

**Conditional Generative Adversarial Networks (cGANs) aided motion correction of dynamic 18F-FDG PET brain studies.**

Lalith Kumar Shiyam Sundar<sup>1\*</sup>, David Iommi<sup>1\*</sup>, Otto Muzik<sup>2</sup>, Zacharias Chalampalak<sup>3</sup>, Eva-Maria Klebermass<sup>4</sup>, Marius Hienert<sup>5</sup>, Lucas Rischka<sup>5</sup>, Rupert Lanzenberger<sup>5</sup>, Andreas Hahn<sup>5</sup>, Ekaterina Patarai<sup>6</sup>, Tatjana Traub-Weidinger<sup>4</sup>, Johann Hummel<sup>1</sup> and Thomas Beyer<sup>1</sup>.

1: QIMP Team, Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, Vienna, Austria

2: Department of Pediatrics, Children's Hospital of Michigan, The Detroit Medical Center, Wayne State University School of Medicine, Detroit, Michigan, USA

3: Service Hospitalier Frédéric Joliot, CEA, Inserm, CNRS, Univ. Paris Sud, Université Paris Saclay, Orsay, France

4: Division of Nuclear Medicine, Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Vienna, Austria

5: Department of Psychiatry and Psychotherapy, Medical University of Vienna, Vienna, Austria

6: Department of Neurology, Medical University of Vienna, Austria

**Corresponding Author**

Otto Muzik, PhD

Wayne State University School of Medicine Detroit, Michigan, USA

Telephone: +1(313) 993-2616

21 [otto@pet.wayne.edu](mailto:otto@pet.wayne.edu)

22 First authors contact information:

23 Lalith Kumar Shiyam Sundar, PhD and David Iommi MSc (\*Equal contribution)

24 QIMP Team, Center for Medical Physics and Biomedical Engineering, Medical University of  
25 Vienna, Vienna, Austria

26 Telephone: +43 1 40400 55450

27 [Lalith.shiyamsundar@meduniwien.ac.at](mailto:Lalith.shiyamsundar@meduniwien.ac.at)

28 **Running title:** cGAN-derived PET navigators

29 **Word count:** 4955

## 30 **Funding**

31 This work was supported by the Austrian Science Fund KLI482-B31. David Iommi and Zacharias  
32 Chalampalakis are funded by a Marie Skłodowska-Curie grant (No. 764458, EU Horizon 2020  
33 research and innovation program). Lucas Rischka is recipient of a DOC Fellowship of the Austrian  
34 Academy of Sciences at the Department of Psychiatry and Psychotherapy, Medical University of  
35 Vienna.

36

Immediate Open Access: Creative Commons Attribution 4.0  
International License (CC BY) allows users to share and adapt with  
attribution, excluding materials credited to previous publications.  
License: <https://creativecommons.org/licenses/by/4.0/>.  
Details: <http://jnm.snmjournals.org/site/misc/permission.xhtml>.



## Abstract

This work set out to develop a motion correction approach aided by conditional generative adversarial network (cGAN) methodology that allows reliable, data-driven determination of involuntary subject motion during dynamic 18F-FDG brain studies. **Methods:** Ten healthy volunteers (5M/5F,  $27 \pm 7$  years,  $70 \pm 10$  kg) underwent a test-retest 18F-FDG PET/MRI examination of the brain ( $N = 20$ ). The imaging protocol consisted of a 60-min PET list-mode acquisition contemporaneously acquired with MRI, including MR navigators and a 3D time-of-flight MR-angiography sequence. Arterial blood samples were collected as a reference standard representing the arterial input function (AIF). Training of the cGAN was performed using 70% of the total data sets ( $N = 16$ , randomly chosen), which was corrected for motion using MR navigators. The resulting cGAN mappings (between individual frames and the reference frame (55-60min p.i.)) were then applied to the test data set (remaining 30%,  $N = 6$ ), producing artificially generated low-noise images from early high-noise PET frames. These low-noise images were then co-registered to the reference frame, yielding 3D motion vectors. Performance of cGAN-aided motion correction was assessed by comparing the image-derived input function (IDIF) extracted from a cGAN-aided motion corrected dynamic sequence against the AIF based on the areas-under-the-curves (AUCs). Moreover, clinical relevance was assessed through direct comparison of the average cerebral metabolic rates of glucose (CMRGlc) values in grey matter (GM) calculated using the AIF and the IDIF. **Results:** The absolute percentage-difference between AUCs derived using the motion-corrected IDIF and the AIF was  $(1.2 \pm 0.9) \%$ . The GM CMRGlc values determined using these two input functions differed by less than 5%  $((2.4 \pm 1.7) \%)$ . **Conclusion:** A fully-automated data-driven motion compensation approach was established and

59 tested for 18F-FDG PET brain imaging. cGAN-aided motion correction enables the translation of  
60 non-invasive clinical absolute quantification from PET/MR to PET/CT by allowing the accurate  
61 determination of motion vectors from the PET data itself.

62 **Keywords:** 18F-FDG brain, Deep learning, Head-motion correction, absolute quantification,  
63 Patlak analysis.

64

## INTRODUCTION

The capability of obtaining fully-quantitative physiological measurements from the human body is a fundamental strength of PET methodology (1,2). However, due to the requirement of an arterial input function (AIF), the adoption of absolute quantification into clinical work has been severely limited, and only semi-quantitative assessments of tracer uptake using standardized uptake expressed as SUV are commonly performed.

In recent years, several methodologies have been proposed to extract an image-derived input function (IDIF) directly from brain PET data in order to avoid arterial cannulation for determination of an AIF (3–10). It was demonstrated that a brain IDIF can be calculated either by using a combined protocol that integrates PET/CT with MR data (3,7,8), or by using a fully-integrated PET/MR protocol (4,6,9,10). The calculation of an IDIF typically entails, in addition to the definition of a suitable blood pool region and accounting for partial volume effects, also an accurate correction for involuntary subject motion. A fully-integrated PET/MR system is ideally suited to perform all these tasks due to its capability of providing detailed anatomical information which also includes MR navigators that track motion. However, due to its high cost, the proliferation of PET/MRI into the clinical realm has been severely limited. In contrast, PET/CT is widespread and cost-effective, thus, motivating the transfer of IDIF methodology from PET/MR to PET/CT. Using co-registered PET/CT and MR data, the definition of a suitable blood pool region, as well as the geometric correction for partial volume effects can be easily accomplished. Nonetheless, the accurate correction for subject motion remains a serious challenge in PET/CT imaging.

Assessment of currently available motion compensation techniques points towards three general approaches: data-driven approaches (11–17), frame-based image-registration (FIR) (18,19) and real-time hardware motion tracking (HMT) (20). Real-time hardware-based motion tracking detects subject motion with excellent temporal resolution (20), but is typically not used in clinical routine due to its complexity and the necessity to integrate external data (motion tracking) with the imaging system (applying the motion vector to images). In contrast, data-driven approaches do not require any external information (such as fiducials or laser positioning system) and they are also less computationally demanding. However, the clinical adoption of frame-based motion correction schemes has been slow due to poor performance when co-registration is applied to frames that display a dissimilar tracer uptake pattern or noise characteristics as well as the difficulty to correct for intra-frame motion in long-duration PET frames (>5 min).

Here, we explored the utility of conditional Generative Adversarial Networks (cGAN) (21) as a data-driven approach to facilitate motion correction for involuntary subject motion in dynamic 18F-FDG PET studies of the brain. Thereby, we build on recent studies that have shown the potential of cGAN methodology in converting low-count PET images to high-count images (22). In general, the objective of cGAN processing is the mapping of a low-count tracer distribution pattern to a high-count pattern based on *a priori* training data, where generic image features - such as overall brain shape and contours – are likely to be correctly reproduced. Put differently, the creation of high-count images enhances image features that are important for the detection of motion artefacts and as a result might improve the performance of subsequently applied conventional rigid body co-registration routines. It is important to note, however, that

the so obtained images are devoid of unique characteristics that are specific for a particular subject and cannot be considered as representing the true (subject-specific) high-count uptake pattern beyond the enhancement of generic image features.

In light of the above, our ultimate objective was to determine the accuracy with which involuntary subject motion occurring during the first part of a dynamic  $^{18}\text{F}$ -FDG PET/CT study can be detected using conventional motion correction routines when images are first pre-processed using cGAN methodology, given that early-frames are subject to both low-count statistics and dynamically changing tracer uptake patterns. The cGAN preprocessed frames can be thought of as PET navigators whose activity distribution are now temporally invariant, similar to that of the MR navigators. Although our focus was geared towards the derivation of an IDIF, the developed methodology appears to be broader in scope, potentially aiding in improved ability to detect both inter- and intra-frame motion. Consequently, our study was guided by the overarching hypothesis that cGAN pre-processing of images can be used to address low count limitations of short time frame motion correction strategies and support an accurate data driven arterial IDIF calculation.

## MATERIALS AND METHODS

Ten healthy volunteers (5M/5F,  $(27 \pm 7)$  years) were included in this study (10,11). The study was approved by the Ethics Committee of the Medical University of Vienna and was performed in accordance with the revised Declaration of Helsinki (1964). All volunteers were deemed to be healthy based on their medical history, physical examinations and vital signs. Written informed consent was obtained from all subjects prior to the examinations.

### Study Design

We studied 10 subjects, each of whom received two PET/MR scans (mean time difference =  $(17 \pm 44)$  days) in a fully-integrated PET/MRI system (Biograph mMR, Siemens Healthineers, Germany). To correct the PET study for involuntary subject motion, cGAN image pre-processing was performed prior to image co-registration, enabling the accurate determination of motion parameters in 3D space. These motion parameters were then used to extract the IDIF from the motion-corrected dynamic PET sequence (Fig. 1). To assess the accuracy of the IDIF, arterial blood samples were obtained from a radial artery. Finally, immediately following the PET/MRI examination, a low-dose CT scan of the brain (120 kVp, 50 mAs) was acquired once using a PET/CT system (Biograph TruePoint TrueView 64, Siemens Healthineers, USA) for the purpose of CT-based attenuation correction.

### Imaging Protocol

All examinations were conducted in the afternoon; subjects were asked to keep their eyes open without performing any task. Prior to each scan, the glucose concentration (mmol/l) in



144 blood was measured and a venous line was established for the injection of the  $^{18}\text{F}$ -FDG tracer.  
145 In addition, an arterial line was established in the contralateral arm for manual arterial blood  
146 sampling. To ensure a high signal-to-noise ratio in the MR images, a head and neck coil was used.  
147 After positioning the brain in the center of the field-of-view (FOV), a 60-min PET list-mode (LM)  
148 acquisition was initiated with the bolus injection of  $^{18}\text{F}$ -FDG ( $(352 \pm 66)$  MBq).  
149 Contemporaneously with the PET data acquisition, multiple MR sequences were acquired: a 3D  
150 time-of-flight MR angiography (TOF-MRA) sequence ( $0.5 \times 0.5 \times 1 \text{ mm}^3$  voxel size,  $\text{TE}=3.6\text{ms}$ ,  
151  $\text{TR}=21\text{ms}$ ,  $25^\circ$  flip angle,  $228 \times 384$  matrix, 220 slices) for the definition of the carotid vasculature  
152 and a T1-w MRI sequence ( $1 \times 1 \times 1 \text{ mm}^3$  voxel size,  $256 \times 256$  matrix, 192 slices) for anatomical  
153 localization. Sparsely sampled MR navigators (2D-EPI,  $3.0 \times 3.0 \times 3.0 \text{ mm}^3$  voxels,  $64 \times 64$  matrix,  
154 36 slices,  $\text{TE}=30 \text{ ms}$ ,  $\text{TR}=3000 \text{ ms}$ ) were interleaved at specific time intervals (0, 2.5, 5, 7.5, 10, 14,  
155 17, 21, 26, 33, 38, 42, 44 and 50.5 min post injection (p.i.)) yielding for each time point a 3D image  
156 volume (23) that allowed the determination of subject motion with 6 degrees of freedom  
157 (translation in x,y,z direction and rotation with respect to the three Euler angles). These six  
158 motion parameter defined a “motion vector”, which was used to ensure spatial correspondence  
159 between early frames and late frames for cGAN training.

