RESEARCH ARTICLE

Comparative test-retest variability of outcome parameters derived from brain $[^{18}\text{F}]$FDG PET studies in non-human primates

Sébastien Goutal, Nicolas Tournier, Martine Guillermier, Nadja Van Camp, Olivier Barret, Mylène Gaudin, Michel Bottlaender, Philippe Hantraye, Sonia Lavisse

1 Laboratoire des Maladies Neurodégénératives, CEA, CNRS, MIRCen, Université Paris-Saclay, Fontenayaux-Roses, France, 2 Laboratoire d’Imagerie Biomédicale Multimodale (BioMaps), CEA, CNRS, Inserm, Service Hospitalier Frédéric Joliot, Université Paris-Saclay, Orsay, France, 3 UNIACT, Neurospin, CEA, Université Paris-Saclay, Gif-sur-Yvette, France

Abstract

Introduction

Knowledge of the repeatability of quantitative parameters derived from $[^{18}\text{F}]$FDG PET images is essential to define the group size and allow correct interpretation. Here we tested repeatability and accuracy of different $[^{18}\text{F}]$FDG absolute and relative quantification parameters in a standardized preclinical setup in nonhuman primates (NHP).

Material and methods

Repeated brain $[^{18}\text{F}]$FDG scans were performed in 6 healthy NHP under controlled experimental factors likely to account for variability. Regional cerebral metabolic rate of glucose (CMRglu) was calculated using a Patlak plot with blood input function Semi-quantitative approaches measuring standard uptake values (SUV, SUV$\times$glycemia and SUVR (SUV Ratio) using the pons or cerebellum as a reference region) were considered. Test-retest variability of all quantification parameters were compared in different brain regions in terms of absolute variability and intra-and-inter-subject variabilities. In an independent $[^{18}\text{F}]$FDG PET experiment, robustness of these parameters was evaluated in 4 naive NHP.

Results

Experimental conditions (injected dose, body weight, animal temperature) were the same at both imaging sessions ($p > 0.4$). No significant difference in the $[^{18}\text{F}]$FDG quantification parameters was found between test and retest sessions. Absolute variability of CMRglu, SUV, SUV$\times$glycemia and normalized SUV ranged from 25 to 43%, 16 to 21%, 23 to 28%, and 7 to 14%, respectively. Intra-subject variability largely explained the absolute variability of all quantitative parameters. They were all significantly correlated to each other and they were all robust. Arterial and venous glycemia were highly correlated ($r = 0.9691; p < 0.0001$).
Conclusion

$^{[18]}$FDG test-retest studies in NHP protocols need to be conducted under well-standardized experimental conditions to assess and select the most reliable and reproducible quantification approach. Furthermore, the choice of the quantification parameter has to account for the transversal or follow-up study design. If pons and cerebellum regions are not affected, non-invasive SUVR is the most favorable approach for both designs.

Introduction

Positron emission tomography (PET) imaging using 2-deoxy-2-$^{[18]}$fluoro-D-glucose (FDG) remains so far the most accurate in vivo method for the investigation of regional glucose metabolism in the living brain. In the brain, glucose metabolism is tightly linked to neuronal and synaptic functions. $^{[18]}$FDG PET therefore provides a convenient tool to non-invasively monitor changes in neuronal activity in pathophysiological states. $^{[18]}$FDG PET is a translational imaging method, which can be performed either in humans or in animals going from rodents to large animal species such as nonhuman primates (NHP) [1]. Though invasive in rodents, quantitative $^{[18]}$FDG PET is only minimally invasive in NHPs, which facilitates its use as a close-human animal model. The high-resolution of preclinical PET scanners in particular allows determination of the level of glucose metabolism in discrete brain regions. Over the last four decades, $^{[18]}$FDG brain PET has mostly been used in the context of neurodegenerative diseases and brain tumor research [2]. As non-invasive technique, $^{[18]}$FDG PET has also been used to monitor the regional effects of potential CNS drugs and/or pharmacological challenges with repeated scans [3–6]. More recently, $^{[18]}$FDG brain PET was also applied to unravel brain networks in healthy and diseased brains [7]. Many quantification approaches exist to determine brain glucose metabolism, hence questioning the accuracy of the different $^{[18]}$FDG PET quantification methods.

In the nuclear medicine practice, $^{[18]}$FDG PET images are often analyzed qualitatively, using visual comparison of static images to identify regions with abnormal glucose uptake compared with normal surrounding brain tissue. PET data can also be interpreted from the time-activity curves (TAC) in selected brain regions which are usually reported as percentage of injected dose (%ID/mL) or standardized uptake value (SUV) versus time. SUV (g/mL) is defined as the tissue concentration (MBq/mL) divided by injected activity per body weight or lean body mass (MBq/g) [8] and represents a normalization approach reflecting the degree of metabolic activity in healthy or pathological tissues.

Peripheral glucose metabolism drastically impacts the plasma concentration of FDG and its brain kinetics, hence influencing the estimation of brain glucose metabolism. Many authors have discussed the multiple factors that affect the SUV index, such as kinetics of tracer uptake in the brain [9]. Methods exist to approach absolute quantification of brain glucose utilization or, at least, reduce the effect of unintended and non-specific variability of brain $^{[18]}$FDG PET data between conditions. Estimation of the cerebral metabolic rate of glucose (CMRglu) can be achieved using full kinetic modelling of $^{[18]}$FDG PET data using the irreversible two-compartment model or the graphical method of Patlak [10]. The outcome parameter integrates glycemia and $^{[18]}$FDG plasma kinetics (i.e the arterial input function), and is technically demanding as it requires arterial blood sampling at multiple time-points. This is invasive, uncomfortable and can even be sometimes considered unethical for the patient. Further simplification or normalization methods for $^{[18]}$FDG PET have therefore been proposed to
overcome this limitation and avoid arterial catheterization in patients [11–13]. However, many factors of variability still exist and full kinetic modelling remains the gold-standard method for brain [18F]FDG PET quantification, notably in preclinical settings.

Knowledge of the repeatability of estimated outcome parameters derived from [18F]FDG PET is particularly relevant for longitudinal follow-up studies assessing brain disease progression and/or therapeutic efficacy. The determination of group sizes is essential to [18F]FDG PET studies design for correct interpretation of changes measured during follow-up. Also, it is essential to control for experimental factors so as to minimize additional sources of variability on data measurement, analysis or interpretation and their respective impact on overall test-retest variability during repeated brain [18F]FDG PET experiments.

