as a group's comparison or as repeated measures in the same individuals. Repeated measures in the same individuals are performed (a) when addressing drug effects in occupancy studies, (b) when

TABLE 4
Test–Retest Characteristics of Tracer DV Computed with Logan Model with Actual ^{18}F-MPPF Data of 10 Subjects

| Actual Logan + DV Region	Test (mean ± SD)	Retest (mean ± SD)	Change in mean (%)	Typical error as CV (%)	ICC	Paired t test P
Hippocampus	2.38 ± 0.23	2.45 ± 0.24	3.01	7.11	0.44	0.344
Amygdala	2.02 ± 0.32	2.17 ± 0.30	8.05	12.85	0.42	0.143
Enthorinal cortex	2.10 ± 0.24	2.23 ± 0.22	6.35	6.53	0.67	0.052
Parhippocampal gyrus	1.78 ± 0.14	1.84 ± 0.16	3.55	5.00	0.60	0.143
Anterior cingulum	1.68 ± 0.22	1.77 ± 0.20	5.12	6.68	0.72	0.119
Posterior cingulum	1.58 ± 0.17	1.61 ± 0.16	1.72	3.45	0.88	0.312
Insula	1.97 ± 0.16	2.04 ± 0.18	−3.38	5.04	0.66	0.147
Temporal pole	1.71 ± 0.14	1.77 ± 0.14	−3.42	4.20	0.75	0.103
Temporal cortex	1.79 ± 0.21	1.85 ± 0.18	−3.25	4.69	0.80	0.162
Raphe	0.56 ± 0.07	0.59 ± 0.10	5.71	9.33	0.63	0.193
Mean limbic area	**1.76 ± 0.19**	**1.83 ± 0.19**	**2.34**	**6.48**	**0.66**	
Temporal neocortex	1.81 ± 0.18	1.86 ± 0.17	−2.59	4.97	0.71	0.273
Lateral occipitotemporal gyrus	1.68 ± 0.16	1.69 ± 0.13	−0.45	6.12	0.47	0.905
Frontal gyrus	1.54 ± 0.19	1.60 ± 0.17	−3.64	5.26	0.80	0.143
Prefrontal cortex	1.48 ± 0.19	1.52 ± 0.16	−3.19	5.03	0.82	0.193
Inferior parietal cortex	1.58 ± 0.18	1.62 ± 0.15	−2.47	4.75	0.78	0.296
Superior parietal cortex	1.55 ± 0.17	1.58 ± 0.13	−2.07	4.24	0.78	0.333
Occipital cortex	1.55 ± 0.19	1.60 ± 0.17	−3.52	4.78	0.83	0.147
Occipital pole	1.12 ± 0.19	1.19 ± 0.27	−4.37	10.47	0.79	0.255
Occipital gyrus	1.49 ± 0.13	1.53 ± 0.16	−1.94	5.27	0.71	0.399
Mean other area	**1.53 ± 0.17**	**1.58 ± 0.17**	**−2.69**	**5.65**	**0.74**	
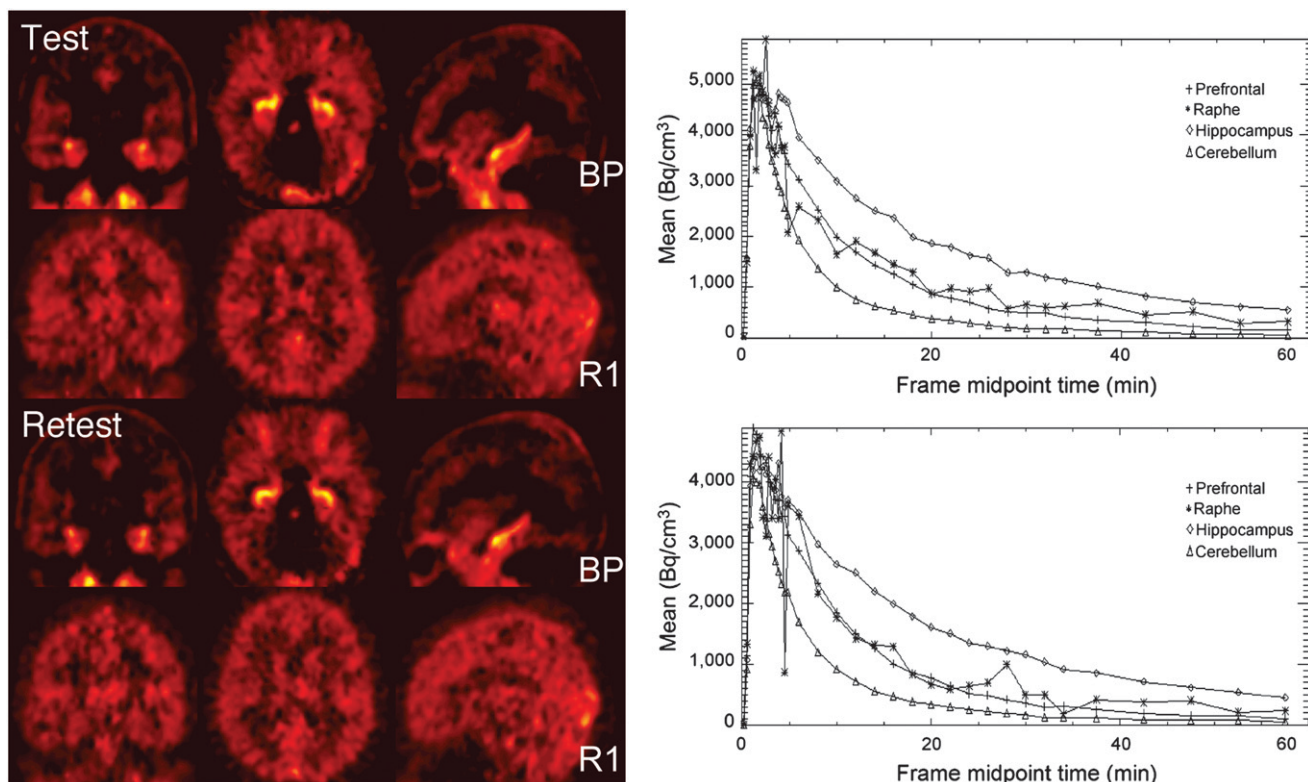
FIGURE 1. (Left) Example of an individual SRTM parametric BP (binding potential) and R_1 (delivery in region of interest relative to cerebellum) images in test and in retest conditions. (Right) Example of time–activity curves of 1 subject after injection of 148 MBq (test) and 138 MBq (retest) of ^{18}F-MPPF. For clarity, only 4 of the 19 regions are included (hippocampus, raphe, prefrontal cortex, and reference cerebellum).