160 PET list-mode data was re-binned into a dynamic frame sequence ( $24 \times 5 \text{ s}$ ,  $1 \times 60 \text{ s}$ ,  $1 \times$   
161  $120 \text{ s}$ ,  $11 \times 300 \text{ s}$ ) and was reconstructed (Siemens e7 tools) into a  $344 \times 344 \times 127$  matrix (voxel  
162 size  $2.08 \times 2.08 \times 2.03 \text{ mm}^3$ ) using the Ordinary Poisson Ordered Subset Expectation-  
163 Maximization (OP-OSEM) 3D algorithm (3 iterations, 28 subsets, 2 mm Gaussian filter).

164 Attenuation and scatter correction were performed using CT-based AC-maps corrected for  
165 motion in each PET frame (see below).

## 166 **Blood Sampling**

167 Arterial blood samples were collected manually at different time points (24 x 5 s, 1 x 60 s,  
168 1 x 120 s, 1 x 300 s, 1 x 600 s, 2 x 1200 s post-injection) from the radial artery. Whole-blood  
169 radioactivity concentrations were measured using a gamma counter (PerkinElmer, 2480  
170 Automatic Gamma counter, Wizard23). To obtain the arterial input function (AIF), whole blood  
171 samples were centrifuged to separate the plasma component, followed by the measurement of  
172 radioactivity in the plasma. The measured whole blood and plasma tracer concentrations were  
173 used to calculate the time-dependent plasma-to-whole blood ratios for each subject.

## 174 **3D-Conditional Generative Adversarial Networks (3D-cGAN)**

175 Generative Adversarial Networks (GANs) are generative algorithms, which belong to the  
176 field of unsupervised learning (24). The architecture of a GAN consists of two convolutional  
177 neural networks (CNNs) that together constitute an opponent-component system: a neural  
178 network (termed the Generator  $G$ ) that generates artificial data based on a training data set, and  
179 a neural network (termed the Discriminator  $D$ ) that classifies the artificially created data as being  
180 either real (i.e. belonging to the training data set) or being artificially generated.

181 Conditional GANs (cGANs) are a supervised extension of the GAN framework (25). While GANs  
182 typically perform a mapping operation from a random noise vector ( $z$ ) to an output vector ( $y$ )  
183 expressed as ( $G: z \rightarrow y$ ), cGANs perform a mapping operation from an observed image ( $x$ ) and a

random noise vector ( $z$ ) to an output image ( $y$ ), expressed as ( $G: \{x, z\} \rightarrow y$ ). Here, the mapping operation is the process of linking two image patterns together and is “learned” from a training set that defines the true correspondence between pairs of input and output images ( $x \rightarrow y$ ). In short, the training data set provides a general mapping of two images with different noise characteristics (e.g., a low-count to a high-count image). It is a generic mapping operation that accounts for broad features in the two images but does not account for subject-specific attributes. Accordingly, such mapping is representative for the transformation of any 18F-FDG uptake image obtained at a particular time (e.g. 2-3min p.i) to any FDG uptake image at a later time (e.g. 55-60min p.i).

In this 3D-cGAN implementation, corresponding pairs of low-count (early) and high-count (late) PET frame images were used to define the mapping operation ( $G$ ) by minimizing a loss function expressed as:

$$\mathcal{L}_{cGAN}(G, D) = E_{x,y}[\log D(x, y)] + E_x \left[ \log \left( 1 - D(x, G(x)) \right) \right] \quad Eqn.1$$

where  $G$  attempts to minimize the loss function ( $\mathcal{L}_{cGAN}$ ), while  $D$  strives to maximize it (i.e.  $G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D)$ ). To create artificially generated high-count images from low-count (early) PET frames, we added an estimation error loss to the Discriminator feedback for the effective training of the Generator ( $G$ ) (23). The final loss function,  $G^*$  is then expressed as:

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G) \quad Eqn. 2$$

where  $\mathcal{L}_{L1}(G)$  is an additional L1-norm based loss-function for the generator and  $\lambda$  is a tunable parameter, which is greater than zero (in our case  $\lambda = 1$ , (24)). The U-net-like architecture (26)

with skip connections was used as the Generator network (Supplemental Fig. 1), taking 3D sub-volumes of the original early-frame PET image as input. The skip connections facilitate the preservation of the local image information that is lost during the initial down-sampling process and transfer this information to the later occurring up-sampling phase of the network. We used a *PatchGAN* (Supplemental Fig. 2) as a Discriminator (22), which classifies each given patch as either true or artificial. In addition, we added two more convolutional layers to the Discriminator architecture. The advantage of this architecture is that the network preserves the high-frequency structures of the high-count (late) PET frames using fewer parameters than would be required using the full-size images. Training of the CNNs was performed using the standard method from Goodfellow et al (24).

#### **cGAN-aided Motion Correction**

A random 70%-to-30% data split of the full data was used for cGAN training (14 scans) and testing (6 scans). Initially, all studies in the training set (14 measurements with 37 frames each) were corrected for motion using motion vectors obtained from the contemporaneously acquired MR navigators (23) (Fig. 1). Real-time data-augmentation (rotation, translation, shearing, additive gaussian noise, brightness, contrast) was performed on the training datasets resulting in 21,000 synthetic datasets. Subsequently, cGAN mapping was carried out between the last high-count PET frame (reference frame #37 representing tracer accumulation at 55-60 min p.i.) and all other PET frames, resulting in 36 mappings ( $x \rightarrow y$ ) with variable quality (Fig. 2).

The obtained mappings were then applied to the test data sets to obtain artificially-generated high-count images (using cGAN) that imitate the distribution of the reference frame from the

original low-count images. Following the generation of the cGAN based high-count images for the test datasets, motion correction was performed by considering the 55-60 min p.i. PET-frame as the reference frame ( $F_{ref}$ ) and all other frames were subsequently registered to the  $F_{ref}$  using a standard multi-scale mutual information routine (Greedy module ITK 1.2.4, Insight Segmentation and Registration Toolkit). For each frame of the dynamic sequence, this approach resulted in a 6-parameter motion vector.

### **Standard PET Frame-based Motion Correction**

To evaluate the added value of cGAN aided motion correction, this methodology was compared to a standard PET frame-based motion correction. PET image frames were aligned using the same multi-scale mutual information based coregistration routine as described above (Greedy module ITK 1.2.4). This routine performs alignment between images starting at a coarse scale which is then used to initialize registration at the next finer scale, a process repeated until it reaches the finest possible scale. As for the early images (< 3 min p.i.), the applied multi-scale mutual information coregistration approach failed due to insufficient count statistics, we summed the first 3 min of the dynamic sequence to create a reference frame with sufficient statistics. Subsequently, all later frames (>3 min p.i.) were rigidly aligned to this summed frame. It is important to point out that this approach (summing of early frames) is frequently implemented in dynamic studies when low-count images are analyzed that do not contain sufficient data that would allow extraction of an accurate motion vector. Because robustness of this coregistration procedure can be improved by low-level smoothing (The ITK Software Guide, Kitware inc.), our standard registration approach therefore consisted of applying a heuristically

chosen 4 mm gaussian filter to the images prior to registration. However, in order to assess the performance of cGAN methodology when processing the original (low-count) images, this smoothing step was omitted when testing cGAN processed images.

## **PET Emission Data and Attenuation Map Alignment**

To account for misalignment between the PET emission data and the attenuation map, we used a dual-reconstruction approach. Specifically, PET attenuation correction was performed based on an attenuation map derived from a CT image acquired immediately following the PET/MR imaging protocol. This CT attenuation map was co-registered to the first MR sequence of the study protocol (TOF-MRA sequence) and this static map was then used to perform attenuation correction for the whole PET dynamic sequence. However, this approach does not take into account PET inter-frame motion. Therefore, non-motion corrected PET frames were initially used to derive a motion vector (using either MR navigators for the training data or cGAN-processed images for the test data) and once the motion vector was determined, this information was used to align the CT attenuation map to each PET dynamic frame. The motion-corrected CT attenuation maps were subsequently used to re-reconstruct the whole dynamic PET sequence.

## **Characterization of cGAN Performance**

In order to assess the degree to which cGAN image processing is able to enhance generic features of brain tracer distribution (such as overall brain shape and contours) cGAN performance of individual frames was assessed based on two measures: first, by the improvement in mutual information (MI) between the reference image and the cGAN generated

high-count images relative to the original images, and, second, by the comparison of the absolute percent difference with respect to the histogram area-under-the-curve ( $\text{Hist}_{\text{AUC}}$ ) between histograms derived from cGAN generated and reference images. It is important to note that these histograms include all image voxels in 3D space and are not affected by subject motion. Moreover, the MI and  $\text{Hist}_{\text{AUC}}$  are complimentary; MI is sensitive to the similarity in image patterns expressed in the two images, while  $\text{Hist}_{\text{AUC}}$  provides information with respect to scale-relationship between voxel intensities in the two images (27,28).

### **Generation of Simulated Test Data Sets**

Given the low number of the original test data sets (6 scans), additional test data sets were generated based on the original six test data sets. These simulated data sets were used to further investigate the potential of the cGAN method to address the problem of inter-frame motion.

Excessive inter-frame motion was simulated by adding to each dynamic frame (except the reference frame) an arbitrary translation (0, 1 or 2 voxels in each direction) and rotation (0, 0.5 or 1 degree for each Euler angle) vector (SimpleITK 1.2.4, Supplemental Fig. 3). Ten repetitions were performed for each dataset with different motion vectors added, resulting in a total of 60 synthetically created test data sets, each consisting of 37 dynamic frames. These data sets were then either pre-processed using cGAN methodology, or were directly co-registered to the reference frame (standard PET-frame based motion correction) using the normalized mutual information alignment routine as described above.

Moreover, in order to assess the potential of cGAN methodology to aid in the detection of intra-frame motion in routine clinical static studies, we partitioned the list mode data of the reference frames (#37 from 55 – 60min p.i) into a series of short sub-frames (10 sub-frames of 30s and 20 sub-frames of 15s). cGAN was used to produce artificially generated high-count images from these sub-frames. This procedure was applied for each measurement of the test data set (6 scans). Improvement in image quality was assessed based on the increase in MI following cGAN processing. In addition, to further demonstrate the ability of cGAN processing in accurately accounting for intra-frame motion, we selected a representative reference frame, which we partitioned into 15s sub-frames and introduced random motion (translation 3-5 mm, rotation < 1 degrees) to the sub-frames. Following motion correction either with (cGAN-aided) or without cGAN pre-processing, the coregistered sub-frames were summed and the resulting images were visually compared for image quality.

#### **cGAN-based IDIF**

To assess the clinical performance of the cGAN method for motion compensation of dynamic PET frames, we extracted the IDIF from the test data set (both original and simulated) and compared the IDIF with the AIF. For this, we replaced the MRI navigator-based motion correction in our previously developed IDIF pipeline with the developed cGAN-aided motion correction. This analysis pipeline was described in detail in (9,10). In brief, it entails automated segmentation of the petrous region of the internal carotid arteries (ICA) from the corresponding TOF-MRA images followed by an automatic multi-scale intermodal NMI co-registration (Greedy ITK 1.2.4) of the TOF-MRA volume and the reference frame (frame #37) for each study. The arterial blood pool



region was defined based on the MR angiography image; clearly identifying the ICA region. This region was subsequently transferred into PET space where it was used to extract the time-activity curve. cGAN-aided motion vectors were applied during the extraction of the time-activity curve in order to adjust the blood pool region for the computed displacements. Finally, an iterative regional partial volume correction procedure was applied in each PET frame to recover the true activity in the internal carotid arteries (10).

### **Post-processing of IDIF**

The motion- and partial volume-corrected IDIF was interpolated with a step length of 1 using a “Piecewise Cubic Hermite Interpolating Polynomial” to match the blood sampling times. All corrections were applied to the IDIF with the AIF being considered as the reference (3,9). First, count rates from sampled arterial blood were scaled using the cross-calibration factor between the PET/MR and the on-site gamma counter in order to express the AIF in the same units as the PET data (Bq/mL). Second, a plasma IDIF was derived based on the individual plasma-to-blood ratios obtained from sampled arterial blood of the study subjects. Third, the delay between the AIF and the IDIF was corrected by shifting the IDIF curve to match the sampling times of the AIF. Finally, due to the difference in sampling location (ICA for IDIF and radial arteries for AIF), a mono-exponential dispersion function with a tau value of 5 s (29) was convolved with the IDIF to mimic the dispersion effects.

