Test-retest reproducibility of cerebral adenosine A(2A) receptor quantification using [C-11]preladenant

Toyohara, Jun; Sakata, Muneyuki; Wagatsuma, Kei; Tago, Tetsuro; Ishibashi, Kenji; Ishii, Kenji; Elsinga, Philip; Ishiwata, Kiichi

Published in:
ANNALS OF NUCLEAR MEDICINE

DOI:
10.1007/s12149-021-01678-5

Publication date:
2022

Annals of Nuclear Medicine (2022) 36:15–23
https://doi.org/10.1007/s12149-021-01678-5

ORIGINAL ARTICLE

Test–retest reproducibility of cerebral adenosine A_{2A} receptor quantification using [^{11}C]preladenant

Jun Toyohara[1] · Muneyuki Sakata[1] · Kei Wagatsuma[1,2] · Tetsuro Tago[1] · Kenji Ishibashi[1] · Kenji Ishii[1] · Philip Elsinga[3] · Kiichi Ishiwata[1,4,5]

Received: 11 July 2021 / Accepted: 13 September 2021 / Published online: 26 September 2021

Abstract

Objective To evaluate the reproducibility of cerebral adenosine A_{2A} receptor (A_{2A}R) quantification using [^{11}C]preladenant ([^{11}C]PLN) and PET in a test–retest study.

Methods Eight healthy male volunteers were enrolled. Dynamic 90 min PET scans were performed twice at the same time of the day to avoid the effect of diurnal variation. Subjects refrained from caffeine from 12 h prior to scanning, and serum caffeine was measured before radioligand injection. Arterial blood was sampled repeatedly during scanning and the fraction of the parent compound in plasma was determined. Total distribution volume (V_T) was estimated using 1- and 2-tissue compartment models (1-TCM and 2-TCM, respectively) and Logan graphical analysis (Logan plot) ($t^* = 30$ min). Plasma-free fraction (f_P) of [^{11}C]PLN was measured and used for correction of V_T values. Distribution volume ratio (DVR) was calculated from V_T of target and reference regions and obtained by noninvasive Logan graphical reference tissue model (LGAR) ($t^* = 30$ min). Absolute test–retest variability (aTRV), and intra-class correlation coefficient (ICC) of V_T and DVR were calculated as indexes of repeatability. Correlation between DVR and serum concentration of caffeine (a nonselective A_{2A}R blocker) was analyzed by Pearson's correlation analysis.

Results Regional time–activity curves were well described by 2-TCM models. Estimation of V_T by 2-TCM produced some erroneous values; therefore, the more robust Logan plot was selected as the appropriate model. Global mean aTRV was 20% for V_T and 14% for V_T/f_P (ICC, 0.72 for V_T and 0.87 for V_T/f_P). Global mean aTRV of DVR was 13% for Logan plot and 10% for LGAR (ICC, 0.70 for Logan plot and 0.81 for LGAR). DVR estimates using LGAR and Logan plot were in good agreement ($r^2 = 0.96$). Coefficients of variation for V_T, V_T/f_P, DVR (Logan plot), and DVR (LGAR) were 47%, 47%, 27%, and 18%, respectively. Despite low serum caffeine levels, significant concentration-dependent effects on [^{11}C]PLN binding to target regions were observed ($p < 0.01$).

Conclusions In this study, moderate test–retest reproducibility and large inter-subject differences were observed with [^{11}C]PLN PET, possibly attributable to competition by baseline amount of caffeine. Analysis of plasma caffeine concentration is recommended during [^{11}C]PLN PET studies.

Trial registration UMIN000030040.

Keywords Adenosine A_{2A} receptor · Preladenant · Positron emission tomography · Reproducibility

✉ Jun Toyohara
toyohara@pet.tmig.or.jp

[1] Research Team for Neuroimaging, Tokyo Metropolitan Institute of Gerontology, 35-2 Sakae-cho, Itabashi-ku, Tokyo 173-0015, Japan

[2] School of Allied Health Science, Kitasato University, 1-15-1 Kitasato, Sagamihara, Kanagawa 252-0373, Japan

[3] Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands

[4] Institute of Cyclotron and Drug Discovery Research, Southern Tohoku Research Institute for Neuroscience, 7-115 Yatsuyamada, Koriyama, Fukushima 963-8563, Japan

[5] Department of Biofunctional Imaging, Fukushima Medical University, 1 Hikariga-oka, Fukushima 960-1295, Japan

Introduction

Cerebral adenosine A_{2A} receptors (A_{2A}Rs) are involved in numerous neuropsychiatric disorders, including Parkinson disease, Alzheimer disease, drug addiction, alcohol abuse, epilepsy seizures, sleep disorders, and schizophrenia [1]. The selective A_{2A}R antagonist istradefylline was approved in Japan in 2013 and US in 2019 as an add-on treatment to levodopa/carbidopa treatment in adults with Parkinson disease patients experiencing "off" episodes [2].

Noninvasive in vivo imaging of A_{2A}Rs with PET represents a potentially useful method for monitoring changes in A_{2A}R density during a disease course and for the assessment of receptor occupancy of investigated drugs. Several tracers for PET imaging of A_{2A}Rs have been developed for use in humans. These include xanthine derivatives [7-methyl-^{11}C]-(E)-8-(3,4,5-trimethoxystyryl)-1,3,7-trimethylxanthine ([^{11}C]TMSX) [3] and [4-O-methyl-^{11}C]-8-[(1E)-2-(3,4-dimethoxyphenyl)ethenyl]-1,3-diethyl-3,7-dihydro-7-methyl-1H-purine-2,6-dione ([^{11}C]KW-6002) [4], non-xanthine ligands 5-amino-7-(3-(4-[^{11}C]methoxy-phenyl)propyl)-2-(2-furyl)pyrazolo[4,3-e]-1,2,4-triazole[1,5-c]pyrimidine ([^{11}C]SCH442416) [5], 2-(furan-2-yl)-7-(2-(4-(4-(2-[^{11}C]methoxethoxy)phenyl)piperain-1-yl)ethyl)-7H-pyrazolo[4,3-e][1,2,4]triazol[1,5-c]pyrimidine-5-amine ([^{11}C]preladenant; [^{11}C]PLN) [6] and its [^{18}F]fluoroethoxy derivative 7-(2-(4-(4-(2-[^{18}F]fluoroethoxy)phenyl)piperazin-2-yl)ethyl)-2-(furan-2-yl)-7H-pyrazolo[4,3-e][1,2,4]triazol[1,5-c]pyrimidine-5-amine ([^{18}F]MNI-444) [7]. Among these, [^{11}C]PLN and [^{18}F]MNI-444 have the highest target-to-nontarget ratios (striatum-to-cerebellum ratio > 5).

