Measurement of PET Quantitative Bias *In Vivo*

Martin A Lodge[1], Wojciech Lesniak[1], Michael A Gorin[1,2], Kenneth J Pienta[2], Steven P Rowe[1,2], Martin G Pomper[1,2]

1. The Russel H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, MD, USA

2. The James Buchanan Brady Urological Institute and Department of Urology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Corresponding author: Martin Lodge, Johns Hopkins PET Center, 600 North Wolfe Street, Baltimore, MD 21287; Telephone 410 614 2416; Email mlodge1@jhmi.edu

Word count: 5380

Running title: PET quantitative bias *in vivo*.

Abstract

Quantitative imaging biomarkers are widely used in PET for both research and clinical applications, yet bias in the underlying image data has not been well-characterized. In the absence of a readily available reference standard for *in vivo* quantification, bias in PET images has been inferred using physical phantoms, even though arrangements of this sort provide only a poor approximation of the imaging environment in real patient examinations. In this study we used data acquired in patient volunteers to assess PET quantitative bias *in vivo*. Image-derived radioactivity concentrations in the descending aorta were compared with blood samples counted on a calibrated gamma counter. **Methods:** Ten patients with prostate cancer were studied using ^{18}F-DCFPyL PET/CT. For each patient, 3 whole-body PET/CT image series were acquired following a single administration of the radiotracer: shortly after injection as well as approximately 1 and 4 hours later. Venous blood samples were obtained at 8 time points over an 8 hour period and whole-blood was counted on a NaI gamma counter. A 10 mm diameter, 20 mm long cylindrical volume-of-interest was positioned in the descending thoracic aorta to estimate the PET-derived radioactivity concentration in blood (C_{PET}). A tri-exponential function was fit to the gamma counter blood data and used to estimate the radioactivity concentration (C_{gamma}) at the time of each PET acquisition. **Results:** C_{PET} and C_{gamma} were linearly related with R^2 = 0.985 over a range of relevant radioactivity concentrations. The mean difference between the PET and gamma counter data corresponded to 4.8 ± 8.6 % with the PET measurements tending to be greater. **Conclusions:** Human image data acquired on a conventional whole-body PET/CT system with a typical clinical protocol differed by an average of around 5 % compared to blood samples counted on a calibrated gamma counter. This bias may be partly attributable to residual uncorrected scatter or attenuation

correction error. These data offer an opportunity for the assessment of PET bias *in vivo* and provide additional support for the use of quantitative imaging biomarkers.

INTRODUCTION

Positron emission tomography (PET) has a rich history of quantitative methods (*1*). While quantitative imaging is a relatively recent phenomenon in other areas of radiology (*2*), PET has been quantitative from its inception. Studies of neuro-receptor binding, myocardial perfusion and tumor response assessment are just some examples of the many quantitative applications of PET. Quantitative data is usually regarded as being less subjective than qualitative image assessment. It is readily amenable to statistical analysis. It allows for graded characterization of processes that might not fall into a binary classification. And in some cases quantification provides information that is not discernible by visual inspection alone, such as myocardial perfusion studies in cases of balanced ischemia (*3*). For these reasons, quantitative methods are not only pervasive in PET research but have been widely incorporated into routine clinical imaging where parameters such as standardized uptake values (SUVs) are extensively employed.

PET allows for a wide range of quantitative techniques with different levels of complexity ranging from simple SUV to elaborate dynamic studies incorporating blood analysis and radiotracer kinetic modeling (*4*). In general, quantification is feasible because all of the key steps in PET image formation are linear (or approximately linear) and most of the effects that degrade the measured data can be corrected or minimized. Although spatial resolution varies slightly across the field-of-view, PET images are generally free from gross distortions. Sophisticated corrections for detector non-uniformity, geometrical issues, detector dead-time, random coincidences, scatter and attenuation have all been meticulously developed (*5*). If all things work as intended, the absolute value of the image voxels should have some meaning. Most PET-based biomarkers assume the voxel intensity reflects the local radioactivity concentration, which in turn reflects the properties of the radiopharmaceutical, the time since radiotracer administration and most importantly, the

characteristics of the patient. Surprisingly, the assumption that PET images reflect radioactivity concentration within the body has not been widely tested. A common concern is that quantitative measurements are not directly comparable between different scanners (*6*), even for regions that are not greatly affected by partial volume issues. Of course, a wealth of experience supports the assumption that PET images <u>do</u> reflect radioactivity concentration but the accuracy (strictly bias) with which radioactivity concentration can be quantified within the body is less clear.