## Quantification of CMRGlc

Calculation of cerebral metabolic rate of glucose (CMRGlc) for the test data sets was performed using a voxel-wise Patlak graphical analysis (lumped constant,  $LC = 0.65$  (30)) using either the AIF or the IDIF. Analyses were performed using in-house developed Matlab tools (Matlab R2018a, MathWorks, USA) that generate parametric images representing CMRGlc ( $\mu\text{mol}/100\text{g}/\text{min}$ ). In particular, a linear function was fitted to the Patlak-transformed data, including data from 25 min p.i. until the end of the scan (8 data points). The resulting slope was then multiplied with the subject's plasma glucose level ( $\mu\text{mol}/\text{L}$ ) and divided by the LC. By applying a grey matter (GM) mask derived from individually co-registered T1-w MR images, the average CMRGlc value in the GM was determined using either the AIF ( $\text{CMRGlc}_{\text{AIF}}$ ) or the IDIF ( $\text{CMRGlc}_{\text{IDIF}}$ ).

## Assessment of cGAN Performance for Motion Compensation

The quality of cGAN-aided motion correction of the dynamic frame sequence was assessed in relation to the sampled AIF. Specifically, IDIFs were determined from the test data sets (both original and simulated) using cGAN-aided motion vectors and compared to the AIFs by utilizing the area under the curve (AUC). Differences in GM CMRGlc values derived from the IDIF and AIF were assessed using the absolute percent difference ( $|\% \Delta|$ ) between CMRGlc values:

$$|\% \Delta| = \left| \frac{\text{CMRGlc}_{\text{IDIF}} - \text{CMRGlc}_{\text{AIF}}}{\text{CMRGlc}_{\text{AIF}}} \right| \times 100 \quad \text{Eqn. 3}$$

cGAN performance was assessed separately for the original data set ( $N = 6$ ) and the simulated data set ( $N = 60$ ).

## RESULTS

Fig. 2 visualizes the results of cGAN processing of the dynamic frame sequence from ~100s (1.7min) p.i. until the penultimate frame in the study (#36). The data indicates a visually excellent quality of the artificially generated images for frames of 60s duration, as early as 2min p.i. of the tracer. In contrast, the quality of cGAN images appears suboptimal for very short frames (5s) prior to 2min p.i.

Quantitative assessment of cGAN performance based on mutual information (MI) is depicted in Fig. 3. A substantial increase in MI of the individual frames after the application of cGANs, in frames as early as 1min p.i. is clearly noted. Fig. 4 indicates a substantial decrease in the  $|\%|$  between histogram AUCs characterizing cGAN-processed images and those characterizing the original images during the very early phase of the study (60s – 120s). During that time, cGAN processed images derived from the very short frames (5s) decreased the difference in histogram AUC by ~80% relative to histogram AUCs obtained from the original images. Subsequent improvements were minor (<5%) for longer and later frames with better count statistics.

Supplemental Fig. 4 demonstrates that cGAN-based IDIFs were closer to the reference standard in comparison to the IDIFs obtained without cGAN processing and using only standard frame-based motion correction (motion profile of the represented subjects in Supplemental Fig. 5). For the original data set ( $N = 6$ ), the  $|\%\Delta|$  between AUCs derived using the motion-corrected IDIF and the AIF was  $(1.2 \pm 0.9)\%$ . The GM CMRGlc values determined using these two input functions differed by less than 5%  $((2.4 \pm 1.7)\%$  (Fig. 5). The quantitative difference in AUC and

GM CMRGlc values between AIF and IDIF (before and after cGAN preprocessing) for individual datasets (N=6) with their respective augmentations (10 iterations) is depicted in Table. 1. For the simulated data sets (N = 6 x 10 iterations), the mean difference in AUC values between those obtained using the AIF and the IDIF using cGAN preprocessing was  $0.9 \pm 0.7\%$ , whereas the difference in AUC values between AIF and the IDIF without cGAN preprocessing was  $2.9 \pm 3.1\%$ . Moreover, IDIFs extracted from cGAN-preprocessed motion compensated data resulted in CMRGlc values closer to those obtained using the AIF, with an absolute difference of  $2.2 \pm 1.8\%$  as compared to CMRGlc values determined without cGAN preprocessing of  $3.9 \pm 3.5\%$ . The improved performance of cGAN aided as compared to non-cGAN aided motion correction can be also inferred from the smaller variance of both AUC and GM CMRGlc values in case of cGAN preprocessing.

Fig. 6 shows representative images that were obtained by partitioning the reference image (55-60 min p.i.) into 15s subframes and the image quality of these sub-frames following cGAN processing. For the (10 x 30s) subframe data set, the MI improved by 135% (from  $0.030 \pm 0.003$  to  $0.070 \pm 0.001$ ), whereas improvement was even greater for the (20 x 15s) data set (improvement of 290%; from  $0.002 \pm 0.003$  to  $0.0700 \pm 0.0001$ ). Moreover, Fig. 7 demonstrates the improvement in image quality of the reference image when 15s sub-frames with artificially introduced random motion underwent standard frame-based motion correction and cGAN preprocessing prior to rigid body motion correction as compared to non-motion corrected summed sub-frames. As expected, motion correction improves image sharpness and one can appreciate a slight improvement of images processed using cGAN methodology as compared to those processed with standard frame-based motion correction.

## DISCUSSION

We present a fully-automated motion correction approach for dynamic 18F-FDG PET studies of the brain that utilizes cGAN pre-processing of low-count images in order to improve the estimation of motion vectors derived using conventional rigid body coregistration algorithms (Fig. 2). Our results suggest that cGAN methodology allows the creation of artificially generated high-count 18F-FDG brain images from early low-count images that closely resemble the 18F-FDG uptake pattern at late (~60min p.i.) scan times. The creation of artificially-generated, high-count images allows then the reliable determination of a motion vector directly from the data, which was verified by the comparison of an IDIF with arterial blood samples. Thus, cGAN-aided motion correction is likely to have a substantial impact on the quality of dynamic low-count 18F-FDG PET brain studies and as a result might contribute to the expansion of absolute brain quantification into clinical routine. Especially in the context of clinical PET/CT imaging, cGAN pre-processing of low-count image frames could play an important role in improving the performance of established motion correction approaches. The developed cGAN methodology also shows promise in addressing the problem of intra-frame motion in long duration (5 - 20 min) PET scans, by allowing the partitioning of a frame into sub-frames which, following cGAN pre-processing, can be accurately corrected for motion.

To prevent motion when imaging the brain, subjects are usually instructed to remain motionless and their head is immobilized using bands that affix the skull to the head rest (31). This approach works reasonably well when imaging cooperative subjects, but frequently fails in the clinical population due to the subjects being either uncomfortable or claustrophobic within the PET gantry (31). As a result, motion artefacts are frequently encountered in both static and

dynamic PET imaging. The problem with motion artefacts is even more severe when an IDIF is extracted from the dynamic PET sequence, given that the magnitude of random displacements and the system resolution are typically larger than the small size of the arteries. Accordingly, accurate motion correction is an important prerequisite for absolute quantification in PET imaging (17).

It has been well recognized that the original low-count/high-noise images are poorly suited for alignment due to the poor definition of image landmarks that could guide the registration procedure. As such, low-count images require some form of pre-processing in order to achieve a satisfactory performance of the subsequently applied coregistration routines (32). This pre-processing step could include various forms of smoothing or morphological operations, but could also consist of more sophisticated forms of processing, such as the here presented cGAN methodology. From a conceptual point of view, cGAN pre-processing might be superior to the previously applied methods, as cGAN processing is based on an automated (i.e. operator-independent) mapping of low-count images to their true high-count match. Stated differently, the calculated mapping is specific to the noise characteristics of the original low-count images and the resulting artificially generated images represent the most likely prediction of the final (high-count) tracer distribution one could expect based on the training data. Overall, our data clearly highlights the strengths of cGAN processing, such as the autonomous optimized smoothing and “smart” inpainting, which substantially enhance the information content of low-count images so that co-registration algorithms are provided with sufficient information to accurately estimate the motion parameters.

In practical terms, the functionality of cGANs consists in the ability to correctly predict the local distribution of measured data based on a statistically insufficient sample. More specifically, the cGAN algorithm extracts the most likely relationships between low-count images (where the underlying distribution is ambiguous) and high-count images (where the underlying distribution is well defined) from the training data set and applies the extracted relationships to new images. As a result, cGAN methodology is able to accurately predict generic image features (such as brain contours) from low-count images. The improved definition of brain contours then allows improved performance of conventional mutual information coregistration routines that strongly depend on well-defined imaging features.

Images with dissimilar uptake pattern are typically encountered in dynamic studies when the tracer uptake pattern changes as a function of time during the frame sequence. Our results showed that structural information inherent to very early low-count images (< 2 min p.i.) is insufficient to generate an acceptable mapping with the reference frame, thus, precluding the generation of a high-count image that could guide the derivation of an accurate motion vector. Conversely, 18F-FDG brain uptake at times >2min p.i. appears to be sufficient for adequate cGAN mapping if the frames are not too short (>30s). However, we acknowledge that the relevance of cGAN processing is strongly diminished in the case of high-count/low noise images that are already characterized by well-defined features. Finally, despite the fact that histograms derived from cGAN-processed images have a similar overall shape with respect to the reference images (Supplemental Fig. 6), they tend to overrepresent high intensities and should not be used in lieu of the original low-count images for clinical diagnosis.

Data from our previous studies suggests that motion magnitude increases with time and for very early time points (<2 min) the motion magnitude tends to be negligible (i.e. translation all axes < 1 mm, rotation < 1 degrees in all axes) (9,10). Moreover, even in the case of subject movement in this very early phase of the study, the error accrued in the integral under the blood time-activity curve remains negligible. We demonstrated this accuracy in our own test data set, showing a mean value of <1.5% for the absolute difference between the IDIF and the sampled arterial blood curve, which translated to an average difference of <3% for the calculated glucose metabolic rate in gray matter (Fig. 5).

An exciting application of cGAN-aided motion correction is the possibility to address intra-frame motion in static clinical PET scans. Clinical 18F-FDG brain scans are usually performed at times >45min with a typical duration of 10 – 20min. Such relatively long frames are sometimes subject to considerable patient motion artefacts, which impair image resolution and reduced image contrast important for differential diagnosis. Our data suggests that due to the high tracer uptake in brain tissue at these late time points, these long frames can be partitioned into sub-frames as short as 15s which can be then processed with cGAN methodology to yield images of sufficient quality for accurate co-registration (Fig. 6). Thus, one can envision a reconstruction protocol in which listmode data is sequentially divided into smaller and smaller sub-frames that are individually corrected for motion by taking advantage of enhanced image features generated by cGAN pre-processing, resulting in an overall improvement in image quality. We would like to point out that the short sub-frames (15s) are only necessary in order to determine the exact time of displacement. Because realistic patient motion occurs in the form of a few distinct shifts in



head position interspersed within a longer time-frame, most motion vector parameters derived from the set of all short sub-frames will be negligible.

In this context, the question arises whether mappings are specific to one particular imaging system or whether it could be also applied to 18F-FDG data obtained from other imaging systems. Our very preliminary data suggests that mappings might be transferrable to other imaging systems based on the application of our mappings to a dynamic 18F-FDG study (12 x 60 s, 4 x 120 s, 5 x 300 s) acquired using an external PET/CT system from a different vendor. The only requirement is that mappings should match their respective PET mid-times. Supplemental Fig. 7 suggests that mappings might possibly be independent from the imaging system, however this issue mandates further investigations that were not the focus of this work.