Uptake of [^{11}C]PLN has been evaluated in healthy human subjects as well as in receptor occupancy studies of Parkinson disease patients [8, 9]. Several outcome measures can be used to detect changes in A_{2A}R binding of [^{11}C]PLN caused by disease progression or therapeutic interventions. Test–retest validation of the tracer is fundamentally important for evaluating its utility for detecting changes in A_{2A}Rs density. The aim of the present study was to evaluate the reproducibility of cerebral A_{2A}R quantification using [^{11}C]PLN and PET in a test–retest study. We also analyzed the effects of the level of serum caffeine, a nonselective A_{2A}R blocker that is consumed daily, on [^{11}C]PLN binding.

Materials and methods

Study subjects and eligibility

Eight healthy male volunteers were enrolled in this study (mean age ± SD, 26 ± 10 y; range 21 − 51 y). The inclusion criteria were as follows: age 20–60 years; male sex; able to provide informed consent; and normal medical history, physical examination, and vital-sign findings. The exclusion criteria were liver or kidney dysfunction, abnormal central nervous system findings, cardiac failure, drug or food allergy, or recent smoking history. The subjects weighed 55.3 − 93.1 kg (mean weight ± SD, 65.4 ± 11.6 kg). Although there were no dietary restrictions, all subjects were required to refrain from drinking caffeinated beverages from the night before PET imaging.

A T1-weighted whole-brain image was acquired for each subject for anatomical co-registration, using a GE Discovery MR750w 3.0 T scanner (GE Healthcare, Wauwatosa, WI) and the following parameters: sagittal three-dimensional (3D) fast spoiled gradient-echo, repetition time = 7.6 ms, echo time = 3.1 ms, inversion time = 400 ms, matrix = 256 × 256 × 196 voxels, voxel size = 1.06 × 1.06 × 1.20 mm. All eight subjects were free of somatic and neuropsychiatric illnesses according to their medical history and findings of physical examination and had no brain abnormalities on MRI.

All experiments were approved by the institutional review board of the Tokyo Metropolitan Institute of Gerontology and were performed in accordance with the institutional review board rules and policies. All subjects gave study-specific informed consent to participate in the study and all experiments were carried out in accordance with the relevant guidelines. The study was registered in UMIN-CTR (Trial ID: UMIN000030040).

Radiochemistry

Radiosynthesis and quality control of [^{11}C]PLN were performed as described previously [6]. [^{11}C]PLN was obtained with radiochemical purity of 98.5% ± 0.3% (range 97.8–98.9%) and molar activity of 101 ± 17 GBq/μmol (range 65–126 GBq/μmol) at the end of synthesis.

Brain PET imaging

Dynamic [^{11}C]PLN PET scanning was performed twice for each subject. The interval between the two scans was 177 ± 128 days, and scans were obtained at 10:00 in four subjects and 14:00 in the remaining four subjects on each day to avoid the effect of diurnal variation.

The PET data were acquired using a Discovery PET/computed tomography 710 scanner (GE Healthcare; axial field of view = 15.7 cm, spatial resolution = 4.5 mm full width at half maximum (FWHM), *Z*-axis resolution = 4.8 mm FWHM) [10].

After low-dose computed tomography scanning to correct for attenuation, [^{11}C]PLN (705 ± 46 MBq, 65 ± 19 MBq/nmol) was injected into the antecubital vein for 1 min as

a bolus, and 90-min dynamic scanning was performed (20 s × three frames, 30 s × three frames, 60 s × five frames, 150 s × five frames, and 300 s × 14 frames). Arterial blood (0.5 ml each) was sampled at 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 135, 150, and 180 s, as well as at 5, 7, 10, 15, 20, 30, 40, 50, 60, 75, and 90 min. The whole blood and separated plasma were weighed, and radioactivity was measured with a NaI (Tl) well scintillation counter (BeWell Model-QS03 F/B; Molecular Imaging Labo, Suita, Japan). To analyze the labeled metabolites, an additional 1.5 ml of blood was obtained at 3, 10, 20, 30, 40, and 60 min. Unaltered [^{11}C]PLN in the plasma was analyzed with high-performance liquid chromatography (HPLC), and the metabolite-corrected time–activity curve (TAC) of plasma was obtained as described previously [6].

Blood analysis

Before injection of the radioligand, 4 ml of arterial blood and 4 ml of venous blood were sampled for the evaluation of free fraction of [^{11}C]PLN in arterial plasma (f_P) and serum caffeine concentration, respectively. f_P was determined by the ultrafiltration method as described previously [11], and caffeine concentrations were determined by liquid chromatography mass spectrometry at Bozo Research Center (Tokyo, Japan).

Kinetic analysis

The 30 frames obtained in dynamic PET imaging (matrix size of 128 × 128 × 47 voxels and voxel size of 2.0 × 2.0 × 3.27 mm) were reconstructed using a 3D ordered-subset expectation maximization algorithm (16 subsets, four iterations) incorporating time-of-flight information. The dynamic images were post-smoothed with a Gaussian filter (4 mm FWHM). Motion correction was applied to the dynamic images with reference to the early phase (0–10 min) using the realign method in SPM8 (Wellcome Trust Center for Neuroimaging, University College London, UK).