after a disease condition over time, and (c) when measuring variability in receptor densities. The time interval between these measurements may well be several weeks or months. Therefore, interpretation of these studies requires understanding of the test–retest reliability of the methodology, particularly when the degree of change is subtle. This is important because 5-HT$_{1A}$ receptor availability can be physiologically or pharmacologically modified and, therefore, could lead to a modification of the apparent binding of ^{18}F-MPPF. For example, preclinical studies suggested that ^{18}F-MPPF was sensitive to 5-HT, because of its affinity close to endogenous 5-HT (*3*). Because it is known that increases in 5-HT release are largely due to several physiologic, environmental, and behavioral manipulations (*28*), it can be hypothesized that the ^{18}F-MPPF binding is reduced when the 5-HT concentration is increased. Recently, a ^{18}F-MPPF study suggested that 5-HT$_{1A}$ receptor availability was increased during sleep (*4*). Other physiologic processes, such as aging, could have a direct impact on the 5-HT$_{1A}$ receptor density (*29*). Therefore, a long period between the test and retest scans could lead to subtle but significant intraindividual variabilities. Finally, as recently described, 5-HT$_{1A}$ receptors in the raphe dorsalis can be partially internalized, leading to a considerable decrease in ^{18}F-MPPF binding (>30%) (*9,30*). It must be noted, that although these multiple factors could theoretically modify the BP of ^{18}F-MPPF,

our results showed that the biologic variability was moderate (<10%). Furthermore, the paired *t* test between test and retest scans in the ROI analysis and in the SPM analysis did not showed significant changes. These results revealed a strong stability of individual measurements and, therefore, opportunity to use ^{18}F-MPPF tracer for testing longitudinal clinical evolutions over periods of a few months.

Use of Reliability Study

Potential Use for Group Comparison. The results of the reliability study may help to determine the minimal size of the sample of a study as far as the delay between consecutive pairs of trials is similar to the delay of the reliability study (around 6 mo in that experiment). Under that condition of delay, we take into account the 2 components of the typical error: the experimental error and the biologic variability. For a test–retest study with short delay, the biologic error may be different and ideally reduced to zero. In that case, variability of measurements is only due to experimental error rather than close to the typical error found in the simulated data presented in this article. But in specific cases of long-term clinical studies, the natural biologic variability of the control population has to be known to determine the minimal sample size allowing optimal conditions for detection. In crossover studies, a simplified formula useful to fix sample size is $n = 8 \ s^2/d^2$, where "d"