One of the main drawbacks of the study is the low number of test datasets. Although synthetic data with variable motion parameters were generated, they were still generated from the test datasets. Moreover, there are several other limitations that need to be considered when applying cGAN methodology in clinical applications. First, current implementations of cGAN processing are highly computationally intensive. The time to generate one (source-to-target) mapping pair is 17 h on a dedicated NVIDIA DGX Workstation with 1x32GB Tesla V100 Volta GPU for a frame size of 344 x 344 x 127 voxels. However, once a mapping specific for a particular tracer is established via the training process, the time to apply this mapping to a low-count image of any individual subject is only 1 min. Another potential source of error might be intra-frame motion in the reference frame. In this study, we did not correct for such motion artefacts, since neither visual inspection of image quality nor close monitoring of the subjects during the last

495 frame did indicate noticeable subject motion. Finally, our cGAN implementation requires a  
496 training set that consists of image pairs that are devoid of motion artefacts. Since our data was  
497 acquired on a fully-integrated PET/MR system, simultaneously acquired MR navigators were used  
498 to correct the training set for motion. When translating these findings to other sites, the use of  
499 cycle-GANs (33), which produce generic mappings from spatially non-corresponding data may be  
500 a potential solution that voids the requirement of motion-corrected image pairs.

## CONCLUSION

We present a data-driven motion correction approach for dynamic  $^{18}\text{F}$ -FDG brain studies that is based on cGAN methodology. The method allows the derivation of an accurate motion vector for low-count frames as early as 2min p.i., thus, facilitating the derivation of an IDIF void of motion artefacts. The developed methodology has the potential to improve the accuracy of non-invasive absolute quantification in the context of clinical PET and PET/CT studies. In addition, cGAN methodology might also facilitate correction for intra-frame motion, thus improving image quality of clinical scans with long duration.

**Dissemination.** To invite swift adoption of this novel approach by the research and clinical community, software codes developed for cGAN processing are free and available online: <https://github.com/LalithShiyam/QIMP-tools/tree/master/PANDA>. This repository contains all software code that has been used to carry out analyses in our work.

## Disclosure

Without relevance to this work, R. Lanzenberger received travel grants and/or conference speaker honoraria within the last three years from Bruker BioSpin MR, Heel, and support from Siemens Healthcare regarding clinical research using PET/MR. He is shareholder of BM Health GmbH since 2019. The other authors report no potential conflicts of interest relevant to this article.

521    **Acknowledgements**

522           We express our gratitude to Vivian Tseng from The University of Applied Arts Vienna for  
523    creating the graphical abstract for our manuscript.

524

525    **Key Points**

526    QUESTION: To explore conditional Generative Adversarial Networks (cGANs) as an image-based  
527    approach to perform motion correction in 18F-FDG dynamic PET brain studies.

528    PERTINENT FINDINGS: The proposed motion correction approach allows accurate non-invasive  
529    determination of an IDIF. The developed non-invasive method yields CMRGlc values that are  
530    within 5% of those determined using arterial sampling.

531    IMPLICATIONS FOR PATIENT CARE: cGAN-aided motion correction of 18F-FDG dynamic PET brain  
532    studies depends only on the acquired emission data and can be applied retrospectively without  
533    additional information. As such, it can be easily implemented in clinical routine and has the  
534    potential to facilitate non-invasive absolute quantification in the context of clinical 18F-FDG  
535    PET/CT patient studies.

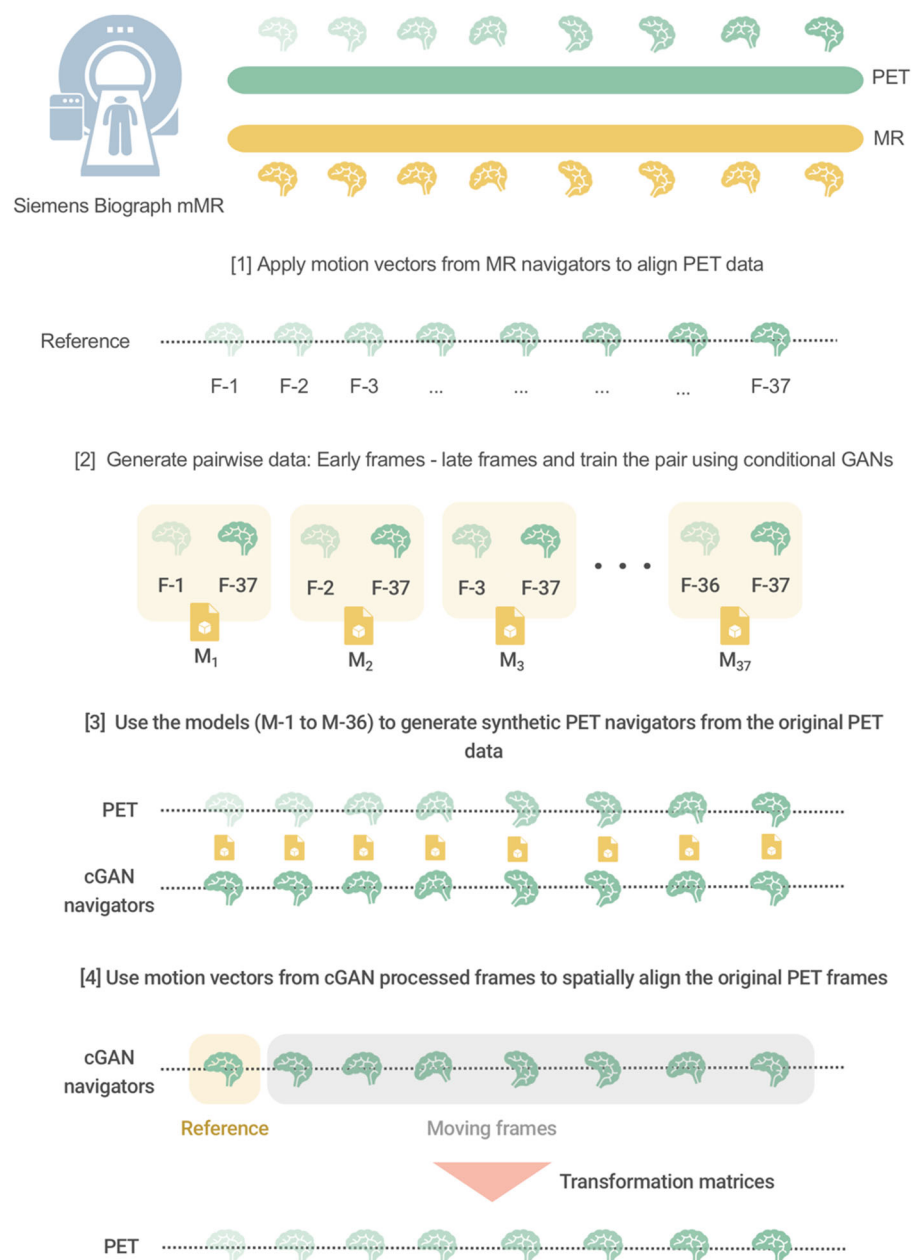
536

## 537 REFERENCES

- 538 1. Lammertsma AA. Forward to the Past: The Case for Quantitative PET Imaging. *J*  
539 *Nucl Med.* 2017;58:1019-1024.
- 540 2. Subramaniam RM. Precision Medicine and PET/Computed Tomography:  
541 Emerging Themes for Future Clinical Practice. *PET Clinics.* 2017;12:xi-xii.
- 542 3. Sari H, Erlandsson K, Law I, et al. Estimation of an image derived input function  
543 with MR-defined carotid arteries in FDG-PET human studies using a novel partial  
544 volume correction method. *J Cereb Blood Flow Metab.* 2017;37:1398-1409.
- 545 4. Jochimsen TH, Zeisig V, Schulz J, et al. Fully automated calculation of image-  
546 derived input function in simultaneous PET/MRI in a sheep model. *EJNMMI Phys.*  
547 2016;3:2.
- 548 5. Su Y, Vlassenko AG, Couture LE, et al. Quantitative hemodynamic PET imaging  
549 using image-derived arterial input function and a PET/MR hybrid scanner. *J Cereb*  
550 *Blood Flow Metab.* 2017;37:1435-1446.
- 551 6. Khalighi MM, Deller TW, Fan AP, et al. Image-derived input function estimation on  
552 a TOF-enabled PET/MR for cerebral blood flow mapping. *J Cereb Blood Flow Metab.*  
553 2018;38:126-135.
- 554 7. Su Y, Shoghi KI. Single-input-dual-output modeling of image-based input function  
555 estimation. *Mol Imaging Biol.* 2010;12:286-294.
- 556 8. Fung EK, Carson RE. Cerebral blood flow with [15O] water PET studies using an  
557 image-derived input function and MR-defined carotid centerlines. *Phys Med Biol.*  
558 2013;58:1903-1923.
- 559 9. Sundar LKS, Muzik O, Rischka L, et al. Towards quantitative [18F] FDG-PET/MRI  
560 of the brain: Automated MR-driven calculation of an image-derived input function for  
561 the non-invasive determination of cerebral glucose metabolic rates. *J Cereb Blood*  
562 *Flow Metab.* 2019;39:1516-1530.
- 563 10. Shiyam Sundar LK, Muzik O, Rischka L, et al. Promise of Fully Integrated  
564 PET/MRI: Noninvasive Clinical Quantification of Cerebral Glucose Metabolism. *J*  
565 *Nucl Med.* 2020;61:276-284.
- 566 11. Carson RE, Barker WC, Jehi-San Liow, Johnson CA. Design of a motion-  
567 compensation OSEM list-mode algorithm for resolution-recovery reconstruction for  
568 the HRRT. In: 2003 IEEE Nuclear Science Symposium. Conference Record (IEEE  
569 Cat. No.03CH37515). Vol 5. ; 2003:3281-3285 Vol.5.

12. Lu Y, Gallezot J-D, Naganawa M, et al. Data-driven voluntary body motion detection and non-rigid event-by-event correction for static and dynamic PET. *Phys Med Biol.* 2019;64:065002.
13. Ren S, Jin X, Chan C, et al. Data-driven event-by-event respiratory motion correction using TOF PET list-mode centroid of distribution. *Phys Med Biol.* 2017;62:4741-4755.
14. Feng T, Yang D, Zhu W, Dong Y, Li H. Real-time data-driven rigid motion detection and correction for brain scan with listmode PET. In: 2016 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD). ; 2016:1-4.
15. Huang C, Petibon Y, Normandin M, Li Q, El Fakhri G, Ouyang J. Fast head motion detection using PET list-mode data. *J Nucl Med.* 2015;56:1827-1827.
16. Thielemans K, Schleyer P, Dunn J, Marsden PK, Manjeshwar RM. Using PCA to detect head motion from PET list mode data. In: 2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC). ; 2013:1-5.
17. Lu Y, Naganawa M, Toyonaga T, et al. Data-Driven Motion Detection and Event-by-Event Correction for Brain PET: Comparison with Vicra. *J Nucl Med.* 2020;61:1397-1403.
18. Costes N, Dagher A, Larcher K, Evans AC, Collins DL, Reilhac A. Motion correction of multi-frame PET data in neuroreceptor mapping: simulation based validation. *Neuroimage.* 2009;47:1496-1505.
19. Picard Y, Thompson CJ. Motion correction of PET images using multiple acquisition frames. *IEEE Trans Med Imaging.* 1997;16:137-144.
20. Kyme AZ, Se S, Meikle SR, Fulton RR. Markerless motion estimation for motion-compensated clinical brain imaging. *Phys Med Biol.* 2018;63:105018.
21. Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-Image Translation with Conditional Adversarial Networks. *arXiv [csCV]*. November 2016.
22. Wang Y, Yu B, Wang L, et al. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. *Neuroimage.* 2018;174:550-562.
23. Keller SH, Hansen C, Hansen C, et al. Motion correction in simultaneous PET/MR brain imaging using sparsely sampled MR navigators: a clinically feasible tool. *EJNMMI Phys.* 2015;2(1):14.
24. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets. In: Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ, eds. *Advances in Neural Information Processing Systems 27*. Curran Associates, Inc.; 2014:2672-2680.