Partially overlapping circular volumes of interest (VOIs) of diameter 10 mm were placed on the cerebellar cortex as the reference region; on the temporal cortex and thalamus as the non-target regions; and on the putamen and head of caudate nucleus as the target regions, with reference to the mean images of the early (0–20 min) phase of the PET images and the T1-weighted MR images co-registered to the mean early-phase PET images (Supplemental Fig. 1). VOIs were calculated as Bq/ml or as standardized uptake value [SUV, (activity/ml tissue)/(injected activity/body weight)]. The rotations for co-registration of the retest scan images with reference to the test scan images were calculated from both static images of the early phases (0–10 min) using the coregistration method in SPM8. VOIs were positioned on a dynamic image of the test scan, and on a dynamic image of the retest scan co-registered to the test scan image. Total distribution volume (V_T) was estimated by 1- and 2-tissue-compartment models (1-TCM and 2-TCM, respectively) and by Logan graphical analysis (Logan plot) [12] ($t^* = 30$ min) using the TACs for these VOIs and the metabolite-corrected TAC of plasma, as described previously [6]. Blood delay in each subject was calculated using the mean TAC of the overall VOIs, and this value was then fixed for each VOI. All models were assessed using both a fixed (5%) fractional blood volume (vB) and using vB as the fitting parameter. Goodness of fit was evaluated using Akaike information criterion (AIC). Distribution volume ratio (DVR) was calculated from the V_T of the target and reference regions and estimated by Logan reference tissue method (LGAR) [13] ($t^* = 30$ min). The absolute test–retest variability (aTRV) and intra-class correlation coefficient (ICC; ICC(1,1) in [14]) of V_T and DVR were calculated as the index of repeatability. The AIC, aTRV, and ICC were calculated as follows:

$$\mathrm{AIC} = n \ln\left(\frac{\mathrm{RSS}}{n}\right) + 2p + \frac{2p(p+1)}{n-p-1},$$

where RSS, n, and p represent the residual sum of squares of the fitting, number of frames, and number of fitted parameters;

$$\mathrm{aTRV} = \left|\frac{R_{\mathrm{test}} - R_{\mathrm{retest}}}{(R_{\mathrm{test}} + R_{\mathrm{retest}})/2}\right| \times 100,$$

where R_{test} and R_{retest} represent the results (V_T or DVR) of the test and retest scan, respectively; and

$$\mathrm{ICC} = \frac{\mathrm{BSMSS} - \mathrm{WSMSS}}{\mathrm{BSMSS} + \mathrm{WSMSS}},$$

where BSMSS and WSMSS represent the mean sum of squares (MSS) of between-subject (BS) and within-subject (WS) differences, respectively.

Correlation between DVR and plasma caffeine concentration was analyzed by Pearson's correlation analysis. Outlier data of V_T (2-TCM) were identified by robust regression followed by the outlier identification (ROUT) method [15] using GraphPad Prism Ver. 9.1.2 software (San Diego, CA). The ROUT method follows three steps. (1) Robust nonlinear regression based on the assumption that scatter follows a Lorentzian distribution is used to fit a curve that is not influenced by outliers. (2) The residuals of the robust fit are analyzed to identify any outliers. This step uses a new outlier test adapted from the False Discovery Rate approach of testing for multiple comparisons. (3) Outliers are removed and ordinary least-squares regression is performed on the remaining data.

Results

Subject dosing

Molar activity (MBq/nmol), injected mass (nmol), and normalized injected mass (pmol/kg) at the time of injection were not significantly different between the test and retest scans ($p = 0.42$, 0.43, and 0.46, respectively; paired t test). Injected dose (MBq) and normalized injected dose (MBq/kg) were significantly higher in the test scan than the retest scan ($p < 0.01$ and $p < 0.05$, respectively; paired t test). The estimated occupancy rate of $A_{2A}R$ in human striatum at the maximum uptake (0.7 ± 0.3 pmol/ml; range 0.4–1.5 pmol/ml) estimated for an $A_{2A}R$ density of 85 pmol/ml [1] was very small (< 1%) and was not different between the test and retest groups: $0.9\% \pm 0.4\%$ for the test group and $0.8\% \pm 0.4\%$ for the retest group ($p = 0.38$, paired t test). Therefore, differences between the injected dose and the normalized injected dose did not affect the $A_{2A}R$ measurement of [^{11}C]PLN. Table 1 summarizes the dose information for each of the test and retest scan groups.

Arterial input function

Figure 1a shows the averaged metabolite-corrected plasma radioactivity of the test and retest scans. Peripheral metabolism of [^{11}C]PLN was slow and did not differ significantly between the test and retest scans (Fig. 1b) ($p = 0.09$, paired t test). There was no significant difference in f_P between the test and retest scans (Table 1) ($p = 0.52$, paired t test). The aTRV of f_P was $12\% \pm 10\%$. The amount of non-specific binding to the filter was $57\% \pm 5\%$ in the test scans and $60\% \pm 2\%$ in the retest scans, indicating high retention of [^{11}C]PLN on the filter. There was no significant difference in serum caffeine concentration between the test and retest scans (Table 1) ($p = 0.97$, paired t test).

Table 1 Dose information for test and retest groups including net injected doses of [^{11}C]preladenant with corresponding injected mass, molar activity, plasma-free fraction, and serum caffeine concentration

Groups	Activity (MBq)	Injected mass (nmol)	MA (GBq/μmol)	Normalized injected dose (MBq/kg)	Normalized injected mass (pmol/kg)	f_P (%)	Serum caffeine (μg/ml)
Test	728 ± 53** (659–809)	13 ± 4 (7–18)	61 ± 19 (38–94)	11 ± 2* (8–14)	203 ± 72 (127–308)	3.3 ± 0.8 (2.3–4.8)	0.69 ± 0.60 (0.01–1.44)
Retest	682 ± 23 (640–706)	11 ± 6 (8–26)	68 ± 19 (26–82)	10 ± 2 (7–13)	177 ± 93 (108–397)	3.4 ± 0.6 (2.5–4.3)	0.70 ± 0.45 (0.32–1.75)

Data obtained in healthy male subjects ($n = 8$), represented as mean ± SD

Significant difference in activity (*$p < 0.05$, **$p < 0.01$) between test and retest groups (paired t-test)