is the minimal difference to be detected between pre- and post test acquisitions, and "s" is the typical error found in the reliability study (*23*). The coefficient 8 is an approximation of 2 times the inverse Student distribution for a confidence level of 95%. As an example, expecting a difference of 5% of the variation of BP in the hippocampus (typical error of 7.4% in our reliability study) in a crossover test–retest study will require a sample size of 17 subjects for a statistical power of 95%. That sample size must be multiplied by 4 when another independent group is used as control. When the expected differences are much larger than noise—for example, looking for a difference of 15% in BP in the hippocampus—only 2 subjects are required (8 × $7.4^2/15^2 = 1.9$). In that case, the only restriction is to ensure that the selected 2 subjects are representative of a wider population. The major advantage of conducting a reliability study is that it allows performing a crossover study without a healthy control group, just from the knowledge of the typical error of the present reliability study.

Assessing for Individual Measurement. In that case, the typical error is used quite differently. One approach consists of assessing the difference between 2 scans: if the difference exceeds a confidence interval based on the typical error found in the reliability study (mean ± 2 SD for a 95% likelihood). Another approach, more powerful and less restrictive, consists of establishing whether an expected difference between test and retest acquisitions is exceeded by measurement of an individual subject, with reasonable likelihood. That approach forces one to make an estimation of what smallest clinically important changes between measurements would have a significant importance. So, a priori knowledge of biologic variability is required. Let's suppose that a 10% modification of ^{18}F-MPPF BP is an expected value for evidence of clinical variation in the serotoninergic system. With that case, if an individual patient presented a scan difference of 14% in the hippocampus, knowledge of the test–retest typical error allows evaluation of the confidence interval of the true value of changes for a determinate likelihood—that is, for an 80% likelihood, the factor to be applied to the typical error around the measurement is 1.81. With that value, the confidence interval around the patient variation of 14% will be [12.2;15.9]. We can assume that, with an 80% likelihood, true change of the measured change of 14% is greater than a 10% change, so that a change actually occurred between test and retest conditions. This approach is less conservative but is more effective and clinically practical for deciding on an effect in therapeutics.

More generally, a general usage of parametric imaging is the established individual or group comparison at a voxel level, thanks to the usual and friendly approach of SPM. In that study, we verified that the parametric images did not present significant bias between test and retest measurements: This database is then suitable for statistical inference in group comparisons and individual assessment via the general linear model.

CONCLUSION

The results of this study demonstrate—to our knowledge, for the first time—that ^{18}F-MPPF is reliable for performing PET of brain 5-HT$_{1A}$ receptors in longitudinal studies: ^{18}F-MPPF parametric imaging of BP, R$_1$, and k$_2$ by noninvasive SRTM and DV by Logan graphical analysis was reproducible with long-term delay. The choice of the analytic method and of the measure of interest can be freely selected and motivated by the purpose of the study, as the results presented in this article showed that both models allow computing indices with similar reproducibility. The parametric data also provide knowledge of noise, allowing the estimation of sample size for group comparison, and confidence interval for individual subject assessment.

REFERENCES

1. Barnes NM, Sharp T. A review of central 5-HT receptors and their function. *Neuropharmacology.* 1999;38:1083–1152.
2. Lanfumey L, Hamon M. 5-HT1 receptors. *Curr Drug Targets CNS Neurol Disord.* 2004;3:1–10.
3. Aznavour N, Zimmer L. [^{18}F]MPPF as a tool for the in vivo imaging of 5-HT1A receptors in animal and human brain. *Neuropharmacology.* 2007;52:695–707.
4. Derry C, Benjamin C, Bladin P, et al. Increased serotonin receptor availability in human sleep: evidence from an [^{18}F]MPPF PET study in narcolepsy. *Neuroimage.* 2006;30:341–348.
5. Merlet I, Ostrowsky K, Costes N, et al. 5-HT1A receptor binding and intracerebral activity in temporal lobe epilepsy: an [^{18}F]MPPF-PET study. *Brain.* 2004;127:900–913.
6. Merlet I, Ryvlin P, Costes N, et al. Statistical parametric mapping of 5-HT1A receptor binding in temporal lobe epilepsy with hippocampal ictal onset on intracranial EEG. *Neuroimage.* 2004;22:886–896.
7. Praschak-Rieder N, Hussey D, Wilson AA, et al. Tryptophan depletion and serotonin loss in selective serotonin reuptake inhibitor-treated depression: an [^{18}F]MPPF positron emission tomography study. *Biol Psychiatry.* 2004;56:587–591.
8. Riad M, Zimmer L, Rbah L, Watkins KC, Hamon M, Descarries L. Acute treatment with the antidepressant fluoxetine internalizes 5-HT1A autoreceptors and reduces the in vivo binding of the PET radioligand [^{18}F]MPPF in the nucleus raphe dorsalis of rat. *J Neurosci.* 2004;24:5420–5426.
9. Aznavour N, Rbah L, Riad M, Reilhac A, Costes N, Descarries LA. PET imaging study of 5-HT1A receptors in cat brain after acute and chronic fluoxetine treatment. *Neuroimage.* 2006;33:834–842.
10. Passchier J, van Waarde A, Pieterman RM, et al. Quantitative imaging of 5-HT$_{1A}$ receptor binding in healthy volunteers with [^{18}F]p-MPPF. *Nucl Med Biol.* 2000; 27:473–476.
11. Costes N, Merlet I, Zimmer L, et al. Modeling [^{18}F]MPPF PET kinetics for the determination of 5-HT1A concentration with multi-injection. *J Cereb Blood Flow Metab.* 2002;22:753–763.
12. Costes N, Merlet I, Ostrowsky K, et al. A ^{18}F-MPPF PET normative database of 5-HT1A receptor binding in men and women over aging. *J Nucl Med.* 2005; 46:1980–1989.
13. Soares JC, van Dyck CH, Tan P, et al. Reproducibility of in vivo brain measures of 5-HT2A receptors with PET and [^{18}F]deuteoralteranserin. *Psychiatry Res.* 2001;106:81–93.
14. Hirvonen J, Kajander J, Allonen T, Oikonen V, Nagren K, Hietala J. Measurement of serotonin 5-HT1A receptor binding using positron emission tomography and [carbonyl-^{11}C]WAY-100635: considerations on the validity of cerebellum as a reference region. *J Cereb Blood Flow Metab.* 2007;27: 185–195.