25. Mirza M, Osindero S. Conditional Generative Adversarial Nets. *arXiv [csLG]*. November 2014.
26. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv [csCV]*. May 2015.
27. Reza FM. An Introduction to Information Theory. Courier Corporation; 1994.
28. Studholme C, Hill DLG, Hawkes DJ. Automated 3-D registration of MR and CT images of the head. *Med Image Anal.* 1996;1:163-175.
29. Iida H, Law I, Pakkenberg B, et al. Quantitation of Regional Cerebral Blood Flow Corrected for Partial Volume Effect Using O-15 Water and PET: I. Theory, Error Analysis, and Stereologic Comparison. *Journal of Cerebral Blood Flow & Metabolism.* 2000;20:1237-1251.
30. Wu H. Measurement of the Global Lumped Constant for 2-Deoxy-2-[18F]Fluoro-D-Glucose in Normal Human Brain Using [15O]Water and 2-Deoxy-2-[18F]Fluoro-D-Glucose Positron Emission Tomography Imaging A Method with Validation Based on Multiple Methodologies. *Molecular Imaging & Biology.* 2003;5:32-41.
31. Huang C, Ackerman JL, Petibon Y, Brady TJ, El Fakhri G, Ouyang J. MR-based motion correction for PET imaging using wired active MR microcoils in simultaneous PET-MR: phantom study. *Med Phys.* 2014;41:041910.
32. Mukherjee JM, Lindsay C, Mukherjee A, Olivier P, Shao L, King MA, Licho R. Improved frame-based estimation of head motion in PET brain imaging. *Med Phys.* 2016 May;43(5):2443.
33. Zhu J-Y, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision.* ; 2017:2223-2232.



634

635 Figure 1. Schematic representation of cGAN methodology for motion correction of dynamic PET

636 frames: [1] Motion vectors from the MR navigators are applied to align the PET data. [2] All

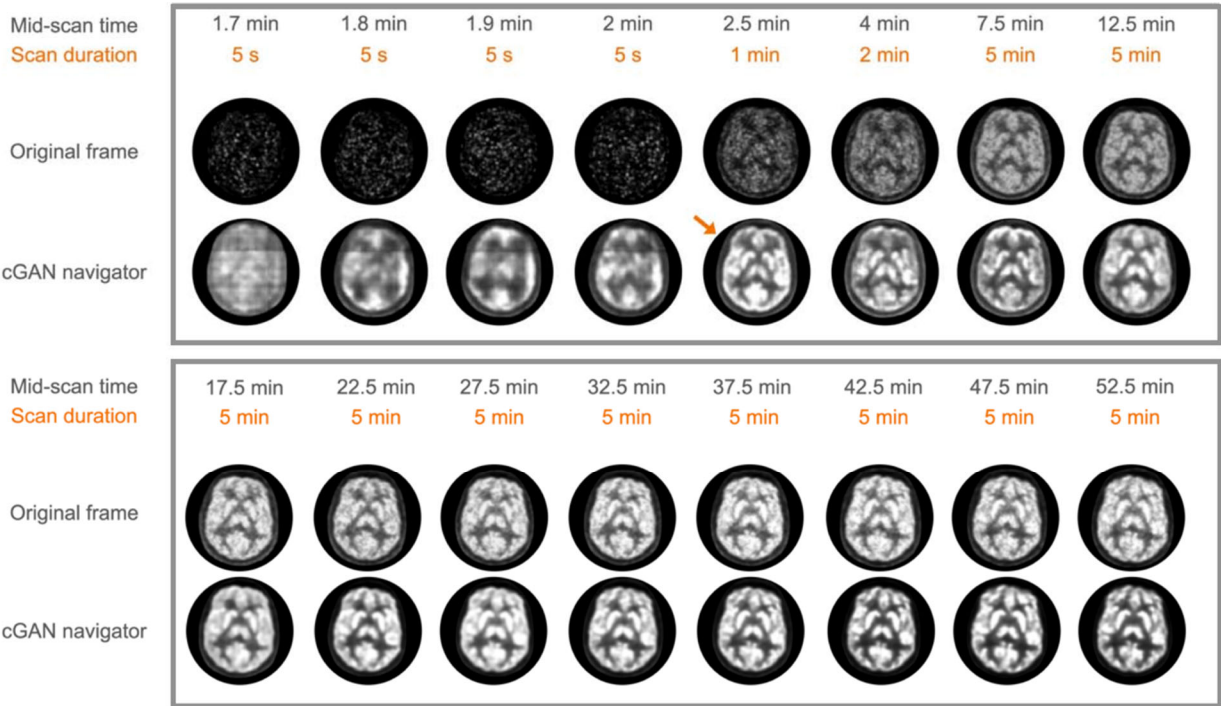
637 training data ( $N = 14$ ) is used to calculate mappings ( $M_i, i = 1, 36$ ) between each individual frame



638 (F-1 to F-36) and the reference frame (F-37). [3] These mappings are subsequently used to derive  
639 a motion vector based on co-registration of cGAN-produced artificially generated images and the  
640 reference image (test-data). [4] Application of the motion vector to either re-bin frames so that  
641 they all correspond spatially or apply the motion vector to the location of region from where a  
642 time-activity curve is extracted.

643

644



645

646 Figure 2. Results of cGAN processing for a representative subject. (Top) Original PET frames at  
647 different mid-scan times (1.7 min to 52.5 min) with various frame durations (5 s, 1 min, 2 min,  
648 and 5 min). (Bottom) Corresponding, artificially-generate, high-count PET images. cGAN  
649 processing is able to produce a tracer distribution pattern that is similar to the reference frame  
650 as early as 2min after injection (arrow, top row).

651

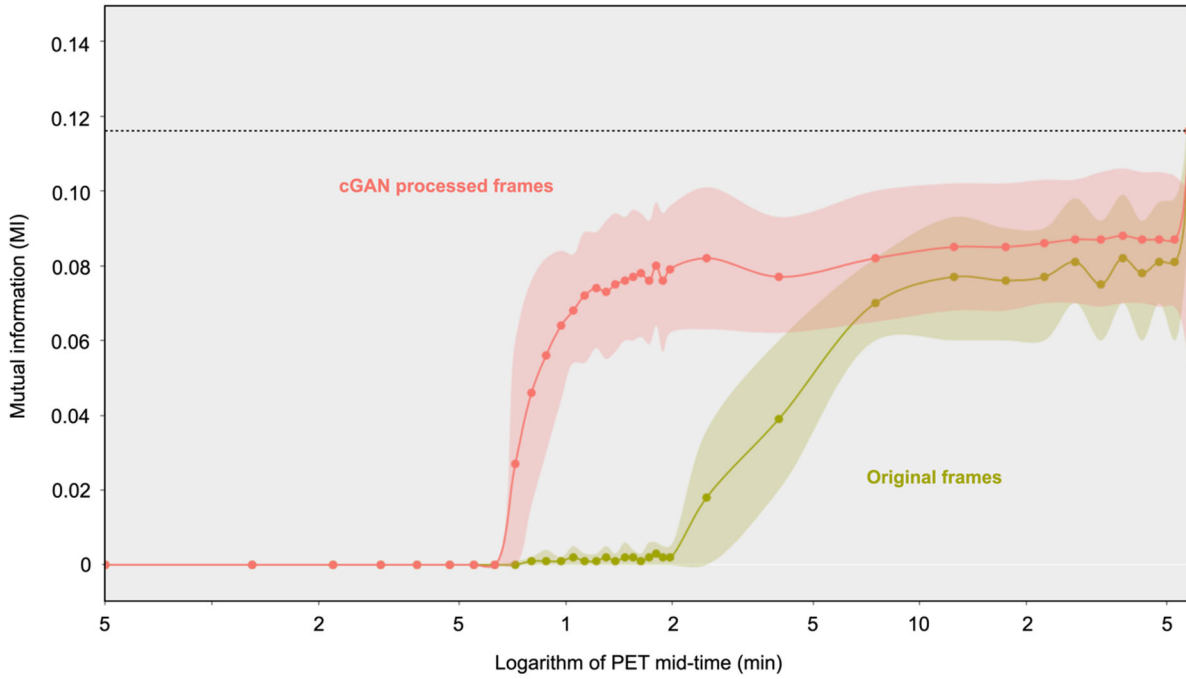


Figure 3. Comparison of mutual information index (MI) between the original (green) and cGAN processed (red) dynamic frames with the reference frame as a (log) function of scan time. Note that neither dynamic sequence is motion corrected. Values represent the average and the std. dev. (shown as shaded area) of the test data set (N = 6). cGAN processing increases mutual information of 18F-FDG brain images with respect to the reference frame close to the mutual information of late frames (dotted line).

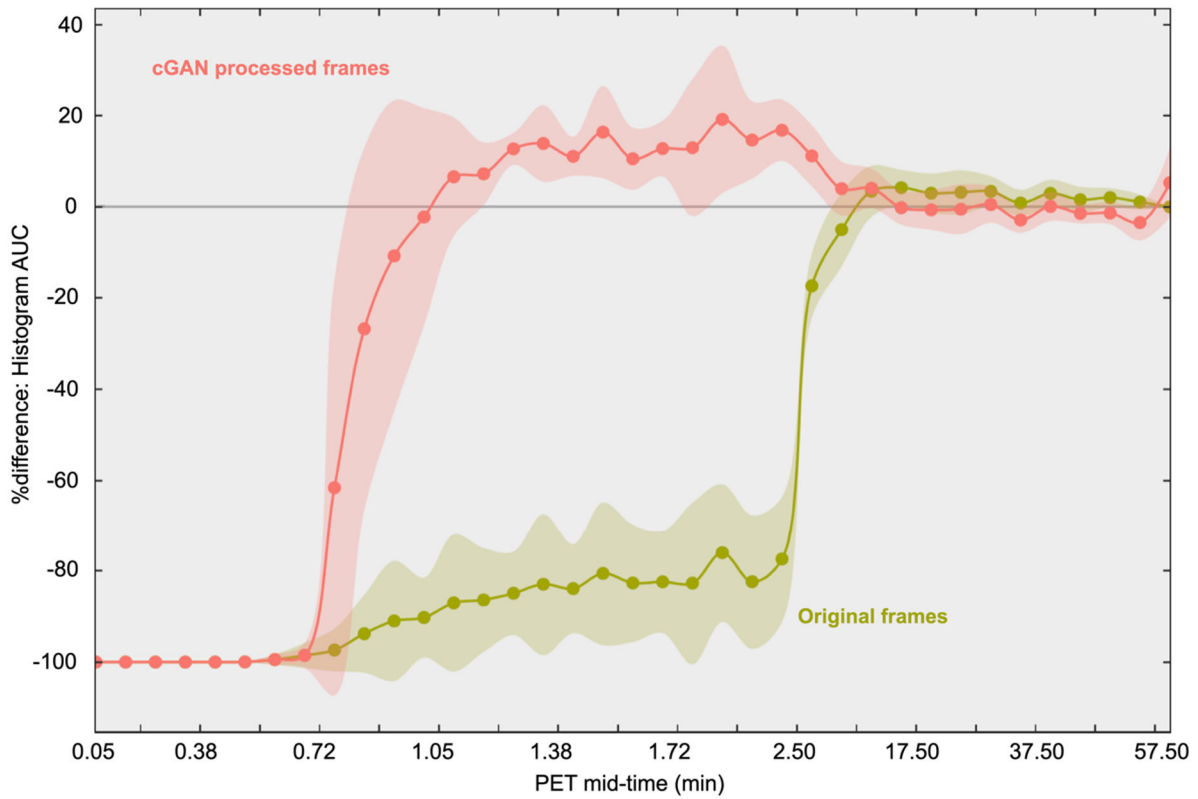
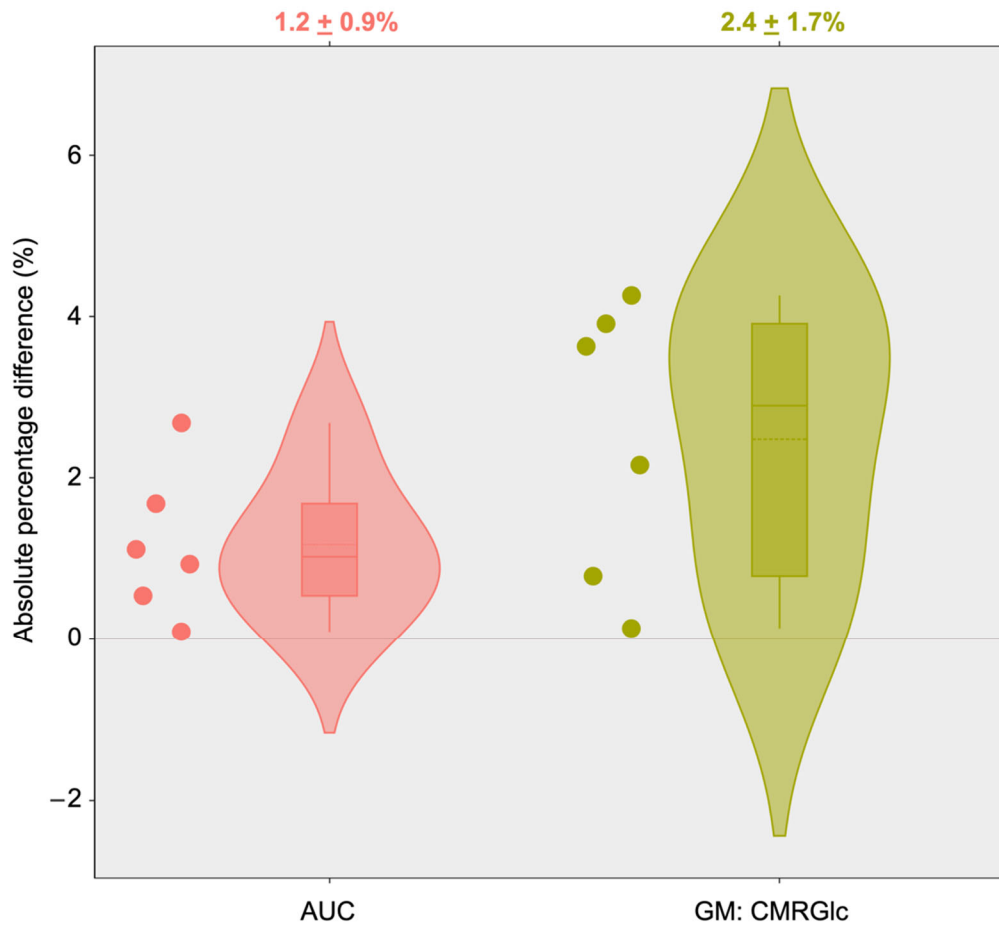