MA molar activity, f_P plasma-free fraction

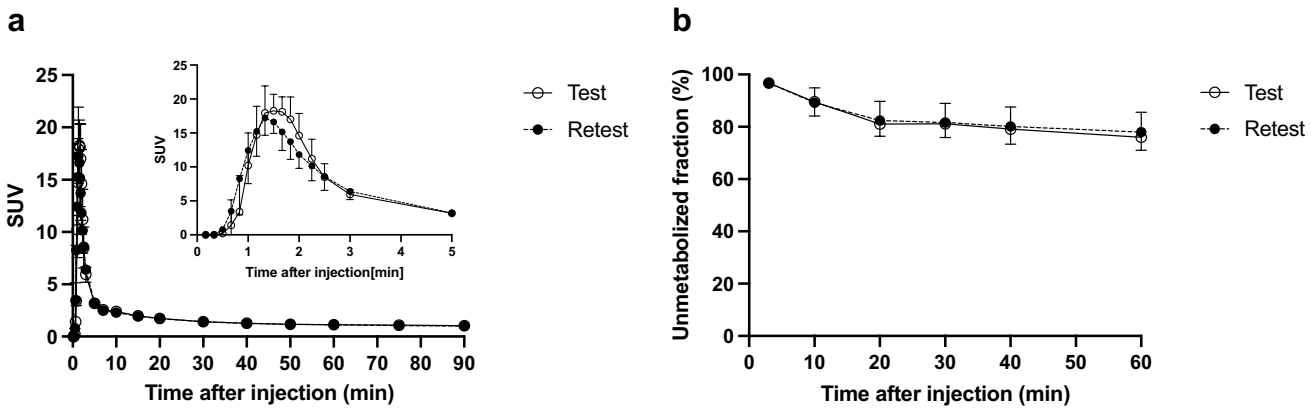

Fig. 1 Metabolite-corrected input function and unmetabolized fraction of [^{11}C]PLN. **a** Mean decay-corrected time–activity curves of metabolite-corrected plasma in test (open circles), and retest (closed circles) subjects. The insert shows extracted values for the first 5 min. **b** Unchanged fraction of [^{11}C]PLN in test (open circles) and retest (closed circles) subjects. Data represent the mean ± SD in 8 subjects

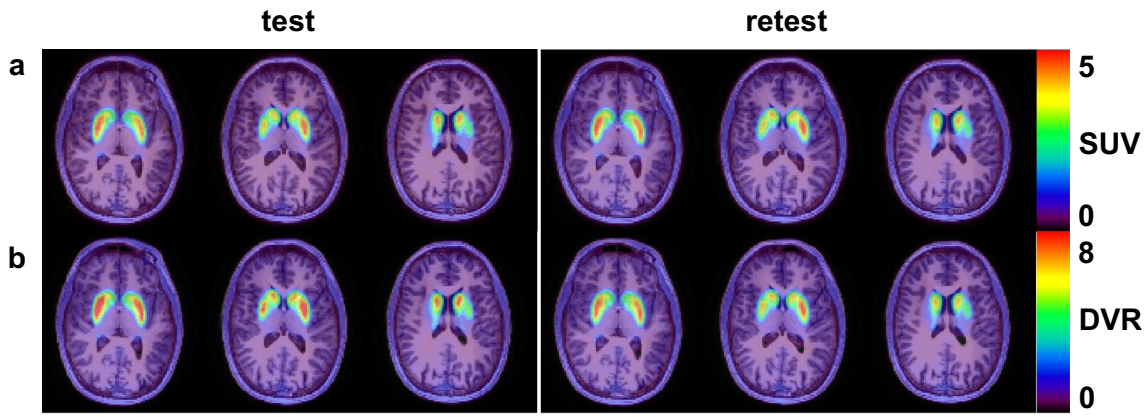

Fig. 2 Representative MR-fused images of [^{11}C]PLN in test and retest conditions. **a** Static [^{11}C]PLN images in the pseudoequilibrium phase (40–60 min). **b** Parametric DVR images

Brain kinetics

Figure 2 shows representative MR-fused static [^{11}C]PLN images of pseudoequilibrium phase (40–60 min) and parametric DVR images for the test and retest scans. According to the AIC scores, the best fits were obtained when vB was used as the fitting parameter (1-TCM: vB fixed = 438 ± 37 versus vB fit = 426 ± 33; 2-TCM: vB fixed = 379 ± 25 versus vB fit = 365 ± 26, $p < 0.0001$, paired t test). The 2-TCM provided significantly lower AIC scores than the 1-TCM (1-TCM = 426 ± 33 versus 2-TCM = 365 ± 26, $p < 0.0001$, paired t test) and good quality of visual assessment for fit. In some individuals, however, 2-TCM provided implausibly large V_T estimates with large standard errors of the nonlinear fitting estimations, most commonly in regions of low uptake (Supplemental Tables 1 and 2). Due to the unreliability of V_T estimation by 2-TCM, the more robust Logan plot was used for evaluation. When excluding outliers of V_T (2-TCM), the V_T values from Logan plot (64/80 regions) matched well with those from the 2-TCM models (Supplemental Fig. 2; V_T [Logan plot] = 0.95V_T [2-TCM] − 0.01, $r^2 = 0.99$).

Tables 2 and 3 list the obtained values of V_T (Logan plot), V_T/f_P (Logan plot), DVR, aTRV, and ICC. There was no significant difference in V_T (Logan plot), V_T/f_P (Logan plot), or DVR between the test and retest scans ($p > 0.3$, paired t test). Global mean aTRV was 20% and 14% for V_T and V_T/f_P, respectively. ICC was 0.72 ± 0.13 for V_T and 0.87 ± 0.08 for V_T/f_P. Global mean aTRV was 13% and 10% for DVR using Logan plot and LGAR, respectively. ICC was 0.70 ± 0.20 for DVR (Logan plot) and 0.81 ± 0.07 for DVR (LGAR). The test and retest values of V_T (Logan plot), V_T/f_P (Logan plot), and DVR were also compared by linear regression analysis (Fig. 3). The correlation coefficients of linear regressions between the test and retest values were good ($r^2 \geq 0.87$). Among these, V_T/f_P (Fig. 3b; $r^2 = 0.95$) and DVR (LGAR) (Fig. 3d; $r^2 = 0.92$) gave the best correlation. The correlation coefficient of linear regressions of DVR between Logan plot and LGAR was good (Fig. 4a) ($r^2 = 0.96$). Bland–Altman plot revealed slightly positive average difference of DVR between LGAR and Logan plot (0.29), meaning that values obtained with LGAR were slightly higher than those obtained with Logan plot (Fig. 4b).