A likely reason for the lack of data assessing bias for *in vivo* quantification is that the true radioactivity concentration within the body is generally not known. Extensive research has evaluated all aspects of the quantitative process but these evaluations have had to use surrogate markers such as phantoms (*7*) or computer simulations (*8*). Computer simulation allows fine control of the relevant parameters and the ability to compare measurements with ground truth. While these methods and results are convincing, there is a sense that the evaluations are not quite complete without data acquired on real scanners. In the absence of a readily available reference standard for *in vivo* quantification, the accuracy of PET images has had to be inferred using physical phantoms, even though arrangements of this sort provide only a poor approximation of the imaging environment in patient studies. Phantom-based evaluations may give an unrealistically optimistic impression of PET quantitative accuracy because they do not reflect the complexity of the scatter and attenuation distributions in real patients. In the present study, we have used data acquired in patient volunteers to assess PET quantitative accuracy *in vivo*. Image-derived radioactivity concentrations in the descending aorta were compared with blood samples counted on a calibrated gamma counter. In this way we were able to compare *in vivo* PET measurements with a reliable external reference.

MATERIALS AND METHODS

Data Acquisition

This prospective study was approved by the Johns Hopkins institutional review board and all participants provided written informed consent. Ten patients with prostate cancer were studied using the prostate-specific membrane antigen PET imaging agent ^{18}F-DCFPyL (*9*) as part of the larger OSPREY phase II/III clinical trial (ClinicalTrials.gov identifier: NCT02981368). Each patient had 3 whole-body PET/CT scans following a single administration of approximately 333 MBq (9 mCi) of ^{18}F-DCFPyL. PET acquisitions started shortly after radiotracer injection and approximately 1 and 4 hours later. Images were acquired from mid-thigh to vertex of skull using a Biograph mCT (Siemens Healthineers, Erlangen, Germany) which had a 21.8 cm axial field-of-view (*10*). Whole-body acquisition employed a step-and-shoot approach for 3 minutes per bed position, with adjacent bed positions overlapping by 10 cm. Patients were allowed to get off the PET/CT bed after each whole-body study and separate CTs were acquired at each of the 3 imaging sessions (120 kV, 64 mA, 0.8 pitch, 0.5 sec rotation time).

PET images were reconstructed using three-dimensional ordered-subsets expectation-maximization with time-of-flight, 2 iterations, 21 subsets and a Gaussian filter with 5 mm full-width-at-half-maximum (FWHM). Standard corrections were applied for deadtime, randoms (offline subtraction of delayed events), attenuation (CT based) and scatter (single scatter simulation with scatter estimates scaled to the tails of the projections). Resolution recovery was not incorporated into the reconstruction algorithm. Image voxel dimensions were 4 × 4 mm and had a 3.3 mm slice thickness. The system was calibrated according to manufacturer recommendations to ensure PET images were generated in units of Bq/mL. This process involved a combination of daily normalization and calibration adjustment with a [68]Ge phantom and annual

cross-calibration using an ^{18}F-filled, 20 cm diameter uniform cylinder phantom. The ^{18}F procedure was performed after replacement of the ^{68}Ge phantom and served to cross-calibrate the scanner to the local dose calibrator. Both phantom and patient ^{18}F activities were measured with the same dose calibrator (CRC 15W, Capintec, Florham Park, NJ, USA) and geometry (including syringe type and sample holder), using calibration factor 447. This calibration factor was derived specifically for this instrument using a ^{68}Ge standard source (BM06S-68, Radqual Global Sources, Idaho Falls, ID, USA), traceable to the National Institute of Standards and Technology, which had itself been cross-calibrated for ^{18}F. The locally derived factor was identical to a previously published value for ^{18}F with this same model of instrument (*11*).

Blood samples were obtained from an indwelling venous catheter at 8 time points over an 8 hour period, approximately 5 min, 15 min, 30 min, 1 hr, 2 hr, 4 hr, 6 hr and 8 hr after radiotracer injection. Three hundred microliter samples of whole-blood were counted on a Wizard 2480 (Perkin Elmer, Waltham, MA, USA) gamma counter (*12*). The counting protocol used a 409-613 keV energy window and acquired data for 60 seconds per sample. The efficiency of the gamma counter for ^{18}F (32.7 %) was experimentally determined using the dose calibrator described above. This calibration procedure involved 300 µL samples of an ^{18}F stock solution, counted using the same protocol, test tubes and sample geometry used for the patient blood samples. Application of this efficiency factor enabled the gamma counter to measure ^{18}F radioactivity and therefore radioactivity concentration in absolute terms (Bq/mL). In this way, the dose calibrator, the PET/CT scanner and the gamma counter were all cross-calibrated with respect to each other and with reference to a national metrology laboratory.

Data Analysis

Radioactivity concentrations obtained from regional analysis of the PET images were compared with blood sample measurements as follows.