Standardization of [18F]FDG PET imaging has previously been discussed in the clinical context [14]. Preclinical test-retest studies of cerebral [18F]FDG PET (with arterial blood sampling) have been performed by various groups but in rodents only [15, 16]. To the best of our knowledge, no preclinical study in NHP has yet reported the evaluation of the repeatability and accuracy of the different [18F]FDG PET quantification approaches in a standardized preclinical setup. In the present study, we performed repeated brain [18F]FDG PET acquisition in six healthy NHP using a standardized procedure to control for experimental factors. First, the test-retest variability of PET data metabolism outcome parameters was assessed, and compared for accurate interpretation of brain [18F]FDG images. Finally, we evaluated the robustness of the outcome parameters by considering their respective confidence interval on an independent group of naive NHP.

Material and methods

Animals and housing

All primate experiments performed in this study were approved by the local ethics committee Comité d’Ethique en Experimentation Animale, CETEA n°44 (authorization #12_074) and the Ministry of Higher Education, Research and Innovation (MESRI) ((APAFIS#389) 2015032716213569.02.), were conducted according to European regulations (EU Directive 2010/63) and in compliance with Standards for Humane Care and Use of Laboratory Animals of the Office of Laboratory Animal Welfare (OLAW–n° A5826-01) in a facility authorized by local authorities (authorization n° B92-032-02).

Animals were originally obtained from approved breeding facilities and vendors (Noveprim; Ferney S.E.; Mahebourg Mauritius Island). The main study used six adult male cynomolgus macaques (Macaca fascicularis) of about 4 years of age (4.2 ± 0.8 yrs) and 5.8 ± 1.1 kg average weight. The robustness study included four adult cynomolgus male primates of about 4 years of age (4.2 ± 0.7 yrs) and 4.5 ± 0.9 kg average weight. Animals were housed in pairs or social groups depending on compatibility in a caging system with multiple compartments that were cleaned at least once every two weeks while sawdust was changed twice a week.

Animals were housed under a 12-hour light-dark cycle, temperature was maintained at 22 ±1°C and humidity at 50% in the animal quarters. Animals were fed with a commercial monkey chow, supplemented daily with fruits and vegetables changed on a daily basis as part of the enrichment program and drinking water was changed daily and was available ad libitum. Throughout the study, animals were checked daily by the technicians to evaluate their physical and clinical condition, food and water consumption. Environmental enrichment activities included grooming contact, perches, opaque and colored cage panels, and access to additional environment enrichment devices such as mirrors, swings, foraging devices, and manipulation and feeding toys. All imaging procedures were performed under ketamine/xylazine anesthetic induction (10 and 1mg/kg, respectively, intramuscular injection) and animals were then
maintained anesthetized using an intravenous (i.v.) infusion of propofol (1mL/kg/h) to minimize pain. Animals were not euthanized at the end of these experiments. None of the animals became severely ill during the course of the study or none required euthanasia prior to their experimental endpoint.

**Magnetic resonance imaging**

Magnetic resonance imaging (MRI) acquisition procedure was previously described [1]. Briefly, MRI was performed on each NHP in order to delineate cerebral regions on co-registered PET images. They were placed in the magnet in a sphinx position with the head fixed in a stereotactic MRI-compatible frame (M2E, France) and heated by a hot air flux with their temperature and respiration parameters monitored remotely.

Acquisitions were performed on a horizontal 7T Agilent scanner (Palo Alto, CA, USA) equipped with a gradient coil reaching 100mT/m (300μs rise time). A surface coil (RAPID Biomedical GmbH, Rimpar, Germany) was used for transmission and reception. T2-weighted images were acquired using a high-resolution 2D fast spin-echo sequence (469×469 μm² in-plane resolution, 1 mm slice thickness, 70 slices), with echo time TE/ Repetition time TR = 20/8000 ms, 5 echoes, effective TE = 52.5 ms and acquisition time Tacq = 43 min.

**PET imaging**

**PET acquisition.** [18F]FDG PET imaging was performed in all NHPs after overnight fasting with free access to water. Animals were initially sedated with the ketamine/xylazine mixture and anesthetized under propofol starting with an i.v. bolus followed by infusion in a saphenous vein. The other saphenous vein was catheterized for venous sampling. The femoral artery opposite to the tracer/propofol injection site was cannulated for arterial blood sampling. Primates were wrapped in a heat blanket to maintain body temperature throughout the protocol. Blood pressure, heart rate, respiration rate and oxygen saturation were remotely monitored and registered continuously before and during the acquisition.

Animals were placed in a PET dedicated stereotactic-like animal holder to immobilize the head by ear bars. PET imaging was performed on a Concorde Focus220 microPET scanner providing high sensitivity and high spatial resolution (FHWM = 2 mm) (Siemens, TN, USA). A 17-min-transmission scan (using a 57Co point source) was performed prior to PET acquisition to allow for attenuation correction, followed by a 60 min dynamic emission scan acquired in 2D mode. Data acquisition started simultaneously with the i.v. bolus injection of [18F]FDG (into the same vein as for propofol injection) using an automated injection pump at a rate of 4 mL/min during 30 seconds. PET images were corrected for radioactive decay, scatter, attenuation and detector inhomogeneity, were rebinned in 27 frames (4×15, 4×30, 2×60, 5×120, 3×300, 9×600 sec.) and reconstructed as previously described (OSEM-2D) [17].

To estimate the test-retest variability, a second scan was performed under exact same conditions after an interval of 27 ± 6 days. The time interval was selected to reduce stress induced by close, repeated food deprivation, anesthesia and animal handling. Overall average of [18F]FDG injected dose (ID) was of 169.2 ± 9.0 MBq.