Table 2 Test–retest variability and reproducibility of V_T and V_T/f_P (Logan plot)

Brain region	V_T						V_T/f_P					
	Test	CV (%)	Retest	CV (%)	aTRV (%)	ICC	Test	CV (%)	Retest	CV (%)	aTRV (%)	ICC
Cerebellar cortex	0.73 ± 0.48	65	0.76 ± 0.42	55	21 ± 13	0.87	22.9 ± 15.3	67	23.9 ± 17.6	74	16 ± 13	0.96
Temporal cortex	0.69 ± 0.49	70	0.74 ± 0.41	56	25 ± 16	0.79	21.2 ± 13.2	62	23.1 ± 17.1	74	20 ± 14	0.91
Thalamus	0.75 ± 0.64	86	0.75 ± 0.43	57	27 ± 16	0.76	22.3 ± 14.9	67	23.4 ± 17.0	73	21 ± 14	0.90
Putamen	2.86 ± 0.71	25	2.81 ± 0.51	18	12 ± 12	0.60	87.3 ± 10.0	11	82.5 ± 13.5	16	8 ± 7	0.79
Caudate	2.16 ± 0.47	22	2.12 ± 0.39	18	13 ± 13	0.56	66.1 ± 7.0	11	62.3 ± 10.2	16	8 ± 7	0.77

Data obtained in healthy male subjects ($n = 8$), represented as mean ± SD

V_T total distribution volume, f_P plasma-free fraction, *aTRV* absolute test–retest variability, *ICC* intra-class correlation coefficient, *CV* coefficient of variation

Table 3 Test–retest variability and reproducibility of DVR

Brain region	DVR (Logan plot)						DVR (LGAR)					
	Test	CV (%)	Retest	CV (%)	aTRV (%)	ICC	Test	CV (%)	Retest	CV (%)	aTRV (%)	ICC
Temporal cortex	0.94 ± 0.07	7	0.96 ± 0.04	4	5 ± 5	0.41	0.94 ± 0.03	3	0.94 ± 0.02	2	1 ± 1	0.73
Thalamus	0.98 ± 0.17	17	0.97 ± 0.09	9	6 ± 7	0.71	0.94 ± 0.07	8	0.94 ± 0.06	6	2 ± 2	0.91
Putamen	4.96 ± 2.09	42	4.63 ± 2.06	45	21 ± 12	0.83	5.57 ± 1.63	29	5.27 ± 1.68	32	18 ± 8	0.80
Caudate	3.80 ± 1.67	44	3.52 ± 1.59	45	21 ± 12	0.83	4.37 ± 1.39	32	4.10 ± 1.36	33	18 ± 9	0.81

Data obtained in healthy male subjects ($n = 8$), represented as mean ± SD

DVR distribution volume ratio with reference to cerebellum, *aTRV* absolute test–retest variability, *ICC* intra-class correlation coefficient, *CV* coefficient of variation, *LGAR* Logan reference tissue method

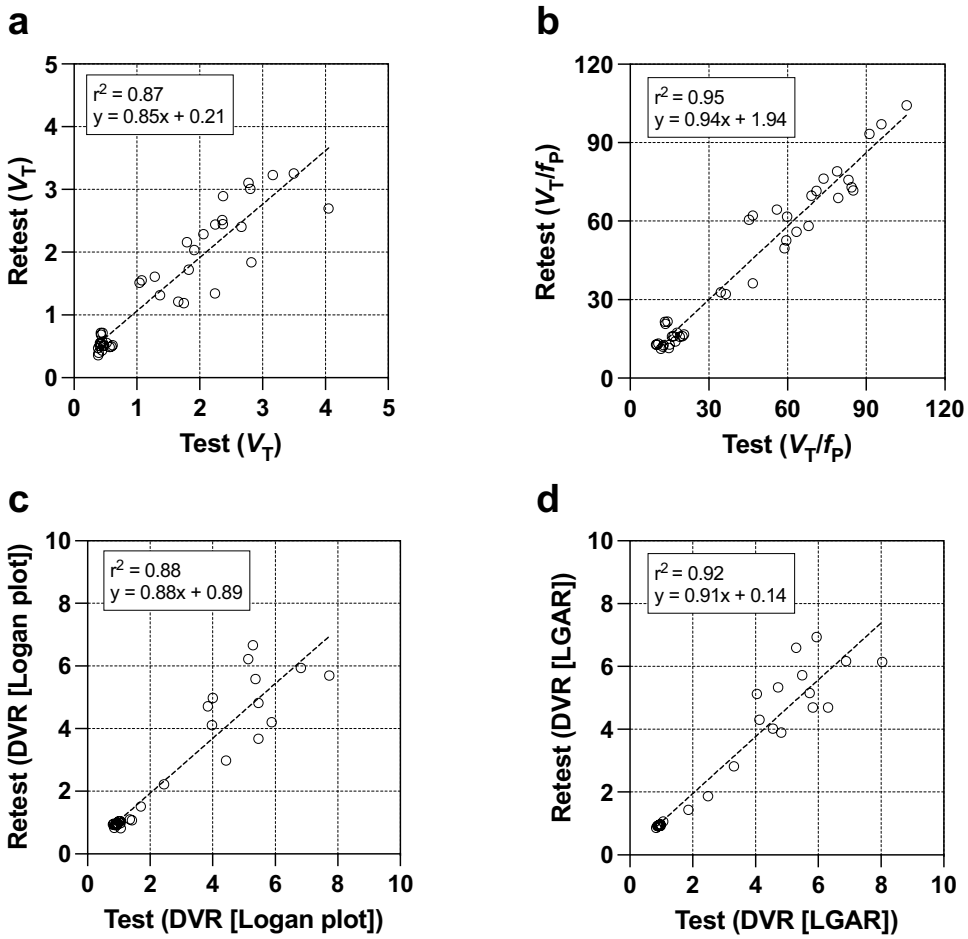

Fig. 3 Correlation of outcome measure values between the test and retest scans. Test–retest values are shown for V_T (**a**), V_T/f_P (**b**), DVR (Logan plot) (**c**), and DVR (LGAR) (**d**)