Figure 4. The  $|\% \Delta|$  of the AUC between the reference image histogram and image histograms derived from cGAN processed images (red) and original images (green) as a function of scan time. Note, images were not motion corrected since histograms are independent from motion. The % difference is substantially decreased for low-count (short) frames between 1 min and 2 min p.i., suggesting that an accurate motion vector can be determined even for early frames.



667

668 Figure 5. The  $|\% \Delta|$  between AUC and GM CMRGlc values (N = 6) obtained using the cGAN-based  
 669 IDIF and AIF. All differences were within 5% of the AIF standard. AUC: Area-under-the-curve of  
 670 the input function; GM: gray matter; CMRGlc: cerebral metabolic rate of glucose.

671

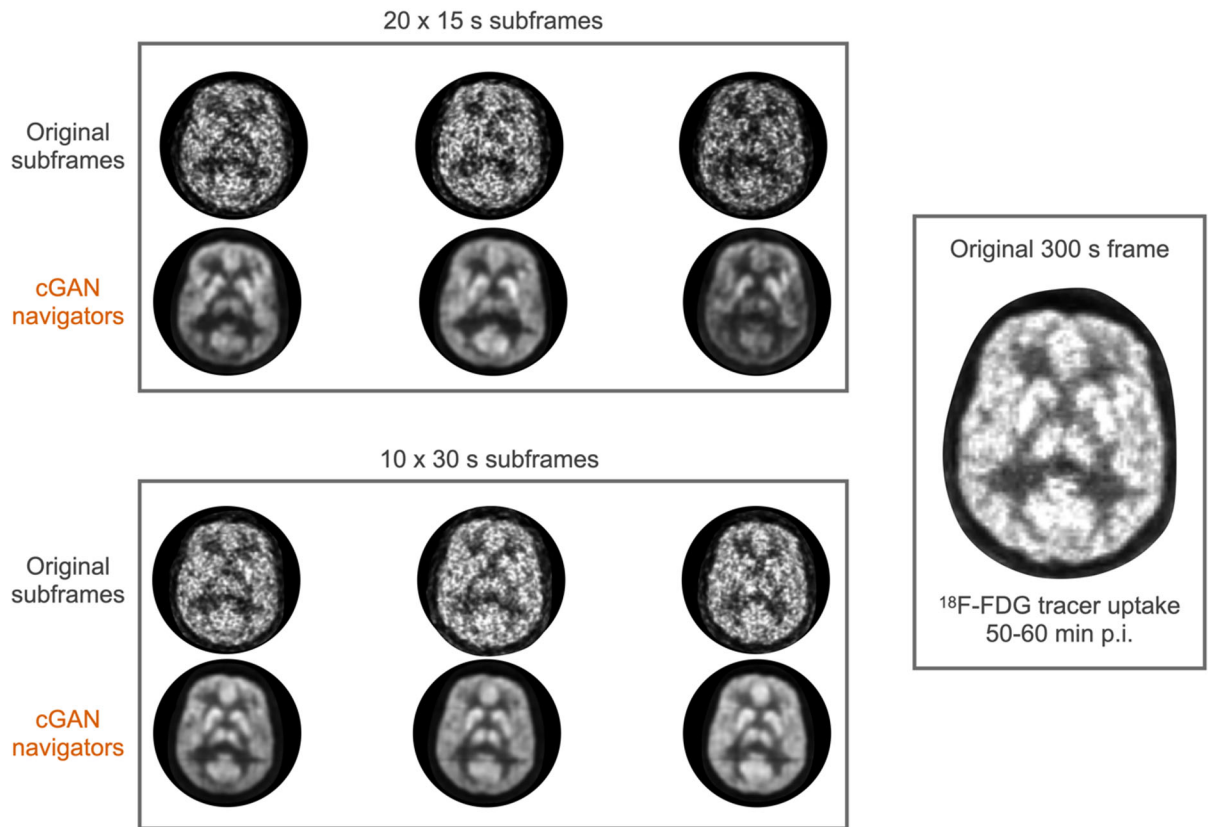


Figure 6. Performance of cGAN processing when applied to a subset of frames that were partitioned into frames of either 15s duration (upper panel) or 30s duration (lower panel) from an original 300s static  $^{18}\text{F}$ -FDG frame at 55 min p.i. (panel on right). The image quality of the processed sub-frames for both subsets is substantially improved when compared to the unprocessed sub-frames, and resembles closely the original 300s frame that includes all data.

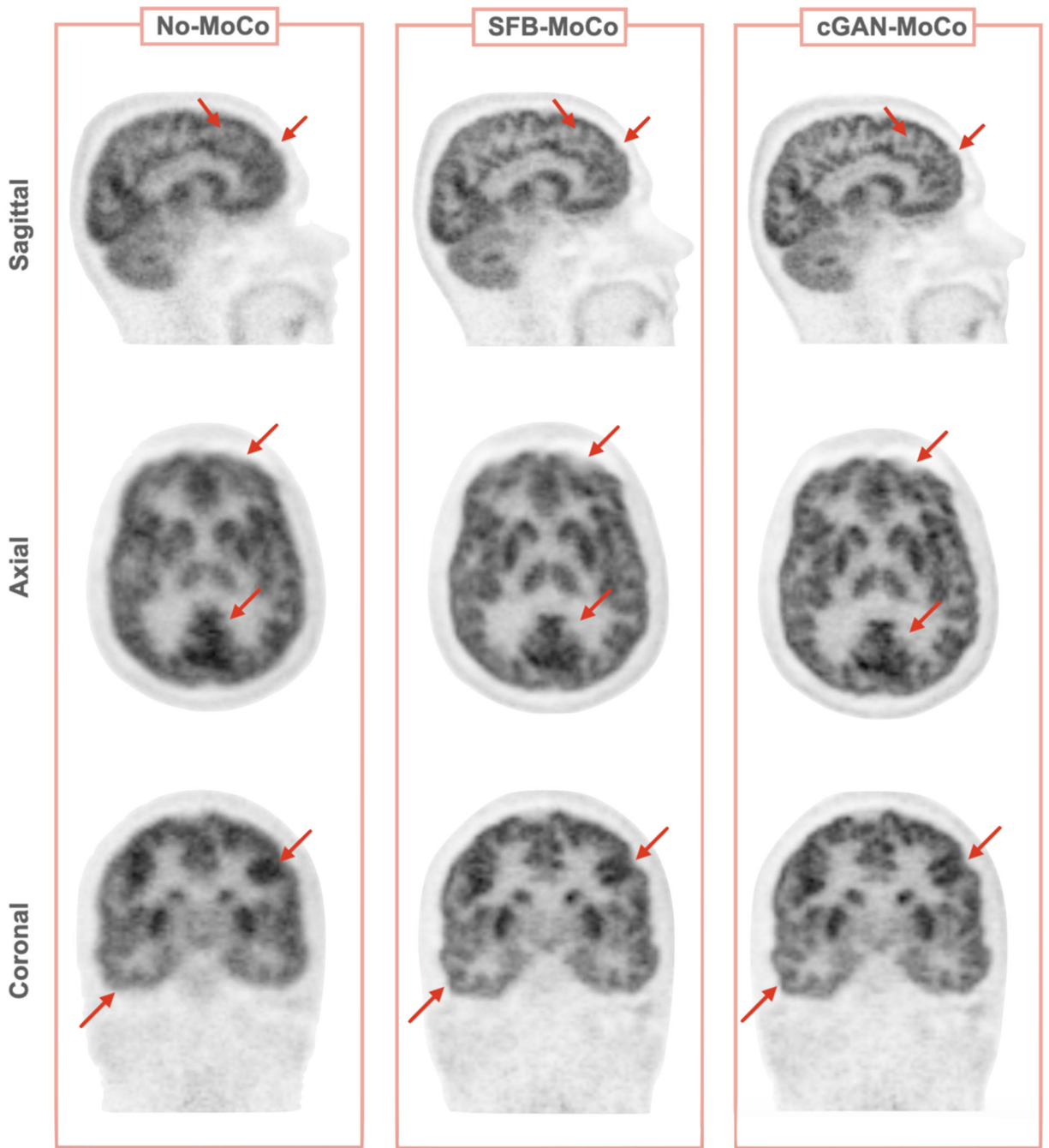


Figure 7. Summed images corresponding to 55-60 min p.i, following intra-frame motion correction using motion vectors obtained from 15s subframes. Left column: summed images without correction for intraframe motion (No-MoCo) . Middle column: summed images following standard frame-based motion correction without cGAN preprocessing (SFB-MoCo). Right

684 column: summed images using cGAN aided motion correction (cGAN-MoCo). The red arrows  
685 indicate the areas of improvement.