Arterial blood samples were manually drawn from the femoral artery just before and during each PET acquisition to establish the kinetics in arterial plasma of [18F]FDG (arterial input function, AIF) and glucose concentration. After [18F]FDG injection, 2 ml of arterial blood was sampled every 15 s in the first 2 min and then at 2.5, 3, 5, 10, 20, 40 and 60 min after the injection. Blood samples were centrifuged and 0.2 mL plasma samples were counted in a gamma counter (Wizard², PerkinElmer) that was cross-calibrated with the PET scanner. Area under the curve of AIF was calculated from 0 to 60 minutes post injection.
Arterial glycemia (Gly) and venous blood samples were withdrawn simultaneously just before [18F]FDG injection. Both measurements were compared at test and retest sessions. All glycemia samples were measured in duplicate with a GM9 glucose analyser (Analox, Stourbridge, UK), which was calibrated at 5 mmol/L using standard solutions. The bias and repeatability of the GM9 analyzer was determined from 10 measures of a glucose standard.

The four additional and independent animals were scanned once under the same conditions so as to address the robustness of the quantification methods.

**PET image analysis.** Time frames were summed to create an integrated image for automatic co-registration with the T2-weighted MRI using a dedicated PMOD software (PMOD Technologies, version 3.6, Zurich, Switzerland). Anatomical regions of interest (ROI) were automatically delineated on individual MRI images using the CIVM atlas from the Primatologist segmentation pipeline for the macaque brain [18] and were as follows: thalamus, globus pallidus, caudate, putamen, global cortex, midbrain, pons, cerebellum, frontal and occipital cortices. Kinetics of the radioactivity of each ROI was then extracted from the dynamic PET data as a time-activity curves (TACs). For each animal, the same MRI segmentation was applied to extract TACs from test and retest PET images.

The Patlak graphical analysis was applied to the regional TACs in PMOD to quantitatively estimate the cerebral metabolic rate of glucose (CMRglu) under the hypothesis that the tissue includes a compartment in which the tracer is irreversibly entrapped [10, 19]. Linearization started at 30 min and the lumped constant representing the ratio of FDG utilization to actual glucose utilization in the brain was set to 0.34 [20, 21].

In each ROI, mean SUVs (g/mL) were calculated from the plateau of [18F]FDG brain kinetics, between 30 and 60 min post-injection. Additional SUV values were calculated accounting for glycemia (SUVxGly; mmol/L) [22]. Normalization to a “reference” brain region was performed using SUV uptake value in the pons (SUVRpons), or in the cerebellum (SUVRcerebellum) as already reported [23, 24].

In total, five quantification parameters were computed and used for statistical analyses: CMRglu, SUV, SUVxGly, SUVRpons and SUVRcerebellum.

**Statistical analyses**
Data are presented as mean ± SD unless otherwise indicated. Statistical analysis was performed using the R project free software version 3.3.1 (https://cran.r-project.org/).

The homogeneity of experimental conditions (body weight, temperature and injected dose at test and retest) were analyzed using Wilcoxon signed-rank test. Plasma exposure (AUCplasma) across animals were compared using a one-way analysis of variance (the distribution of residues was normal and homogeneous). Comparison of plasma exposure between test and retest was performed using a paired Wilcoxon rank sum test. The stability of the brain concentration of [18F]FDG was assessed using a two-way ANOVA where macaques and times were defined as explanatory variables (normal and homogeneous distribution of residues), followed by Tukey post hoc testing in TACs of all anatomical regions of interest. Comparison of the brain concentration of [18F]FDG was performed using a paired Wilcoxon rank sum test. Test-retest variability was estimated for each brain region and quantification parameter using the absolute variability across animals, defined as the absolute difference between values from the first and the second scan divided by the mean of both values (“AbsVar”, × 100 when expressed in %). The total variability was evaluated and explained by the contribution of intra-subject and inter-subject variabilities. Intra-subject variability was expressed 1)- through the coefficient of variation (CV%) and 2)- in relative proportion of the total variability. Inter-subject variability (expressed through the CV% and in relative proportion) was deduced from the total
variability, as proposed by Mermet [25]. Exceptional negative inter-subject variability referred to random fluctuations only and was then set to 0. Finally, a 95% confidence interval (95% CI; absolute and relative value) was built. In 3 (out of 68) cases, lower limits of the CI were negative and were thus arbitrarily set to 0. All correlations were performed using Spearman’s rank correlation coefficient.

The sample size \( n \) to estimate the cerebral metabolic rate of glucose consumption with a desired precision level, arbitrary set at 20%, was determined for each quantification parameter:

\[
\frac{t \times s}{\alpha}
\]

with \( n \) being the sample size, \( t \) the coefficient of confidence (= 1.96 for a 5% alpha risk), \( \alpha \) the acceptable imprecision of the mean (precision level of 20% of mean value) and \( s \), the estimated standard deviation of the population.

**Results**

**Homogeneity of experimental conditions**

Body weight (5.8 ± 1.4 kg and 6.3 ± 1.3 kg) and injected dose (171.8 ± 9.7 MBq and 166.6 ± 8.7 MBq) were not significantly different between the test and retest scans, respectively \( (p = 0.4633; p = 0.6905) \). During PET acquisition, body temperature remained stable for both test- and the retest scan (37.5 ± 0.4 and 37.3 ± 0.5, respectively, \( p = 0.53 \)).

**Arterial input function**

Bolus injection of \([^{18}\text{F}]\text{FDG}\) resulted in a sharp peak at 0.73 ± 0.07 min followed by a slow distribution phase. Plasma exposures (AUC\text{plasma} = 94.1 ± 13.7 SUV.min) were not significantly different across animals \( (p = 0.07) \), nor scan sessions \( (p = 1) \), with mean variability at test and re-test sessions of 14%CV and 15%CV, respectively.

**Glycemia**

The standard calibration analysis showed that determination of glucose levels in plasma obtained with GM9 was highly accurate \( (\text{bias} < 1\%) \) and reproducible \( (\text{CV} < 1\%) \). Glycemia measures from arterial samples were similar between test and retest conditions \( (4.5 ± 0.7 \, \text{mmol/L} \text{ and } 5.1 ± 1.5 \, \text{mmol/L}; \, p = 0.75) \). Glycemia measures from arterial and venous samples at both sessions were highly correlated \( (4.6 ± 1.3 \, \text{mmol/L} \text{ and } 4.7 ± 1.4 \, \text{mmol/L}, \text{respectively. } \text{Spearman } r = 0.9691; \, p < 0.0001, \text{ slope } = 0.972) \).