Inter-subject variations (coefficient of variation: CV) was 47% ± 24% for V_T, 47% ± 29% for V_T/f_P, 27% ± 27% for DVR (Logan plot), and 18% ± 14% for DVR (LGAR) (Tables 2 and 3). Although serum caffeine concentrations were low (0.70 ± 0.52 µg/ml; range, 0.01–1.75 µg/ml), significant concentration-dependent effects on [^{11}C]PLN binding to the target regions (Pearson's correlation coefficient: $p < 0.01$; putamen: $r^2 = 0.49$, caudate: $r^2 = 0.47$) were observed (Fig. 5).

Discussion

In the previous first-in-human study, 2-TCM was chosen as the method of choice for analysis of TACs and estimation of V_T [6]. In the present study, 2-TCM also provided better fit than 1-TCM, but implausibly large V_T estimates with large standard errors were observed in some individuals. These occurred most commonly in regions of low uptake and were caused by the gradual uptake of radioactivity in

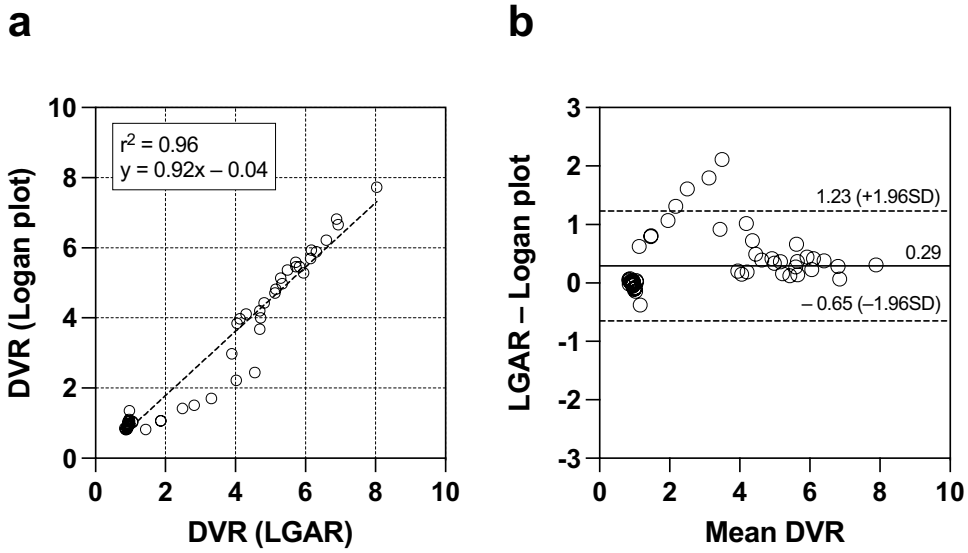

Fig. 4 Comparison of DVR data calculated by Logan plot and LGAR. **a** Correlation of LGAR versus Logan plot. **b** Bland–Altman plots with 95% confidence intervals represent the difference in DVR versus mean DVR for Logan plot versus LGAR

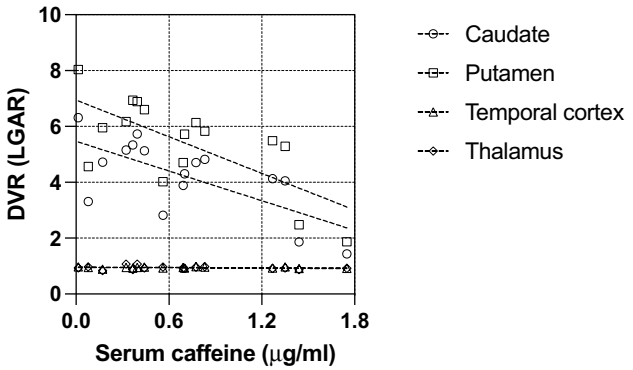

Fig. 5 Correlation between DVR (LGAR) and serum caffeine concentration (μg/ml). Negative linear regression was observed for caudate (DVR [LGAR] = − 1.8 caffeine [μg/ml] + 5.5) and putamen (DVR [LGAR] = − 2.2 caffeine [μg/ml] + 6.9). No correlation was found between DVR and serum caffeine concentration (μg/ml) in the non-target regions (temporal cortex and thalamus)

the brain, which makes it difficult to estimate k_4 (close-to-zero). Erroneously high V_T values were found in 16 regions of low uptake in 6/16 scans (Supplemental Fig. 3). The gradual uptake affected the V_T values even from the Logan plot, which was more robust in V_T estimation. The V_T of cerebellar cortex (reference region) was over-estimated in these scans, whereas DVR was under-estimated (Logan plot) in target regions (DVR of around 2 to 4 in Fig. 4). The effects were offset or diminished DVR from LGAR. Mean metabolism of [^{11}C]PLN was significantly more rapid in these six scans than in the well fitted 2-TCM scans ($p < 0.05$, paired t test of the mean value of each time point) (Supplemental Fig. 4). A previous study in rats found significant amounts of radiolabelled metabolite of [^{11}C]PLN in the brain [16]. Because the radiolabelled metabolite profile in human plasma is identical to that in rodents, peripherally generated radiolabelled metabolites may penetrate the human brain. Due to the above-mentioned convergence issues, which were also experienced in a previous study [6], we did not further investigate 2-TCM as an analytical method.