15. Gunn RN, Lammertsma AA, Grasby PM. Quantitative analysis of [carbonyl-^{11}C]WAY-100635 PET studies. *Nucl Med Biol.* 2000;27:477–482.

16. Parsey RV, Slifstein M, Hwang DR, et al. Validation and reproducibility of measurement of 5-HT1A receptor parameters with [carbonyl-^{11}C]WAY-100635 in humans: comparison of arterial and reference tissue input functions. *J Cereb Blood Flow Metab.* 2000;20:1111–1133.

17. Goldberg DP, Hillier VF. A scaled version of the general health questionnaire. *Psychol Med.* 1979;9:139–145.

18. Le Bars D, Lemaire C, Ginovart N, et al. High-yield radiosynthesis and preliminary in vivo evaluation of p-[^{18}F]MPPF, a fluoro analog of WAY-100635. *Nucl Med Biol.* 1998;25:343–350.

19. Le Bars D, Bonmarchand G, Alvarez G, Lemaire C, Mosdzianowski C. New automation of MPPF using a coincidence synthesizer [abstract]. *J Labelled Compd Radiopharm.* 2001;44(suppl):S1045.

20. Reilhac A, Evans A, Gimenez G, Costes N. Creation and application of a simulated database of dynamic [^{18}F]MPPF PET acquisitions incorporating inter-individual anatomical and biological variability. *IEEE Trans Med Imaging.* 2006; 25:1431–1439.

21. Gunn RN, Lammertsma AA, Hume SP, Cunningham VJ. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *Neuroimage.* 1997;6:279–287.

22. Logan J. Graphical analysis of PET data applied to reversible and irreversible tracers. *Nucl Med Biol.* 2000;27:661–670.

23. Hopkins WG. A new view of statistics. *Internet Society for Sport Science.* 2000; Available at: http://www.sportsci.org/resource/stats/. Accessed June 29, 2007.

24. Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC. A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp.* 1996;4:58–73.

25. Snedecor G, Cochran W, Cox D. *Statistical Methods.* Ames, IA: The Iowa State University Press; 1989.

26. Kim JS, Ichise M, Sangare J, Innis RB. PET imaging of serotonin transporters with [^{11}C]DASB: test-retest reproducibility using a multilinear reference tissue parametric imaging method. *J Nucl Med.* 2006;47:208–214.

27. Frankle WG, Slifstein M, Gunn RN, et al. Estimation of serotonin transporter parameters with ^{11}C-DASB in healthy humans: reproducibility and comparison of methods. *J Nucl Med.* 2006;47:815–826.

28. Rueter LE, Fornal CA, Jacobs BL. A critical review of 5-HT brain microdialysis and behavior. *Rev Neurosci.* 1997;8:117–137.

29. Meltzer CC, Drevets WC, Price JC, et al. Gender-specific aging effects on the serotonin 1A receptor. *Brain Res.* 2001;895:9–17.

30. Zimmer L, Riad M, Rbah L, et al. Toward brain imaging of serotonin 5-HT1A autoreceptor internalization. *Neuroimage.* 2004;22:1421–1426.