\([^{18}\text{F}]\text{FDG PET kinetics and quantification}\)

Visual inspection of the \([^{18}\text{F}]\text{FDG}\) brain PET images (in Bq/cc and in SUV) revealed qualitative difference between test and retest scans, which was confirmed in the TACs (S1 Fig). This difference appears to be global rather than regional as it was greatly reduced in SUVr images. TACs showed rapid uptake of the tracer in all regions with a maximal uptake 3 minutes after injection and a plateau starting at 30 minutes post-injection, with no significant difference between radioactivity values at 30, 40, 50 and 60 min, whatever brain regions \( (p > 0.1) \) (S1 Fig).

No significant difference of regional \([^{18}\text{F}]\text{FDG PET quantitative measures was found between the test and the retest sessions} \, (p > 0.1) \). All quantification approaches resulted in similar regional rank order, measures being the highest in the caudate, putamen and frontal
regions and the lowest in the pons. All regional mean SUVs, SUVR, CMRglu values are reported in Table 1.

**CMRglu estimate.** The statistical values of the regional CMRglu are reported in Table 1A. Overall, absolute variability was high and the range of CMRglu variability across regions was

| Table 1. Mean, confidence interval, total, intra- and inter-subject variabilities of outcome parameters between [18F]FDG test and retest scans. |
|---------------------------------------------------------------|
| A) CMRglu (μmol/min/100g)                                    |
| Mean ± sd, 95% CI, Abs var, IntraSV, InterSV                  |
| Thalamus (n = 12)                                            |
| 20.6±10.1 [0.0;42.8] 42.8±28.5 [86] 18.6 ±14 [14] 2.5±0.5 [1.4;3.5] |
| Globus pallidum (n = 0.40.0)                                 |
| 19.5±9.3 [0.0;4.8] 34.2±27.3 [67] 27.7 [33] 2.2±0.4 [1.3;3.1] |
| Caudate nucleus (n = 1.2;48.6)                               |
| 24.9±10.8 [1.2;18.6] 41.8±25.7 [89] 14.3 ±11 [11] 2.7±0.5 [1.5;3.8] |
| Putamen (n = 2;203)                                          |
| 26.2±10.9 [2.3;20.3] 42.3±22.0 [82] 17.8 ±18 [18] 2.8±0.5 [1.7;3.8] |
| Cortex (n = 1.5;37.5)                                        |
| 19.4±8.2 [1.5;37.5] 39.9±24.1 [80] 18.9 ±20 [20] 2.2±0.3 [1.5;2.9] |
| Midbrain (n = 0.0;41.3)                                      |
| 20.4±9.5 [0.0;41.3] 37.1±29.2 [85] 17.6 ±14 [14] 2.3±0.4 [1.4;3.3] |
| Pons (n = 15.8;69)                                           |
| 15.8±6.9 [0.7;30.9] 25.4±29.2 [75] 22.0 ±25 [25] 1.9±0.3 [1.3;2.6] |
| Cerebellum (n = 0.2;41.5)                                    |
| 21.1±9.2 [0.8;41.5] 33.2±25.6 [65] 26.1 ±35 [35] 2.3±0.4 [1.6;3.1] |
| Frontal cortex (n = 2.0;50.6)                                |
| 26.3±11.1 [2.0;50.6] 40.4±26.8 [90] 13.1 ±10 [10] 2.8±0.5 [1.5;4.0] |
| Occipital cortex (n = 0.0;36.8)                              |
| 17.0±9.0 [0.0;36.8] 38.3±29.6 [75] 26.6 ±25 [25] 2.1±0.4 [1.2;3.0] |
| B) SUV (g/mL)                                                |
| Mean ± sd, 95% CI, Abs var, IntraSV, InterSV                  |
| Thalamus (n = 12)                                            |
| 15.8±6.9 [0.7;30.9] 25.4±29.2 [75] 22.0 ±25 [25] 1.9±0.3 [1.3;2.6] |
| Globus pallidum (n = 0.40.0)                                 |
| 19.5±9.3 [0.0;4.8] 34.2±27.3 [67] 27.7 [33] 2.2±0.4 [1.3;3.1] |
| Caudate nucleus (n = 1.2;48.6)                               |
| 24.9±10.8 [1.2;18.6] 41.8±25.7 [89] 14.3 ±11 [11] 2.7±0.5 [1.5;3.8] |
| Putamen (n = 2;203)                                          |
| 26.2±10.9 [2.3;20.3] 42.3±22.0 [82] 17.8 ±18 [18] 2.8±0.5 [1.7;3.8] |
| Cortex (n = 1.5;37.5)                                        |
| 19.4±8.2 [1.5;37.5] 39.9±24.1 [80] 18.9 ±20 [20] 2.2±0.3 [1.5;2.9] |
| Midbrain (n = 0.0;41.3)                                      |
| 20.4±9.5 [0.0;41.3] 37.1±29.2 [85] 17.6 ±14 [14] 2.3±0.4 [1.4;3.3] |
| Pons (n = 15.8;69)                                           |
| 15.8±6.9 [0.7;30.9] 25.4±29.2 [75] 22.0 ±25 [25] 1.9±0.3 [1.3;2.6] |
| Cerebellum (n = 0.2;41.5)                                    |
| 21.1±9.2 [0.8;41.5] 33.2±25.6 [65] 26.1 ±35 [35] 2.3±0.4 [1.6;3.1] |
| Frontal cortex (n = 2.0;50.6)                                |
| 26.3±11.1 [2.0;50.6] 40.4±26.8 [90] 13.1 ±10 [10] 2.8±0.5 [1.5;4.0] |
| Occipital cortex (n = 0.0;36.8)                              |
| 17.0±9.0 [0.0;36.8] 38.3±29.6 [75] 26.6 ±25 [25] 2.1±0.4 [1.2;3.0] |
| E) SUV* Gly (mmol/L)                                         |
| Mean ± sd, 95% CI, Abs var, IntraSV, InterSV                  |
| Thalamus (n = 12)                                            |
| 12.2±4.6 [2.1;22.4] 27.8±31.4 [34] 39.4 ±100 [100] 0.0 |
| Globus pallidum (n = 1.9;19.8)                               |
| 10.9±4.1 [1.9;19.8] 27.3±30.9 [38] 38.8 ±100 [100] 0.0 |
| Caudate nucleus (n = 3.2;23.5)                               |
| 13.3±4.6 [3.2;23.5] 30.7±30.1 [36] 36.8 ±100 [100] 0.0 |
| Putamen (n = 3.5;24.2)                                       |
| 13.8±4.7 [3.5;24.2] 26.2±30.8 [36] 36.8 ±100 [100] 0.0 |
| Cortex (n = 3.2;18.6)                                        |
| 10.3±3.5 [3.2;18.6] 24.9±29.1 [33] 33.8 ±100 [100] 0.0 |
| Midbrain (n = 3.2;20.0)                                      |
| 11.7±4.2 [3.2;20.0] 22.8±27.6 [33] 33.0 ±100 [100] 0.0 |
| Pons (n = 3.5;23.8)                                          |
| 9.6±3.1 [3.8;16.3] 24.1±29.2 [33] 33.8 ±100 [100] 0.0 |
| Cerebellum (n = 3.2;20.0)                                    |
| 11.6±3.8 [3.2;20.0] 22.8±27.6 [33] 33.0 ±100 [100] 0.0 |
| Frontal cortex (n = 3.5;23.8)                                |
| 13.7±4.6 [3.5;23.8] 27.9±30.3 [36] 36.2 ±100 [100] 0.0 |
| Occipital cortex (n = 1.0;19.9)                              |
| 10.5±4.3 [1.0;19.9] 26.7±35.1 [43] 43.0 ±100 [100] 0.0 |