686



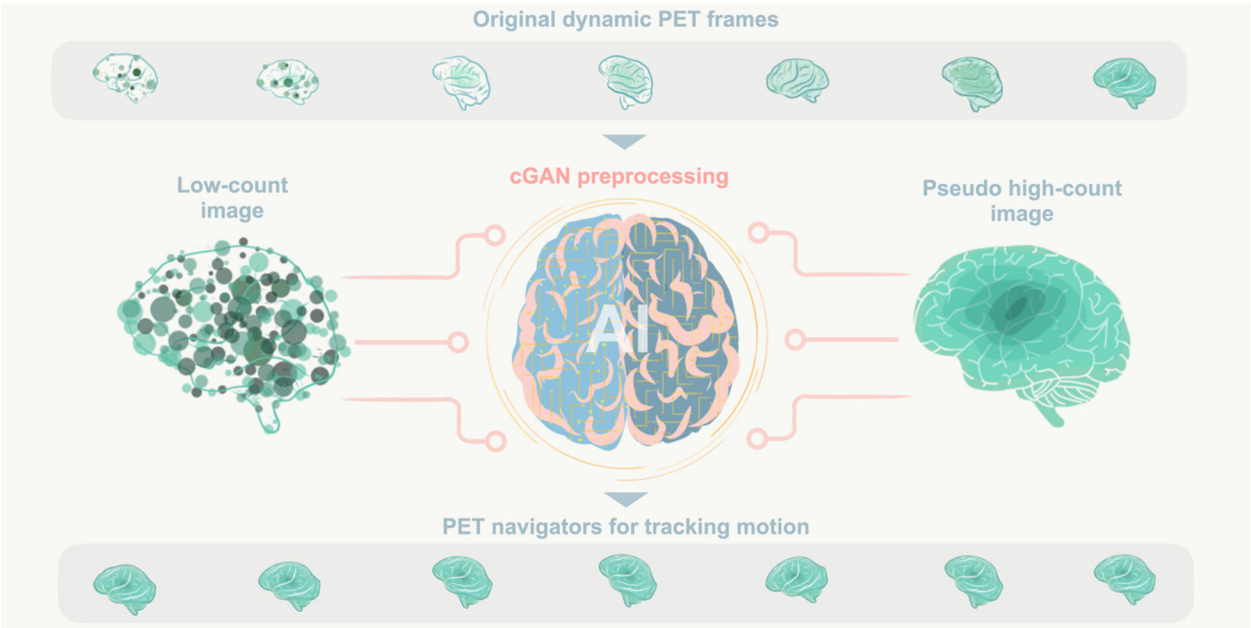
687 **Tables**

688 Table 1. The  $|\% \Delta|$  between AUC or GM CMRGlc values (N = 6 x 10 synthetic datasets) using the  
 689 IDIF (motion corrected with cGAN and without cGAN processing) and values obtained using  
 690 arterial blood sampling. AUC: Area-under-the-curve of the input function; GM: gray matter;  
 691 CMRGlc: cerebral metabolic rate of glucose.

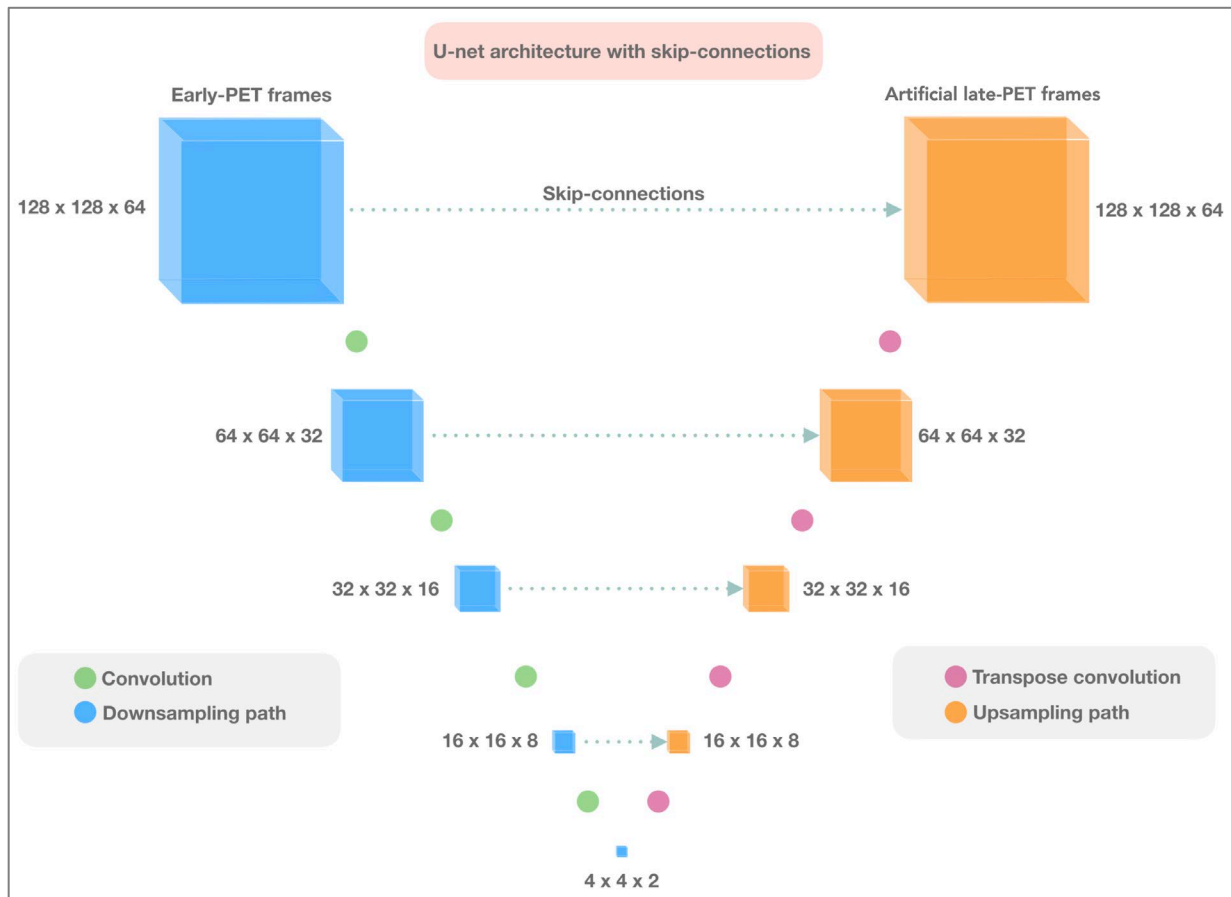
Patient ID	AUC Original PET frames (mean $\pm$ sd) %	AUC cGAN processed frames (%) (mean $\pm$ sd) %	GM CMRGlc Original PET frames (mean $\pm$ sd) %	GM CMRGlc cGAN processed frames (mean $\pm$ sd) %
P-01	(1.4 $\pm$ 0.8)	(0.8 $\pm$ 0.3)	(3.8 $\pm$ 2.3)	(3.9 $\pm$ 0.4)
P-02	(2.0 $\pm$ 1.4)	(0.4 $\pm$ 0.3)	(2.7 $\pm$ 2.5)	(1.3 $\pm$ 1.0)
P-03	(8.3 $\pm$ 1.7)	(1.7 $\pm$ 0.4)	(10.1 $\pm$ 2.6)	(0.9 $\pm$ 0.6)
P-04	(4.3 $\pm$ 2.9)	(0.8 $\pm$ 0.7)	(3.1 $\pm$ 2.3)	(0.9 $\pm$ 0.8)
P-05	(1.1 $\pm$ 0.6)	(0.2 $\pm$ 0.2)	(1.9 $\pm$ 0.6)	(1.0 $\pm$ 0.3)
P-06	(1.3 $\pm$ 0.6)	(0.5 $\pm$ 0.4)	(4.9 $\pm$ 0.9)	(1.9 $\pm$ 0.6)