In each PET/CT image series, a cylindrical volume-of-interest (VOI) was defined in the descending thoracic aorta to estimate the PET-derived radioactivity concentration in blood (C_{PET}). These VOIs were 10 mm diameter and 20 mm long and were automatically positioned using commercial software (Syngo Via, Siemens Healthineers, Erlangen, Germany). The mean radioactivity concentration of all voxels within the VOI was recorded and the measurement time was taken to be the acquisition time of a slice passing through the region. As adjacent bed positions in a multi-bed acquisition overlap each other, a single slice would typically be formed from data acquired at two different bed positions, in which case interpolation was used to estimate the acquisition time. VOI measurements from each of the 3 whole-body studies were decay-corrected back to radiotracer injection time.

As blood sample collection times did not necessarily coincide with PET VOI measurement times, an analytic model was fit to the gamma counter data for each patient. Gamma counter measurements were decay-corrected back to injection time and plotted as a function of sample collection time. A tri-exponential model was fit to the gamma counter data using least squares minimization. The model was optimized for each patient and used to estimate the radioactivity concentration (C_{gamma}) at the time of each PET measurement.

In this way 3 pairs of data were generated for each patient, corresponding to the image-derived (C_{PET}) and gamma counter-derived (C_{gamma}) blood radioactivity concentration measurements at each scan. The extent of the agreement between the two methods was assessed using a Bland-Altman approach (*13*). The difference (*d*) between corresponding data points was calculated as,

$$d = C_{PET} - C_{gamma} \qquad (1)$$

and the relative difference (D) was calculated as,

$$D = 100 \times (C_{PET} - C_{gamma}) / 0.5(C_{PET} + C_{gamma}) \qquad (2)$$

The difference data (d) were plotted as a function of the mean, $0.5(C_{PET} + C_{gamma})$, and Kendall's tau was used to test for proportionality. When the difference (d) was dependent on the magnitude of the measurement, the relative difference (D) was used for subsequent analysis. If these data were consistent with the normal distribution (Shapiro-Wilk test), 95 % limits of agreement were defined as,

$$95\text{ \% limits} = \text{mean}(D) \pm 1.96 \times \text{SD}(D) \qquad (3)$$

Where mean(D) and SD(D) indicates the mean and standard deviation of D respectively. Statistical analysis was performed using SPSS Statistics version 25 (IBM, Armonk, NY, USA) and a P value of 0.05 was taken as the threshold for significance.

Phantom Experiments

Throughout the period of patient accrual, PET/CT phantom data were acquired on a quarterly basis. The purpose of these phantom studies was to monitor the accuracy of scanner calibration and to assess the partial volume effect for a cylindrical insert, comparable in size to the descending thoracic aorta. The American College of Radiology PET phantom (*14*) was prepared with an aqueous solution of ^{18}F and scanned with the same acquisition and reconstruction protocol used for patient imaging. A VOI in a large uniform region was used to measure quantitative accuracy and an SUV of 1.0 would indicate perfect agreement with the expected value, based on the dose calibrator. SUV was calculated using the mass of water in this phantom compartment which was measured using an accurate balance. In addition, a 25 mm diameter cylindrical insert within the phantom was filled with a known radioactivity concentration, 2.5 times that of the background, and used to assess the partial volume effect. A 10 mm diameter cylindrical VOI was placed in the center of the insert and the ratio of the mean VOI measurement to the known radioactivity concentration was defined as the recovery coefficient, expected to be 1.0 in cases of negligible partial volume underestimation. This approach was intended to mirror patient image analysis in terms of the dimensions of the VOI and the size and shape of the descending thoracic aorta.

RESULTS

Between October 2017 and March 2018, 10 patients underwent ^{18}F-DCFPyL PET/CT with data collection as described above. The mean injected activity was 337 ± 11 MBq. Figure 1 shows example images for a particular patient, with the position of the aorta VOIs clearly displayed. Figure 2 shows corresponding gamma counter measurements, the resulting tri-exponential model

fit and PET VOI measurements. No data were lost or compromised and, in all, 30 pairs of quantitative measurements were available for statistical analysis.

Figure 3 shows that the PET and gamma counter data were linearly related ($R^2 = 0.985$) over a range of relevant radioactivity concentrations. Figure 4A shows the difference data (d) were proportional to the mean (Kendall's tau = 0.338, $P = 0.009$). However when the data were expressed in relative units (D), the proportionality was not significant (Kendall's tau = 0.149, $P = 0.246$). In other words, bias expressed in percentage form was approximately constant across a range of radioactivity concentrations (see Figure 4B). The mean bias (the mean of D over all patients and studies) corresponded to 4.8 ± 8.6 %, with the PET measurements tending to be greater. These data (D) were not normally distributed (Shapiro-Wilk test, $P = 0.03$). However for a subset of the data in a clinically relevant range above 5,000 Bq/mL (corresponding to an SUV> 1 in a 74 kg patient following a 370 MBq administration) the relative difference data (D) were normally distributed (Shapiro-Wilk test, $P = 0.38$). In this range, the mean bias was 5.2 ± 4.0 % with 95 % limits of agreement of -2.6 % and +13.0 %.