[.] refers to explained variance (EV). Estimates and absolute variability expressed as mean ± SD. InterSV = inter-subject variability; IntraSV = intra-subject variability.

https://doi.org/10.1371/journal.pone.0240228.t001
from 25.4 to 43%. Lowest variability (25.4%) was found in the pons region. No significant CMRglu difference was reported across regions and intra-subject variability explained between 65 and 90% of the total variability (see Fig 1).

**SUVs and SUVR.** All statistical measures of SUV and SUVR parameters are reported in Table 1B–1E.

**SUV.** Absolute variability of SUV data ranged from 16 to 21% with no regional distinction (global mean = 18.1 ± 1.7 g/mL) and was lower compared to CMRglu absolute variability. The intra-subject CV% (global mean = 17.5 ± 1.8%) was 2.6-fold higher than the inter-subject coefficient of variation in all anatomical ROIs (global mean = 6.7 ± 2.5%) and 2-fold lower than for CMRglu intra-subject variability). In other terms, intra-subject variability of SUV explained between 76 and 100% of the total variability while it explained 65 to 90% of the total variability for CMRglu (Fig 1).

**SUVR.** Normalization of SUV by a reference region decreased the absolute variability down to 7 to 14% for SUVRpons and SUVRcerebellum. Overall variability of SUVR was essentially explained by intra-subject variability ranging from 78 to 100% for SUVRcerebellum and almost 100% for SUVRpons across regions, whereas much lower inter-subject variability was estimated (lower than 5% for SUVRpons and SUVRcerebellum) (Fig 1).
SUV×Gly. Correction for arterial glycemia did not improve the absolute variability as compared to SUV: absolute variability ranged between 23 to 28% across brain regions. In all brain regions, intra-subject variability was higher while inter-subject variability was lower in comparison to SUV. Finally, the absolute variations of SUV×Gly (~ 26%) was lower than those obtained for CMRglu (~ 37%) but inter-subject variability was negligible in comparison to CMRglu.

**Correlations between the quantification parameters.** All quantification parameters were significantly correlated to each other: CMRglu correlated similarly to each of the other [18F] FDG PET measures but correlation was weaker with SUV (r = 0.35). All correlations and statistical significances are provided in Table 2.

**Sample size**
Precise estimation (mean±20%) of outcome parameters derived from brain [18F]FDG PET requires sample size of n = 22, n = 5 and n = 6 NHPs for CMRglu, SUV and SUV×Gly, respectively. On the other hand, SUVR had a low variability which allowed to reach this precision with only 2 NHP.

**Robustness of the outcome parameters**
The 95% confidence interval was determined on the test-retest dataset for each outcome parameter (Table 1). All SUV×Gly and CMRglu values of the robustness group lied within the confidence interval. In addition, 95% of the regional values of SUV, SUVRpons and SUVRcerebellum were found within the confidence interval (Fig 2).

**Discussion**
[18F]FDG has been widely used in clinical and preclinical studies for few decades but this study provides for the first time test-retest variabilities of different quantification parameters in NHP under standardized conditions. Many experimental settings can affect the estimation of glucose consumption in the brain using [18F]FDG PET [22, 26, 27]. In this test-retest study, we aimed at controlling most of these settings to study their impact on various outcome parameters derived from [18F]FDG PET images in NHP. [18F]FDG PET scans were performed on six healthy NHP under the following conditions: experiments were performed in male macaques only [16] with constant and monitored temperature throughout the experiments; the same anesthesia protocol and experimental procedures were applied [28]; all animals were fasted for the same time period before each imaging session [22]; the time interval between PET sessions was constant; the same reconstruction method was applied for all scans [9] and finally, automatic tracer injection and automatic image segmentation were performed to limit

|         | CMRglu | SUV | SUVRpons | SUVRcerebellum |
|---------|--------|-----|----------|----------------|
| SUV     | 0.35** |     |          |                |
| SUVRpons| 0.66****| 0.67****|          |                |
| SUVRcerebellum | 0.53**** | 0.69**** | 0.91**** |                |
| SUV×Gly | 0.66****| 0.67****| 0.68**** | 0.60***        |

**** p-value < 0.0001
*** p-value < 0.001 and
** p-value < 0.01 (n = 6 at test and at retest sessions).

https://doi.org/10.1371/journal.pone.0240228.t002
Fig 2. Robustness study. Grey boxes represent the 95%IC of each parameter. Four naive animals were considered for this study and regional estimates are represented for each animal and region.

https://doi.org/10.1371/journal.pone.0240228.g002
the investigator-induced variability. Under these highly standardized conditions, quantification of outcome parameters were expected to be largely reproducible and indeed no significant differences were observed in any of these parameters between test- and retest PET sessions.