We then evaluated Logan plot as a more robust method for estimation of V_T values. When excluding outliers of V_T (2-TCM), the V_T values from Logan plot (64/80 regions) matched well with those from the 2-TCM models. Normalization of V_T (Logan plot) applied with f_P improved the global mean aTRV and ICC values, and good aTRV (8%) and ICC (≥ 0.77) were observed for the target regions. In contrast, non-target regions showed moderate aTRV (16%–21%) but higher ICC (≥ 0.90). Although f_P was consistently low, normalization of V_T by f_P was effective for improving variability and reliability in [^{11}C]PLN measurement.

In this study, we selected the cerebellar cortex as the reference region for quantification of DVR. The cerebellum has at least a tenfold lower $A_{2A}R$ density than that of striatum (ca. 85 nM) [1]; thus, it is theoretically possible to measure their density using [^{11}C]PLN ($K_i = 1.1$ nM; results in $B_{max}/K_i \sim 7.7$). In vivo, however, [^{11}C]PLN binding was detected only in tissues with high $A_{2A}R$ density regions. Because only the free fraction of the radioligand in tissue (f_{ND}) can bind to the receptor, the specific in vivo binding of the ligand (BP_{ND}) can be approximated as $BP_{ND} = f_{ND} \times (B_{max}/K_d)$ [17]. As $K_1/k_2 = f_P/f_{ND}$, f_{ND} was calculated as 0.04 from the mean K_1/k_2 value of 0.68 in 2-TCM and mean f_P value of 0.03. Taking these consideration into account, it is difficult to measure in vivo density of $A_{2A}R$ regions that have low density, such as the cerebellum, using [^{11}C]PLN ($BP_{ND} \sim 0.3$). Furthermore, the cerebellum has been used as the reference in previous studies of [^{11}C]PLN in rodents [16], non-human primates [18], and also humans

[6]. We found good agreement in terms of DVR between invasive (Logan Plot) and non-invasive (LGAR) methods of quantification ($r^2 = 0.96$), and low bias associated with inevitable over-estimation of V_T values in the reference region.

Global mean aTRV and ICC of DVR were improved by applying non-invasive LGAR. Moderate aTRV (18%) and high ICC (≥ 0.8) were observed in target regions, whereas good aTRV ($\leq 2\%$) and moderate (0.7) to high ICC (0.9) were observed in non-target regions. Furthermore, the correlation coefficient of linear regressions between the test and retest values was good ($r^2 = 0.92$). These results and the finding of the smallest CV for non-invasive LGAR among the outcome measures confirm non-invasive LGAR as the analytical model of choice.

In this test–retest study, the mean scan interval was about 25 weeks with a large SD of 18 weeks. These scan interval conditions may have been the reason for the moderate aTRV of DVR in the target regions. However, there was no significant effect of the length of test–retest scan intervals on aTRV of DVR. Similarly, a previous study has reported no effect of a longer mean scan interval of 23 weeks and a large SD of 10 weeks on test–retest reproducibility of [^{18}F]PBR111 [19].

We found a large CV for DVR, particularly in the target regions, and the CV was larger than found in our previous studies [6, 8]. There was significant negative correlation between [^{11}C]PLN binding in the target regions and serum caffeine concentration. In contrast, the non-target regions (lower A_{2A}R density regions) were not affected by these lower levels of serum caffeine concentration. We found no effect of diurnal variation in the start time of scans on the target regions (Supplemental Table 3). Although the differences were very small, significantly lower uptake was observed in the afternoon groups for the non-target regions: temporal cortex and thalamus. We speculate that the large CV in the target regions was generated by differences in individual serum caffeine concentrations. Although the subjects were not allowed any caffeine intake for at least 12 h prior to PET scanning, low but significant amounts of caffeine were still present in the serum. The caffeine concentrations measured in the present study were the same as the baseline serum caffeine concentrations (0.88 µg/ml, 4.5 µM) previously reported in subjects who had been caffeine restricted for over 12 h [20]. From our results, the maximum blocking dose of [^{11}C]PLN binding by serum caffeine can be estimated as ≥ 3 µg/ml (≥ 15 µM) by linear extrapolation analysis. Considering caffeine affinity to human A_{2A}R ($K_i = 10$ µM) [21], the dose-dependent blocking effects on [^{11}C]PLN binding in our observed serum caffeine concentrations are reasonable. Pharmacokinetic analysis of a single shot of a typical caffeinated beverage (160 mg) revealed peak plasma caffeine concentration of 3–4 µg/ml (15–20 µM) [22]. Because caffeine has a long half-life in healthy adult humans (4–5 h) [23], stricter limitations (e.g., caffeine restriction of at least > 36 h [24]) are needed for reliable measurement of A_{2A}R with [^{11}C]PLN PET. Correction of DVR with serum caffeine concentrations is a challenging issue that requires further investigation in future studies.

Conclusions

In this study, moderate test–retest reproducibility and large inter-subject differences in [^{11}C]PLN PET were observed. These variances might be attributable to competition from endogenous A_{2A}R binding ligands such as caffeine. Further analysis of the effect of serum caffeine concentration on data interpretation is recommended in future studies .

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12149-021-01678-5.

Acknowledgements We thank Mr. Masanari Sakai and Mr. Kosuke Nishino for technical support with cyclotron operation and radiosynthesis, Ms. Kimiko Yokoyama for care of the subjects during PET scanning, and Ms. Airin Onishi for coordination of the clinical study.

Funding This work was supported in part by Grants-in-Aid for Scientific Research (B) No. 16H05396 (K.I.) and (C) No. 21K07663 (J.T.) from the Japan Society for the Promotion of Science.

Availability of data and materials The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Code availability Not applicable.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval All procedures performed in studies involving human participants were approved by the Ethics Committee of the Tokyo Metropolitan Institute of Gerontology (H28-46), and in accordance with the principles of the 1964 Declaration of Helsinki and its later amendments.

Consent to participate Written informed consent was obtained from all individual participants prior to their inclusion in the study.

Consent to publication Not applicable.