Four phantom studies were performed around the study period. The mean SUV in the uniform region was 0.993 ± 0.005, indicating high accuracy and stability with this phantom arrangement. The mean recovery coefficient for the 25 mm insert was 1.002 ± 0.014, indicating negligible partial volume underestimation for a 10 mm VOI and a cylindrical object with a diameter of at least 25 mm. For comparison, the mean diameter of the descending thoracic aorta at the level of the VOI was 27.4 ± 2.4 mm, based on measurements from the 30 patient CT images.

DISCUSSION

In this study we assessed the quantitative accuracy of PET images in real human studies. More specifically, we compared PET measurements of the radioactivity concentration in the descending thoracic aorta to blood samples counted on a gamma counter, carefully calibrated with reference to a national metrology laboratory. On average the PET image data were 4.8 ± 8.6 % greater than the gamma counter data. These results indicate the bias that should be expected in basic PET measurements and confirms the relatively high accuracy that can be achieved in human imaging.

Early work establishing the foundations for PET quantification of radioactivity concentration were based on phantom experiments (*15*) and subsequently supported by *in vivo* measurements in animals. Using a similar approach to the one used in the present study, PET and gamma counter blood measurements were shown to be highly correlated after partial volume correction based on organ dimensions measured at post-mortem (*16*). The importance of partial volume correction reflects the limited spatial resolution of the PET systems available at that time, estimated to be approximately 16 mm FWHM. In contrast, the modern PET system used in our work had an effective spatial resolution of approximately 7 mm FWHM (*17*) and the partial volume effect was not expected to be a significant source of error in our aorta VOI measurements. This was confirmed by phantom experiments that involved a 25 mm cylindrical insert of approximately the same size as the descending aorta (27 mm) at the level of the VOI measurement. Note that the bias results reported in the present paper apply to large organs for which partial volume errors can be neglected. It is worth emphasizing that much greater bias should be expected for smaller lesions, less than roughly three times the effective FWHM.

The current assessment of the quantitative accuracy of PET for *in vivo* human imaging is particularly relevant as it was performed on a modern PET/CT system. Modern scanners

incorporate many components that could potentially degrade quantitative accuracy including three-dimensional data acquisition, scintillators with intrinsic radioactivity, iterative reconstruction, CT-based attenuation correction and scatter correction models that include various assumptions. While each aspect has been tested individually, the overall effect on quantitative accuracy has not been extensively studied. Most likely this is due to the lack of a readily available reference standard for *in vivo* quantification. Here we assumed that the radiotracer activity concentration in samples of whole-blood from a vein in the arm was comparable to arterial blood in the aorta. Unlike for ^{18}F-FDG where arterial and venous concentrations differ substantially as a function of time (*18*, *19*), this is a very reasonable assumption for ^{18}F-DCFPyL (supplemental data in *20*). In previous work by a separate group (*21*, *22*), radioactive urine samples were used as a reference for assessing bias *in vivo*. PET and PET/CT images were found to underestimate radioactivity concentration by 7-12 %. Note this bias is in the opposite direction to our own results and may be related to the extremely high radioactivity concentration in the bladder, possibly causing incomplete convergence of the iterative reconstruction algorithm and other problems.

The discrepancy of around 5 % between our PET and gamma counter measurements could be due to a combination of various technical factors, although many of these sources of error are not expected to have a large impact. The phantom data indicated that partial volume is not likely to be a major effect. The related effect of signal overestimation due to proximity to other organs with a higher radioactivity concentration (spillover) is also expected to be negligible as ^{18}F-DCFPyL is not taken up to any great extent in nearby organs. Detector dead-time was not expected to be a problem for either the PET scanner or gamma counter, based on the measured count rates. The effect of respiratory motion was expected to be minimal due to the long extent of the aorta in the craniocaudal direction and, in any event, would not readily explain the overestimated PET measurements. A perhaps more significant source of error is the possibility of imperfect scatter

correction. Scatter is a major problem in three-dimensional PET and in this study cohort the scatter correction algorithm indicated a scatter fraction of 38.6 ± 2.5 % at the level of the descending aorta. Perfect compensation in software cannot be assumed, especially in the chest where the scatter distribution is particularly complex, and may result in a residual uncorrected scatter component that increased the measured PET signal. Another possibility is that errors in the CT-based attenuation correction could also contribute to the observed bias. CT images are acquired at energies well below 511 keV and need to be scaled so as to reflect the attenuation appropriate for 511 keV annihilation radiation. Small errors in this process cannot be excluded, particularly as the chest includes a complex distribution of organs (bone, lung and soft tissue) with very different attenuating properties.