Various groups have reported test-retest \(^{18}\text{F} \)FDG variability in preclinical studies but in rodents only. Sijbesma et al. [16] have focused on the rat gender difference in cerebral metabolism and reported non negligible absolute variations in different brain regions for CMRglu (mean of 13.4 and 19.6%), SUV (1.61 and 9.12%) and SUV\(_{\text{xGly}}\) (9.3 and 11.9%) in anesthetized male and female rats, respectively. Huang et al. estimated the CMRglu parameter and associated variability in two anesthetized rats with 6 repeated \(^{18}\text{F} \)FDG PET scans and arterial input function measurements [15]. They concluded that repeated PET procedure could be feasible in small anesthetized animals when implementing the surgical methods they developed.

In NHPs, arterial catheterization is relatively feasible and repeated blood sampling is well tolerated after a recovery period [29]. Therefore, NHPs are excellent models for full kinetic modeling of radioligands, allowing the comparison of alternative simplified measures with the gold-standard CMRglu calculated by the Patlak method. The CMRglu equation involves the influx rate of \(^{18}\text{F} \)FDG (\( K_i \)), the blood glucose concentration and the lumped constant, accounting for the differences in transport and phosphorylation between glucose and deoxyglucose [30]. Similar glucose levels in male cynomolgus macaques (of the same age as in our study) have been previously reported in few studies: Wu et al. [31] reported a glycemia range of 2 to 4 mmol/L and Guo et al. found glycemia values around 4.7 mmol/L [32]. Brain CMRglu estimates in the present study were also found in the expected and previously reported range [33] but encompassed non negligible variability. Inter and intra-subject variabilities could be partly due to anesthesia injected to the animals. Indeed, the effects of different anesthetic drugs on glucose metabolism in the human brain have been recently reported [34] showing that propofol decreased glucose metabolism compared to saline placebo. In preclinics, studies using isoflurane anesthesia reported a higher glucose level post-scan compared to pre-scan in rats, suggesting a different impact of isoflurane on glycemia [22, 35]. Furthermore, some studies have reported that ketamine combined with xylazine induces transient acute hyperglycemia lasting up to 25 minutes in rats [36, 37]. This was confirmed in the rhesus monkey where glucose levels were increased up to 60 minutes after ketamine/dexmedetomidine anesthesia as compared to ketamine only that did not elevate glucose level [38].

Ideally, metabolic brain studies should be performed on unanesthetized awake animals. But routine imaging of a conscious NHP is out of reach and the vast majority of PET studies in NHP are conducted in anesthetized animals. Therefore a standardized sedation protocol requires to maintain stable glycemia and hence reliable measurement of glucose metabolism in primate research. The test-retest procedure would have to be reiterated for each anesthesia condition to determine the intrinsic variability of each quantification parameter and select the most reliable and robust one(s) under experimental standardized conditions and for a given experimental design.

The choice of the processing and quantification methods has also to be considered to explain variability. In our study, the normalized SUV (SUVR) and especially the SUVR\(_{\text{cerebellum}}\) showed low test-retest variability, indicating a global rather than regional difference between PET sessions. The residual variability might be related in part to our processing pipeline, particularly to the potential PET/MRI misalignment remaining between the two PET sessions of each animal. However, we would expect this local variability to be small compared to other factors affecting more globally the test/retest variability of SUV and CMRglu parameters.

Efforts to decrease the intra- and inter-subject variability in \(^{18}\text{F} \)FDG brain PET imaging may help gaining statistical power to detect therapeutic effects or disease progression patterns over time within the same subject panel or between groups with a limited number of subjects.
In our experimental setup, we calculated that the sample size required to accurately determine brain glucose consumption was very different depending on the quantification method applied. Normalization of SUV by a reference region (such as pons or cerebellum) reduced the sample size down to 2 animals because of their much lower inter- and intra-individual variabilities. We noticed here that accounting for glycemia in SUV (SUV×Gly) did not reduce variability, in agreement with previous report [39].

We found that [18F]FDG PET outcome measures were significantly inter-correlated across regions and animals. Within-animal correlations (as its own control) were higher than global correlations, favouring individual follow-up experimental design. Nevertheless, significant global correlations (across all animals) between SUV/SUVR and CMRglu parameters underline the possible use of non-invasive procedures instead of arterial sampling to estimate change in brain glucose metabolism over time or treatments. Furthermore, we compared glycemia measurements from arterial and venous blood to avoid arterial collection in animals and found that both measurements were highly correlated. This means that glycemia may be measured in NHP either from venous or arterial blood as long as samples are collected similarly throughout the entire study.

Outcome quantification parameters of [18F]FDG images have to be selected upon the design of the imaging protocol. For cross-sectional studies, it is preferable to use a quantification parameter with a low inter-subject variability and for that, SUV, SUVRpons and SUVRcerebellum are excellent candidates. CMRglu showed a reasonable inter-subject variability as well (~20%). On the other hand, SUVRpons and SUVRcerebellum can only be considered in patients that do not suffer from a pathology affecting the cerebellum or the pons region. For longitudinal studies, a low intra-subject variability is needed to accurately follow-up each individual. SUVRpons and SUVRcerebellum fulfill also these criteria, but alternatively SUV can be considered if the cerebellum or the pons are involved in the patient’s pathology.

Conclusion

Altogether, this study suggests that preclinical [18F]FDG PET imaging requires highly controlled and standardized experimental conditions to minimize variability and reduce the size of animal cohorts. Definition of standardized protocols enables consideration of preclinical multicentric studies to further reduce animal cohorts in each research site. Appropriate choice of anesthetic drugs and outcome parameters is crucial and should be selected with respect to the study design with the aim to reduce variability and improve the sensitivity and accuracy of [18F]FDG PET images to detect changes in glucose brain metabolism.

Supporting information

S1 Fig. Representative time activity curves and SUV PET images. A)- Time activity curves of a representative animal in the pons and thalamus regions; B)- SUV PET images of the same animal at test and retest sessions.

(TIFF)

S1 File.