References

1. van Waarde A, Dierckx RAJO, Zhou X, Khanapur S, Tsukada H, Ishiwata K, et al. Potential therapeutic application of adenosine A_{2A} receptor ligands and opportunities for A_{2A} receptor imaging. Med Res Rev. 2018;38:5–56.
2. Paton DM. Istradefylline: adenosine A_{2A} receptor antagonist to reduce "off" time in Parkinson's disease. Drugs Today (Barc). 2020;56:125–34.
3. Naganawa M, Kimura Y, Mishina M, Manabe Y, Chihara K, Oda K, et al. Quantification of adenosine A_{2A} receptors in the human

brain using [^{11}C]TMSX and positron emission tomography. Eur J Nucl Med Mol Imaging. 2007;34:679–87.
4. Brooks DJ, Doder M, Osman S, Luthra SK, Hirani E, Hume S, et al. Positron emission tomography analysis of [^{11}C]KW-6002 binding to human and rat adenosine A_{2A} receptors in the brain. Synapse. 2008;62:671–81.
5. Ramlackhansingh AF, Bose SK, Ahmed I, Turkheimer FE, Pavese N, Brooks DJ. Adenosine 2A receptor availability in dyskinetic and nondyskinetic patients with Parkinson disease. Neurology. 2011;76:1811–6.
6. Sakata M, Ishibashi K, Imai M, Wagatsuma K, Ishii K, Zhou X, et al. Initial evaluation of an adenosine A_{2A} receptor ligand, ^{11}C-preladenant, in healthy human subjects. J Nucl Med. 2017;58:1464–70.
7. Barret O, Hannestad J, Vala C, Alagille D, Tavares A, Laruelle M, et al. Characterization in humans of ^{18}F-MNI-444, a PET radiotracer for brain adenosine 2A receptors. J Nucl Med. 2015;56:586–91.
8. Ishibashi K, Miura Y, Wagatsuma K, Toyohara J, Ishiwata K, Ishii K. Occupancy of adenosine A_{2A} receptors by istradefylline in patients with Parkinson's disease using ^{11}C-preladenant PET. Neuropharmacology. 2018;143:106–12.
9. Ishibashi K, Miura Y, Wagatsuma K, Toyohara J, Ishiwata K, Ishii K. Adenosine A_{2A} receptor occupancy by long-term istradefylline administration in Parkinson's disease. Mov Disord. 2021;38:268–9.
10. Wagatsuma K, Miwa K, Sakata M, Oda K, Ono H, Kameyama M, et al. Comparison between new-generation SiPM-based and conventional PMT-based TOF-PET/CT. Phys Med. 2017;42:203–10.
11. Toyohara J, Yamamoto H, Tago T. Searching for diagnostic properties of novel fluorine-18-labeled D-allose. Ann Nucl Med. 2019;33:855–65.
12. Logan J, Fowler JS, Volkow ND, Wolf AP, Dewey SL, Schlyer DJ, et al. Graphical analysis of reversible radioligand binding from time-activity measurement applied to [*N*-^{11}C-methyl]-(–)-cocaine PET studies in human subjects. J Cereb Blood Flow Metab. 1990;10:740–7.
13. Logan J, Fowler JS, Volkow ND, Wang GJ, Ding YS, Alexoff DL. Distribution volume ratios without blood sampling from graphical analysis of PET data. J Cereb Blood Flow Metab. 1996;16:834–40.
14. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rather reliability. Psychol Bull. 1979;86:420–8.
15. Motulsky HJ, Brown RE. Detecting outliers when fitting data with nonlinear regression—a new method based on robust nonlinear regression and false discovery rate. BMC Bioinformatics. 2006;7:123.
16. Zhou X, Khanapur S, de Jong JR, Willemsen ATM, Dierckx RAJO, Elsinga PH, et al. In vivo evaluation of [^{11}C]preladenant positron emission tomography for quantification of adenosine A_2A receptors in the rat brain. J Cereb Blood Flow Metab. 2017;37:577–89.
17. Eckelman WC, Kilbourn MR, Mathis CA. Specific to nonspecific binding in radiopharmaceutical studies: it's not so simple as it seems! Nucl Med Biol. 2009;36:235–7.
18. Zhou X, Boellaard R, Ishiwata K, Sakata M, Dierckx RAJO, de Joung JR, et al. In vivo evaluation of ^{11}C-preladenant for PET imaging of adenosine A2A receptors in conscious monkey. J Nucl Med. 2017;58:762–7.
19. Ottoy J, De Picker L, Verhaeghe J, Deleye S, Wyffels L, Kosten L, et al. 18F-PBR111 PET imaging in healthy controls and schizophrenia: test–retest reproducibility and quantification of neuroinflammation. J Nucl Med. 2018;59:1267–74.
20. Mase H, Tanaka A, Sasou T, Nozaki A, Asai S, Miyachi H. Simultaneous measurement of concentrations of caffeine and ibuprofen in blood by LC-MS/MS and pharmacokinetic analysis after oral intake of canned coffee. Jpn J Med Technol. 2015;64:413–20.
21. Jacobson KA, Gao Z-G, Matricon P, Eddy MT, Carisson J. Adenosine A_{2A} receptor antagonists: from caffeine to selective non-xanthines. Br J Pharmacol. 2020. https://doi.org/10.1111/bph.15103.
22. White JR Jr, Padowski JM, Zhong Y, Chen G, Luo S, Lazarus P, et al. Pharmacokinetic analysis and comparison of caffeine administered rapidly or slowly in coffee chilled or hot versus chilled energy drink in healthy young adults. Clin Toxicol (Phila). 2016;54:308–12.
23. Caffeine for the Sustainment of Mental Task, Performance: Formulations for Military Operations. Committee on Military Nutrition Research, Food, and Nutrition Board. National Academy of Sciences; [cited 2015 Oct]. Available from: https://iom.nationalacademies.org/Reports/2001/Caffeine-for-the-Sustainment-of-Mental-Task-Performance-Formulations-for-Military-Operations.aspx. Accessed 8 June 2021.
24. Elmenhorst D, Meyer PT, Matusch A, Winz OH, Bauer A. Caffeine occupancy of human cerebral A_1 adenosine receptors: in vivo quantification with ^{18}F-CPFPX and PET. J Nucl Med. 2012;53:1723–9.