It is worth noting similarities between the present work and previous studies that have used PET images to derive input functions for radiotracer kinetic modeling. In studies of that kind, the goal was to estimate the time course of radiotracer activity in arterial plasma without the need for invasive blood sampling. Correction for the partial volume effect is essential for brain studies because the only available blood vessels in the field-of-view are small compared to the spatial resolution of the system. While the need for partial volume correction is clear, the problems involved with successful application have also been noted (*23*). When imaging included the chest, larger blood vessels were available such as the aorta which was often preferred over the left ventricular cavity due to spillover from the myocardial walls (*24*). Evaluation of these image-derived input functions did not necessarily involve direct comparison of the original image data with arterial blood samples counted on a gamma counter. For example, in many cases the image-derived data were scaled with reference to one or more blood samples (*25, 26*) so did not reflect the quantitative accuracy of the original image data. When input functions derived from unscaled PET data and arterial blood sampling were available for comparison, the assessment was

commonly characterized in terms of the resulting kinetic parameters (*24*, *27*), as opposed to a direct comparison of radioactivity concentration estimates.

In the present study, we directly compared image and sample-based measurements of blood radioactivity concentration. The purpose of this evaluation was to provide an estimate of the bias that can be expected in PET-derived quantitative metrics. Of course some PET metrics such as metabolic tumor volume are not dependent on calibration accuracy because they reflect the volume, rather than the concentration of radiotracer. Also if both the input function and tissue time-activity data are derived from the same image, as is commonly the case with cardiac flow quantification, errors due to scanner miscalibration cancel out and the need for accurate cross-calibration of equipment is avoided. However for many PET biomarkers including SUV, accurate image calibration is critical. In the present study we used a gamma counter, calibrated to a national metrology laboratory, as a reference. Both PET and gamma counter measurements will have errors but we expect the gamma counter to be accepted as a reliable reference method. Note that these results are strictly applicable only to the radiotracer used in this study, ^{18}F-DCFPyL. While similar bias might be expected with other ^{18}F-labeled radiotracers, more complex isotopes such as ^{124}I or ^{82}Rb would require further evaluation due to the presence of problematic prompt gammas. Different results might also be expected in other parts of the body but we suspect the chest may represent one of the more difficult settings because of its complex scatter and attenuation environment.

The results presented here have particular relevance for efforts to standardize quantitative PET biomarkers across different scanners and institutions (*28*). While patient test-retest studies provide information regarding repeatability (*29*), bias has been harder to characterize *in vivo*. Here we propose an approach that addresses this problem, allowing bias to be estimated directly from

patient images. Furthermore, since the amount of the radiotracer in blood was continuously changing, we were able to assess bias over a range of relevant radioactivity concentration levels. The method is generally applicable to other PET/CT models, as well as other PET devices. For example the method may be particularly useful for evaluating combined PET/MR systems where the absence of linear attenuation coefficient measurements makes attenuation correction problematic (*30*). Using this method we have shown that low bias is feasible with conventional clinical PET/CT scanners under normal operating conditions. Of course our specific results reflect the performance of only one particular scanner system and variability between different scanners and institutions is to be expected.

CONCLUSION

Human image data acquired on a conventional whole-body PET/CT system with a typical clinical protocol differed by an average of around 5 % compared to blood samples counted on a calibrated gamma counter. This relatively low bias is encouraging, particularly as it was measured in the complex imaging environment encountered in the chest, and may be partly attributable to residual uncorrected scatter or attenuation correction error. These data offer an opportunity for the assessment of PET bias *in vivo* and provide additional support for the use of quantitative imaging biomarkers.

DISCLOSURE

ACKNOWLEDGMENTS

KEY POINTS

QUESTION: PET quantitative metrics such as SUV assume image voxels reflect the local radioactivity concentration but to what extent is this true and what bias should be expected in human imaging ?

PERTINENT FINDINGS: Data obtained as part of a clinical trial showed PET measurements of radioactivity concentration were approximately 5 % higher than an external blood-based reference standard.

IMPLICATIONS FOR PATIENT CARE: These results indicate modern PET/CT systems are capable of low bias in human (as opposed to phantom) imaging, providing additional support for the use of quantitative PET biomarkers.

REFERENCES

1. Lammertsma AA. Forward to the past: The case for quantitative PET imaging. *J Nucl Med.* 2017;58:1019-1024.

2. Sullivan DC. Imaging as a quantitative science. *Radiology*. 2008;248:328-332.

3. Parkash R, deKemp RA, Ruddy TD, et al. Potential utility of rubidium 82 PET quantification in patients with 3-vessel coronary artery disease. *J Nucl Cardiol.* 2004;11:440-449.

4. Morris ED, Endres CJ, Schmidt KC, Christian BT, Muzic RF, Fisher RE. Kinetic modeling in positron emission tomography. *Emission Tomography: The Fundamentals of PET and SPECT.* San Diego: Academic; 2004:499-540.