(XLSX)

Acknowledgments

Authors are thankful to Leopold Eymin and Sophie Lecourtois for animal handling and PET acquisition. They also thank Romina Aron-Badin for animal care support and MRI acquisitions.
Author Contributions

Conceptualization: Nicolas Tournier, Philippe Hantraye, Sonia Lavisse.
Formal analysis: Sébastien Goutal, Sonia Lavisse.
Investigation: Martine Guillermier, Nadja Van Camp, Mylène Gaudin, Sonia Lavisse.
Methodology: Martine Guillermier, Sonia Lavisse.
Project administration: Sonia Lavisse.
Resources: Philippe Hantraye.
Supervision: Nicolas Tournier, Philippe Hantraye, Sonia Lavisse.
Validation: Sébastien Goutal, Nadja Van Camp, Olivier Barret, Michel Bottlaender.
Visualization: Sébastien Goutal, Nicolas Tournier, Michel Bottlaender, Sonia Lavisse.
Writing – original draft: Sébastien Goutal, Nicolas Tournier, Olivier Barret, Michel Bottlaender, Sonia Lavisse.
Writing – review & editing: Sébastien Goutal, Nicolas Tournier, Martine Guillermier, Nadja Van Camp, Olivier Barret, Mylène Gaudin, Michel Bottlaender, Philippe Hantraye, Sonia Lavisse.

References

1. Lavisse S, Williams S, Lecourtois S, van Camp N, Guillermier M, et al. (2019) Longitudinal characterization of cognitive and motor deficits in an excitotoxic lesion model of striatal dysfunction in non-human primates. Neurobiol Dis 130: 104484. https://doi.org/10.1016/j.nbd.2019.104484 PMID: 31132407

2. Varrone A, Asenbaum S, Vander Borght T, Booij J, Nobili F, et al. (2009) EANM procedure guidelines for PET brain imaging using [18F]FDG, version 2. Eur J Nucl Med Mol Imaging 36: 2103–2110. https://doi.org/10.1007/s00259-009-1264-0 PMID: 19838705

3. Cieeren E, Casteels C, Goffin K, Koole M, Van Laere K, et al. (2018) Positron emission tomography imaging of cerebral glucose metabolism and type 1 cannabinoid receptor availability during temporal lobe epileptogenesis in the amygdala kindling model in rhesus monkeys. Epilepsia 59: 959–970. https://doi.org/10.1111/epi.14059 PMID: 29663353

4. Peng S, Ma Y, Flores J, Comfieldt M, Mitrovic B, et al. (2016) Modulation of Abnormal Metabolic Brain Networks by Experimental Therapies in a Nonhuman Primate Model of Parkinson Disease: An Application to Human Retinal Pigment Epithelial Cell Implantation. J Nucl Med 57: 1591–1596. https://doi.org/10.2987/jnumed.15.161513 PMID: 27056614

5. Tournier N, Saba W, Goutal S, Gervais P, Valette H, et al. (2015) Influence of P-Glycoprotein Inhibition or Deficiency at the Blood-Brain Barrier on (18)F-2-Fluoro-2-Deoxy-D-glucose ((18)F-FDG) Brain Kinetics. AAPS J 17: 652–659. https://doi.org/10.1208/s12248-015-9739-3 PMID: 25716150

6. Zhang X, Newport GD, Callicott R, Liu S, Thompson J, et al. (2016) MicroPET/CT assessment of FDG uptake in brain after long-term methylphenidate treatment in nonhuman primates. Neurotoxicol Teratol 56: 68–74. https://doi.org/10.1016/j.ntt.2016.06.005 PMID: 27307090

7. Sala A, Perani D (2019) Brain Molecular Connectivity in Neurodegenerative Diseases: Recent Advances and New Perspectives Using Positron Emission Tomography. Front Neurosci 13: 617. https://doi.org/10.3389/fnins.2019.00617 PMID: 31258466

8. Sarikaya I, Albatineh A, Sarikaya A (2019) Re-visiting SUV-weight and SUV-lean body mass in FDG PET studies. J Nucl Med Technol.

9. Buvat I (2007) Understanding the limitations of SUV. Médecine Nucléaire 31: 165–172.

10. Patlak CS, Blasberg RG, Fenstermacher JD (1983) Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. J Cereb Blood Flow Metab 3: 1–7. https://doi.org/10.1038/jcbfm.1983.1 PMID: 6822610

11. Sundar LK, Muzik O, Rischka L, Hahn A, Rausch I, et al. (2019) Towards quantitative [(18)F]FDG-PET/MRI of the brain: Automated MR-driven calculation of an image-derived input function for the non-invasive determination of cerebral glucose metabolic rates. J Cereb Blood Flow Metab 39: 1516–1530. https://doi.org/10.1177/0271678X18776820 PMID: 29790820
12. Vanzi E, Berti V, Polito C, Freddi I, Comis G, et al. (2016) Cerebral metabolic rate of glucose quantification with the aortic image-derived input function and Patlak method: numerical and patient data evaluation. Nucl Med Commun 37: 849–859. https://doi.org/10.1097/MNM.0000000000000512 PMID: 27058363

13. Jansen MH, Kloet RW, van Vuurden DG, Veldhuijzen van Zanten SE, Witte BI, et al. (2014) 18 F-FDG PET standard uptake values of the normal pons in children: establishing a reference value for diffuse intrinsic pontine glioma. EJNMMI Res 4: 8. https://doi.org/10.1186/2191-219X-4-8 PMID: 24472395

14. Antunovic L, Rodari M, Rossi P, Chiti A (2014) Standardization and quantification in PET/CT imaging: tracers beyond FDG. PET Clin 9: 259–266. https://doi.org/10.1016/j.cpet.2014.03.002 PMID: 25030389

15. Huang CC, Wu CH, Huang YY, Tzen KY, Chen SF, et al. (2017) Performing Repeated Quantitative Small-Animal PET with an Arterial Input Function Is Routinely Feasible in Rats. J Nucl Med 58: 611–616. https://doi.org/10.2967/jnumed.116.182402 PMID: 27789717

16. Sijbesma JWA, van Waarde A, Vallez Garcia D, Boersma HH, Slart R, et al. (2019) Test-Retest Stability of Cerebral 2-Deoxy-2-[(18)F]Fluoro-D-Glucose ([18]F FDG) Positron Emission Tomography (PET) in Male and Female Rats. Mol Imaging Biol 21: 240–248. https://doi.org/10.1007/s11307-018-1245-4 PMID: 29987619