5. Meikle SR, Badawi RD. Quantitative techniques in PET. *Positron Emission Tomography.* London: Springer; 2005:93-126.

6. Fahey FH, Kinahan PE, Doot RK, Kocak M, Thurston H, Poussaint TY. Variability in PET quantitation within a multicenter consortium. *Med Phys.* 2010;37:3660-3666.

7. Byrd D, Christopfel R, Arabasz G, et al. Measuring temporal stability of positron emission tomography standardized uptake value bias using long-lived sources in a multicenter network. *J Med Imaging.* 2018;5:011016.

8. Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. *J Nucl Med.* 2004;45:1519-1527.

9. Szabo Z, Mena E, Rowe SP, et al. Initial evaluation of [18F]DCFPyL for prostate-specific membrane antigen (PSMA)-targeted PET imaging of prostate cancer. *Mol Imaging Biol.* 2015;17:656-574.

10. Jakoby BW, Bercier Y, Conti M, Casey ME, Bendriem B, Townsend DW. Physical and clinical performance of the mCT time-of-flight PET/CT scanner. *Phys Med Biol.* 2011;56:2375-2389.

11. Zimmerman BE, Bergeron DE, Cessna JT. Impact of recent change in the National Institute of Standards and Technology standard for 18F on the relative response og 68Ge-based mock syringe dose calibrator standards. *J Nucl Med.* 2015;56:1453-1457.

12. Lodge MA, Holt DP, Kinahan PE, Wong DF, Wahl RL. Performance assessment of a NaI(Tl) gamma counter for PET applications with methods for improved quantitative accuracy and greater standardization. *EJNMMI Physics*. 2015;2:11.

13. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999;8:135-160.

14. MacFarlane CR. ACR accreditation of nuclear medicine and PET imaging departments. *J Nucl Med Technol.* 2006;34:18-24.

15. Eichling JO, Higgins CS, Ter-Pogossian M. Determination of radionuclide concentrations with positron CT scanning (PET): concise communication. *J Nucl Med.* 1977;18:845-847.

16. Henze E, Huang SC, Ratib O, Hoffman E, Phelps ME, Schelbert HR. Measurements of regional tissue and blood-pool radiotracer concentrations from serial tomographic images of the heart. *J Nucl Med.* 1983;24:987-996.

17. Lodge MA, Leal JP, Rahmim A, Sunderland JJ, Frey EC. Measuring PET spatial resolution using a cylinder phantom positioned at an oblique angle. *J Nucl Med.* 2018;59:1768-1775.

18. Wakita K, Imahori Y, Ido T, et al. Simplification for measuring input function of FDG PET: Investigation of 1-point blood sampling method. *J Nucl Med.* 2000;41:1484-1490.

19. van der Weerdt AP, Klein LJ, Visser CA, Visser FC, Lammertsma AA. Use of arterialised venous instead of arterial blood for measurement of myocardial glucose metabolism during euglycaemic-hyperinsulinaemic clamping. *Eur J Nucl Med.* 2002;29:663-669.

20. Jansen BH, Yaqub M, Voortman J, et al. Simplified methods for quantification of 18F-DCFPyL uptake in patients with prostate cancer. *J Nucl Med.* 2019;60:1730-1735.

21. Maus J, Hofheinz F, Schramm G, et al. Evaluation of PET quantification accuracy in vivo. Comparison of measured FDG concentration in the bladder with urine samples. *Nucl. Med.* 2014;53:67-77.

22. Maus J, Schramm G, Hofheinz F, et al. Evaluation of in vivo quantification accuracy of the Ingenuity-TF PET/MR. *Med Phys.* 2015;42:5773-5781.

23. Zanotti-Fregonara P, Chen K, Liouw JS, Fujita M, Innis RB. Image-derived input function for brain PET studies: many challenges and few opportunities. *J Cereb Blood Flow Metab.* 2011;31:1986-1998.

24. van der Weerdt AP, Klein LJ, Boellaard R, Visser CA, Visser FC, Lammertsma AA. Image-derived input functions for determination of MRGlu in cardiac 18F-FDG PET scans. *J Nucl Med.* 2001;42:1622-1629.

25. Dhawan V, Takikawa S, Robeson W, et al. Quantitative brain FDG/PET studies using dynamic aortic imaging. *Phys Med Biol.* 1994;39:1475-1487.

26. Schroeder T, Vidal Melo MF, Musch G, Harris RS, Venegas JG, Winkler T. Image-derived input function for assessment of 18F-FDG uptake by the inflamed lung. *J Nucl Med.* 2007;48:1889-1896.

27. de Geus-Oei LF, Visser EP, Krabbe PF, et al. Comparison of image-derived and arterial input functions for estimating the rate of glucose metabolism in therapy-monitoring 18F-FDG PET studies. *J Nucl Med.* 2006;47:945-949.

28. Kinahan PE, Perlman ES, Sunderland JJ, et al. The QIBA profile for FDG PET/CT as an imaging biomarker measuring response to cancer therapy. *Radiology*. 2020;294:647-657.

29. Lodge MA. Repeatability of SUV in oncologic 18F-FDG PET. *J Nucl Med.* 2017;58:523-532.

30. Hofmann M, Pichler B, Scholkopf B, Beyer T. Towards quantitative PET/MRI: a review of MR-based attenuation correction techniques. *Eur J Nucl Med Mol Imag.* 2009;36:93-104.
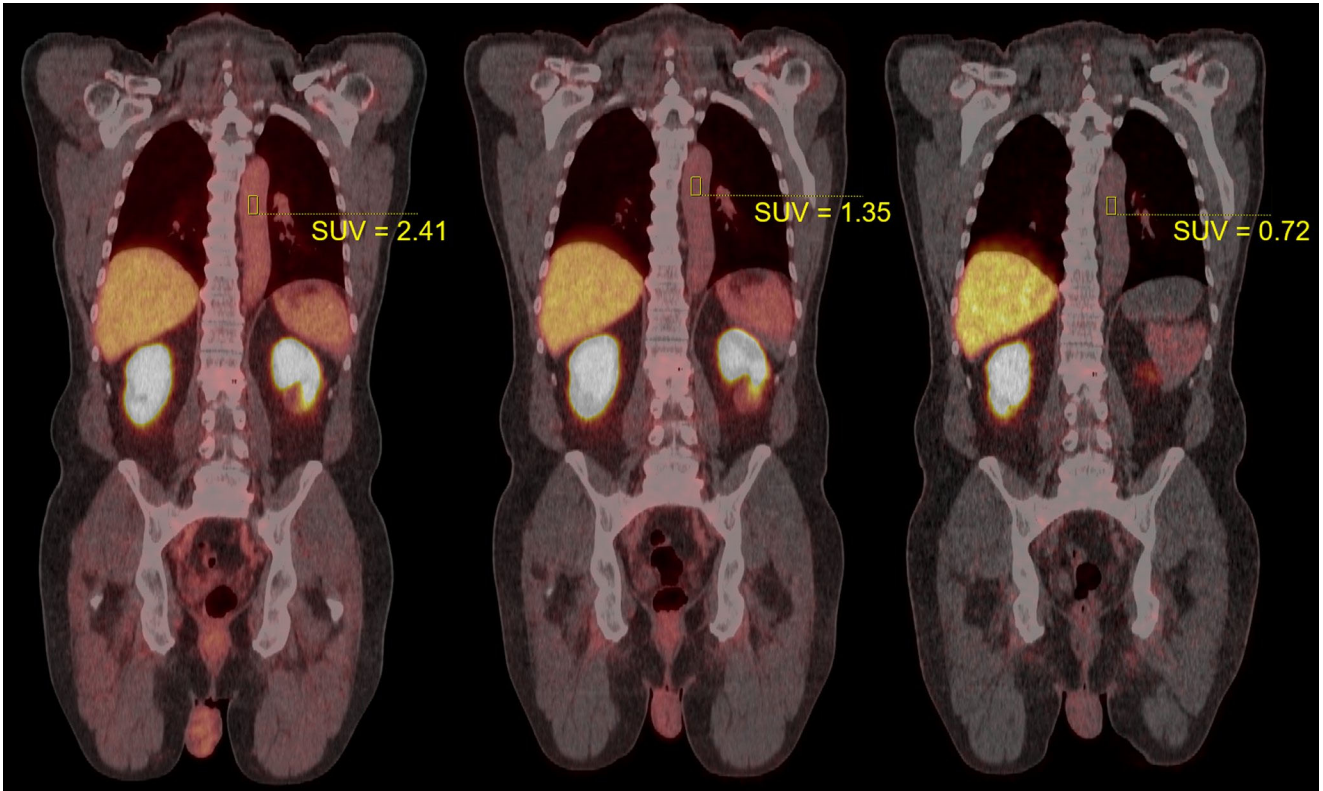
Figure 1. Example PET/CT images for a typical patient. PET data were acquired 24, 79 and 264 minutes post-injection (from left to right). Cylindrical VOIs can be seen in the descending aorta with SUV as indicated. The fused display shows PET in a "hotbody" color map, scaled from 0 – 20 SUV units in each image. The patient weighed 90.5 kg and was administered 339.7 MBq of ^{18}F-DCFPyL.
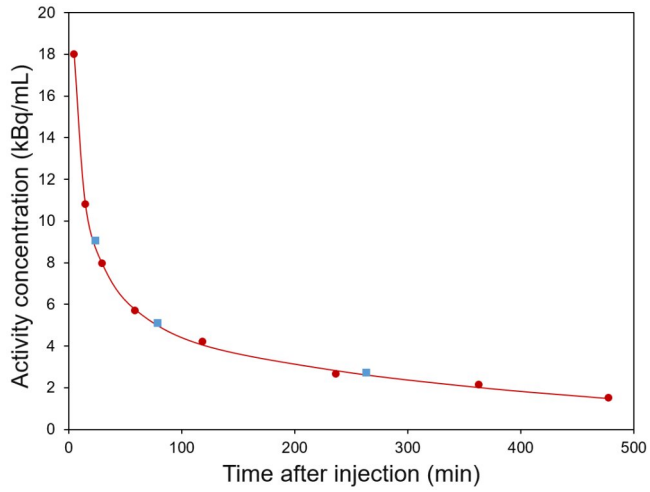