17. Lavisse S, Inoue K, Jan C, Peyronneau MA, Petit F, et al. (2015) [18F]DPA-714 PET imaging of translocator protein TSPO (18 kDa) in the normal and excitotoxically-lesioned nonhuman primate brain. Eur J Nucl Med Mol Imaging 42: 478–494. https://doi.org/10.1007/s00259-014-2962-9 PMID: 25488184

18. Balbastre Y, Riviere D, Souedet N, Fischer C, Herard AS, et al. (2017) Primatologist: A modular segmentation pipeline for macaque brain morphometry. Neuroimage 162: 306–321. https://doi.org/10.1016/j.neuroimage.2017.09.007 PMID: 28899745

19. Patlak CS, Blasberg RG (1985) Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. Generalizations. J Cereb Blood Flow Metab 5: 584–590. https://doi.org/10.1038/jcbfm.1985.87 PMID: 4055928

20. Brooks RA, Hatazawa J, Di Chiro G, Larson SM, Fishbein DS (1987) Human cerebral glucose metabolism determined by positron emission tomography: a revisit. J Cereb Blood Flow Metab 7: 427–432. https://doi.org/10.1038/jcbfm.1987.86 PMID: 3497163

21. Kennedy C, Sakurada O, Shinohara M, Jehle J, Sokoloff L (1978) Local cerebral glucose utilization in the normal conscious macaque monkey. Ann Neurol 4: 293–301. https://doi.org/10.1002/ana.410040402 PMID: 103488

22. Deleye S, Verhaeghe J, wyffels L, Dedeurnevaerde S, Stroobants S, et al. (2014) Towards a reproducible protocol for repetitive and semi-quantitative rat brain imaging with (18) F-FDG: exemplified in a memantine pharmaceutical challenge. Neuroimage 96: 276–287. https://doi.org/10.1016/j.neuroimage.2014.04.004 PMID: 24736171

23. Buchert R, Wilke F, Chakrabarti B, Martin B, Brenner W, et al. (2005) Adjusted scaling of FDG positron emission tomography images for statistical evaluation in patients with suspected Alzheimer’s disease. J Neuroimaging 15: 348–355. https://doi.org/10.1177/1052682205280169 PMID: 16254400

24. Son YD, Cho ZH, Choi EJ, Kim JH, Kim HK, et al. (2014) Individually differentiated serotonergic raphe nuclei measured with brain PET/MR imaging. Radiology 272: 541–548. https://doi.org/10.1148/radiol.14131547 PMID: 24854972

25. Mermet JM (2012) Potential of accuracy profile for method validation in inductively coupled plasma spectrochemistry. Spectrochimica Acta (Part B: Atomic Spectroscopy) 76.

26. Boellaard R (2009) Standards for PET image acquisition and quantitative data analysis. J Nucl Med 50 Suppl 1: 115–20S.

27. Schiffer WK, Mirrione MM, Bieggon A, Alexoff DL, Patel V, et al. (2006) Serial microPET measures of the metabolic reaction to a microdialysis probe implant. J Neurosci Methods 155: 272–284. https://doi.org/10.1016/j.jneumeth.2006.01.027 PMID: 16519945

28. Itoh T, Wakahara S, Nakano T, Suzuki K, Kobayashi K, et al. (2005) Effects of anesthesia upon 18F-FDG uptake in rhesus monkey brains. Ann Nucl Med 19: 373–377. https://doi.org/10.1007/ BF03027401 PMID: 16164193

29. Diehl KH, Hull R, Morton D, Pfister R, Rabernampanina Y, et al. (2001) A good practice guide to the administration of substances and removal of blood, including routes and volumes. J Appl Toxicol 21: 15–23. https://doi.org/10.1002/jat.727 PMID: 11180276

30. Wienhard K (2002) Measurement of glucose consumption using [(18)F]Fluorodeoxyglucose. Methods 27: 218–225. https://doi.org/10.1016/s1046-2023(02)00077-4 PMID: 12183109
31. Wu D, Yue F, Zou C, Chan P, Alex Zhang Y (2012) Analysis of glucose metabolism in cynomolgus monkeys during aging. Biogerontology 13: 147–155. https://doi.org/10.1007/s10522-011-9364-1 PMID: 22057901

32. Guo S, Qian W, Du F, Bingdi W, Xiaoli W, et al. (2014) Proteinuria in Cynomolgus macaques (Macaca fascicularis) with Spontaneously Developed Metabolic Disorder and Diabetes: Transcriptome Analysis of Biopsy Kidney. Journal of Diabetes and Metabolism 5.

33. Harper AM (1990) Cerebral Blood Flow and Metabolism; Alexander Murray Harper SJ, editor: Manchester Univ Pr 110 p.

34. Laaksonen L, Kallioinen M, Langsjo J, Laitio T, Scheinin A, et al. (2018) Comparative effects of dexmedetomidine, propofol, sevoflurane, and S-ketamine on regional cerebral glucose metabolism in humans: a positron emission tomography study. Br J Anaesth 121: 281–290. https://doi.org/10.1016/j.bja.2018.04.008 PMID: 29935583

35. Vodovar D, Duchene A, Wimberley C, Leroy C, Pottier G, et al. (2018) Cortico-Am ygdala-Stria tal Activation by Modafinil/Flecainide Combination. Int J Neuropsychopharmacol 21: 687–696. https://doi.org/10.1093/ijnp/pyy027 PMID: 29635319

36. Saha JK, Xia J, Grondin JM, Engle SK, Jakubowski JA (2005) Acute hyperglycemia induced by ketamine/xyazine anesthesia in rats: mechanisms and implications for preclinical models. Exp Biol Med (Maywood) 230: 777–784.

37. Guarino MP, Santos AI, Mota-Carmo M, Costa PF (2013) Effects of anaesthesia on insulin sensitivity and metabolic parameters in Wistar rats. In Vivo 27: 127–132. PMID: 23239861

38. Vaughan KL, Szarowicz MD, Herbert RL, Mattison JA (2014) Comparison of anesthesia protocols for intravenous glucose tolerance testing in rhesus monkeys. J Med Primatol 43: 162–168. https://doi.org/10.1111/jmp.12104 PMID: 24697511

39. Prando S, Carneiro CG, Robilotto CC, Sapienza MT (2019) Comparison of different quantification methods for 18F-fluorodeoxyglucose-positron emission tomography studies in rat brains. Clinics (Sao Paulo) 74: e1273.