Figure 2. Example time-activity data for the patient shown in figure 1. Red circles indicate whole blood samples counted on a calibrated gamma counter. Blue squares indicate PET-derived data from VOIs in the descending aorta. The continuous red line is a tri-exponential model fit to the gamma counter data. All data were decay-corrected to the time of injection.
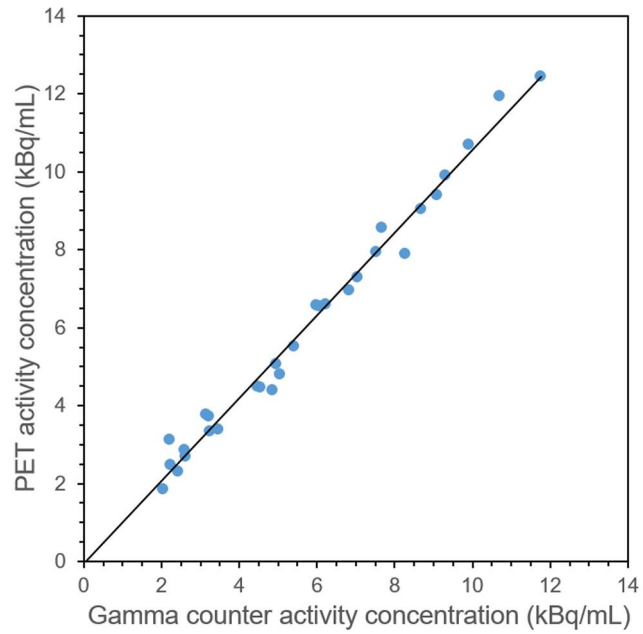
Figure 3. PET and gamma counter radioactivity concentration data for 10 patients, each with 3 corresponding measurements. The solid line indicates a linear fit to the data: y = 1.06 x – 64, R^2 = 0.985.
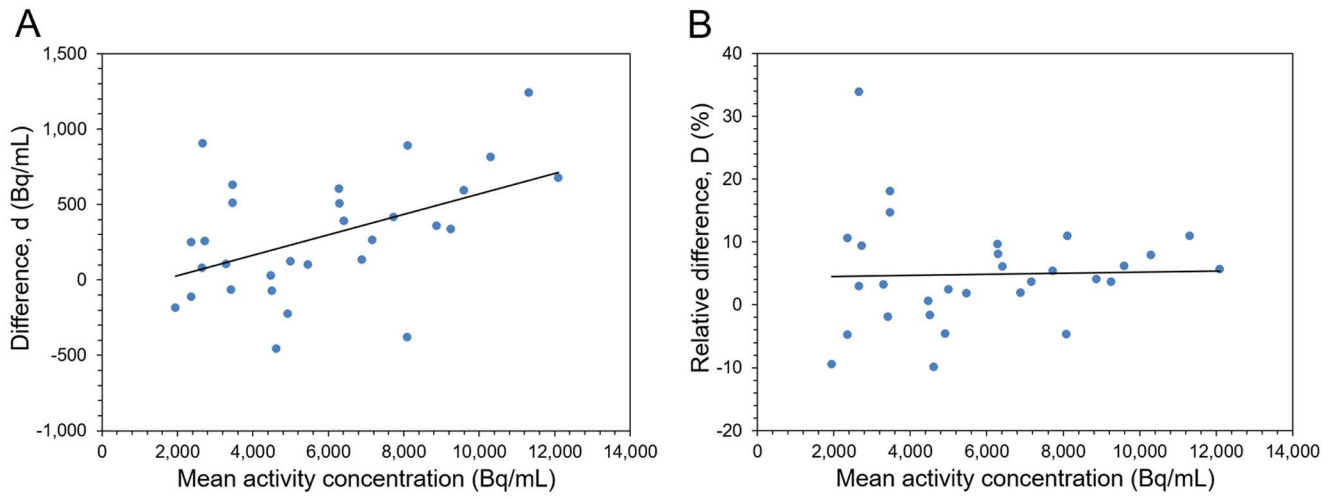
Figure 4. Bland-Altman plots showing the difference between PET and gamma counter measurements in units of radioactivity concentration (A) and in relative units (B). Solid lines indicate linear regression. The relative difference D was not proportional to radioactivity concentration and had a mean value of 4.8 ± 8.6 %.