692

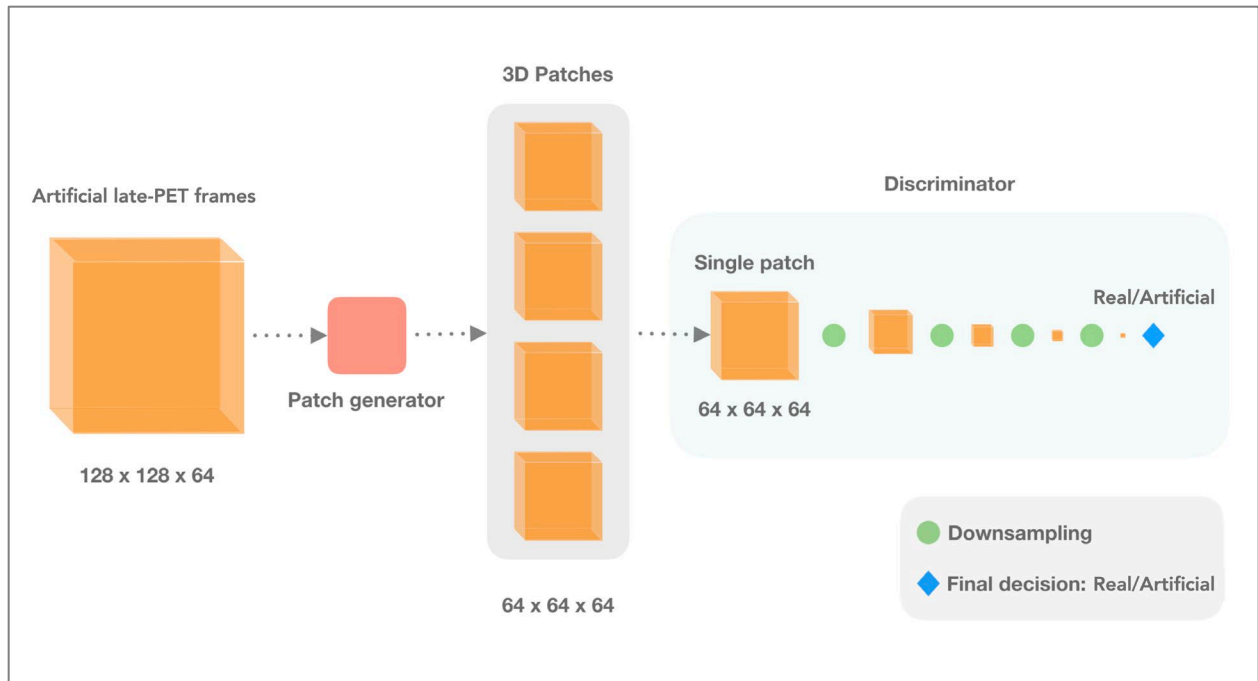
693



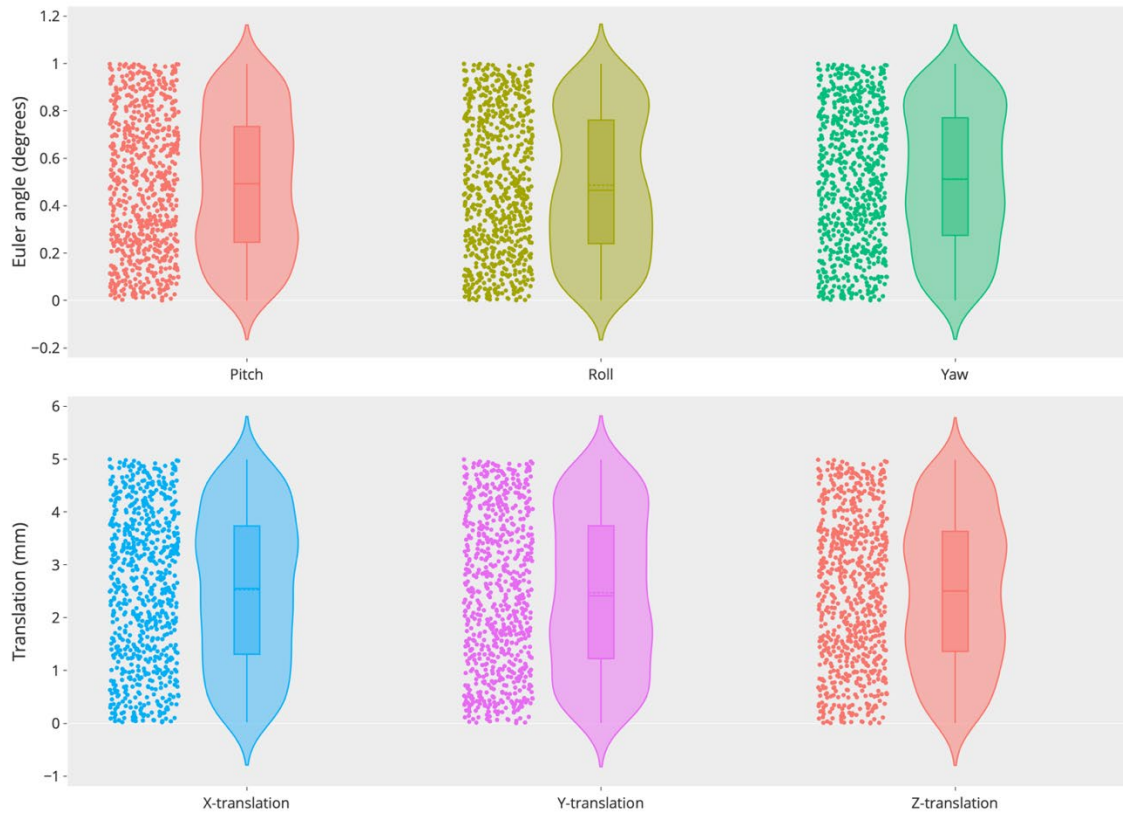
## SUPPLEMENTARY FIGURES



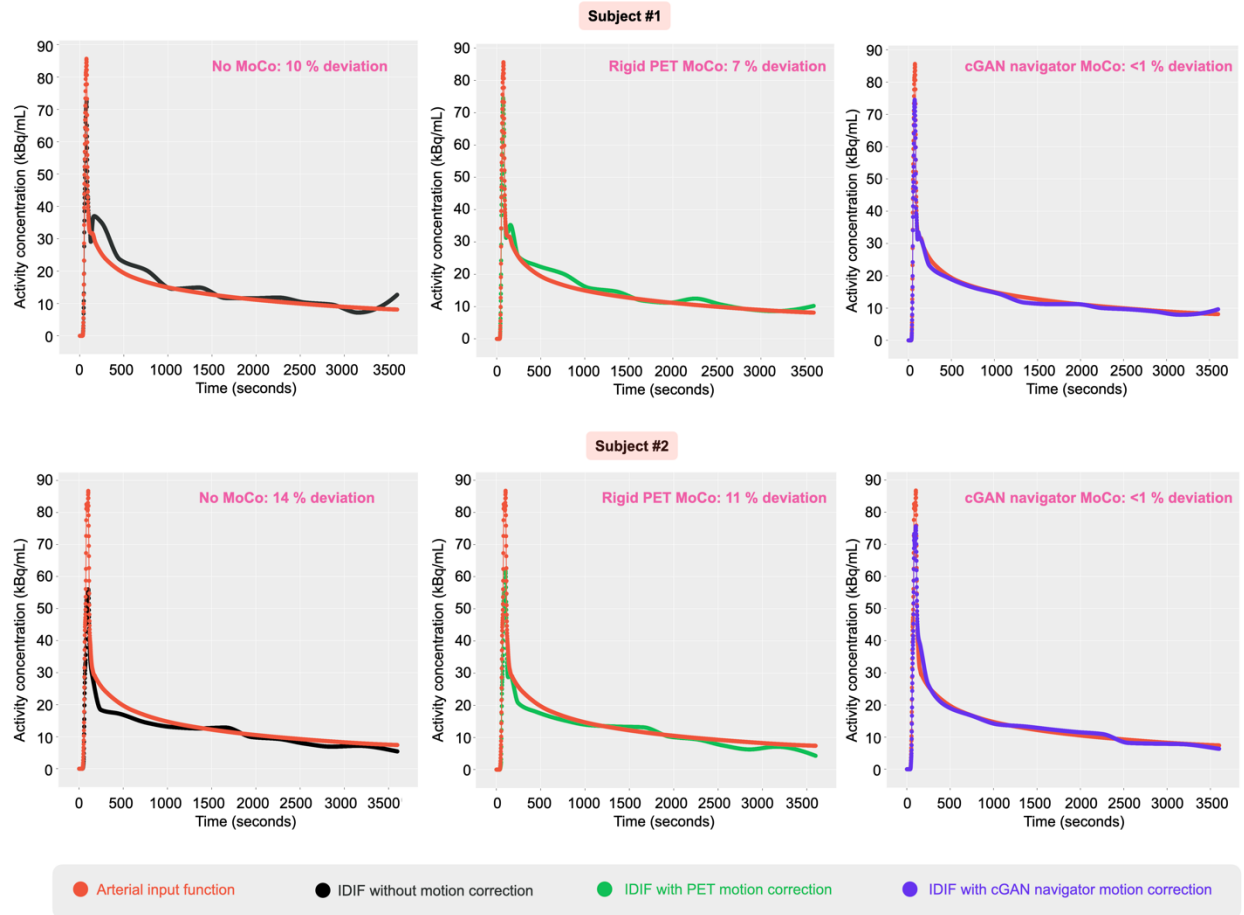
Supplementary Figure 1. Schematic representation of the 3D U-net generator architecture. The blue cubes correspond to a progressive down-sampling path, while the orange cubes correspond to the step-wise up-sampling path. The skip-connections are represented by dotted arrows and enable the preservation of high-frequency information.



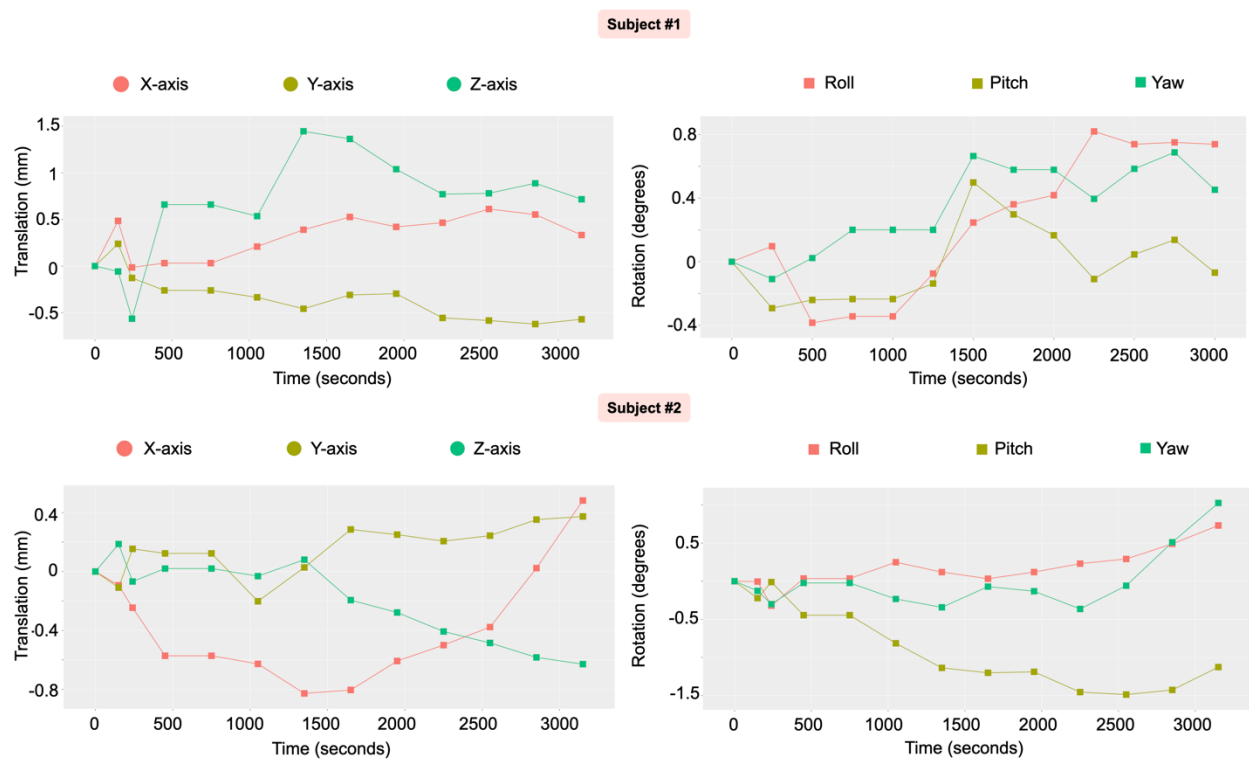
Supplementary Figure 2. Schematic representation of the patchGAN discriminator architecture. The inputs to the discriminator were patches of size 64 x 64 x 64 of the generator's output. The architecture consisted of 5 convolutional layers. The first down-sample layer produced 64 feature maps; this number was doubled at each subsequent down-sampling step. After the last layer, a convolution was applied to map to a 1-dimensional output, followed by a sigmoid activation to determine whether the input is a “real” full-dose PET image or an “artificial” image.



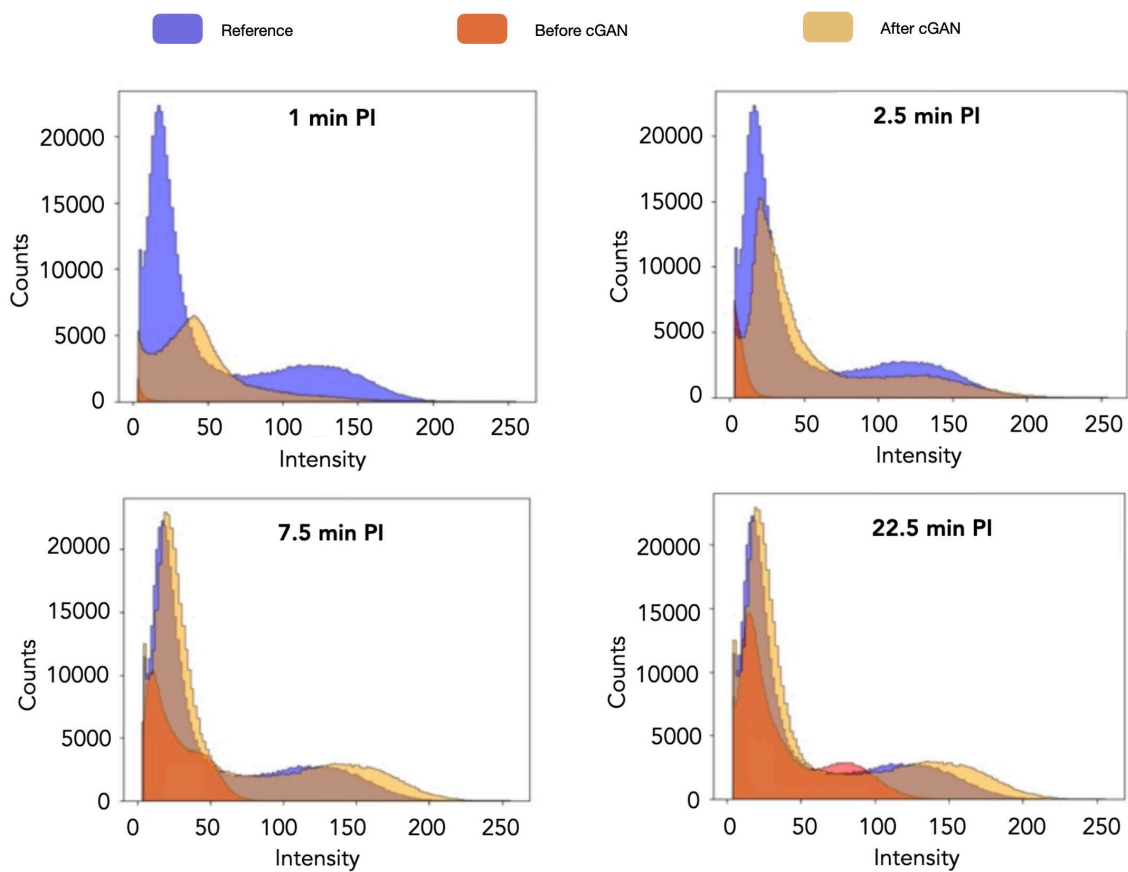
Supplementary Figure 3. Distribution of the motion parameters (rotation and translation), which were used to create synthetically generated data sets that included additional motion. The rotational parameters varied between 0 to 1 degree in all 3 orientations, and the translational parameters varied between 0 mm to 5 mm in x-, y- and z-directions.



Supplementary Figure 4. Representative input functions derived from two test subjects (#1 and #2) demonstrating the ability of cGAN processing to yield an accurate IDIF. Substantial motion artefacts were present in the early stages of the study (0 – 15min p.i., black time-activity curve) as determined by comparison with the reference standard of arterial blood sampling (orange time-activity curve). Application of frame-based motion correction results in an improved, but still suboptimal time-activity curves (green). cGAN-aided motion correction allows the non-invasive extraction of an IDIF (blue) that closely reproduces arterial blood samples.

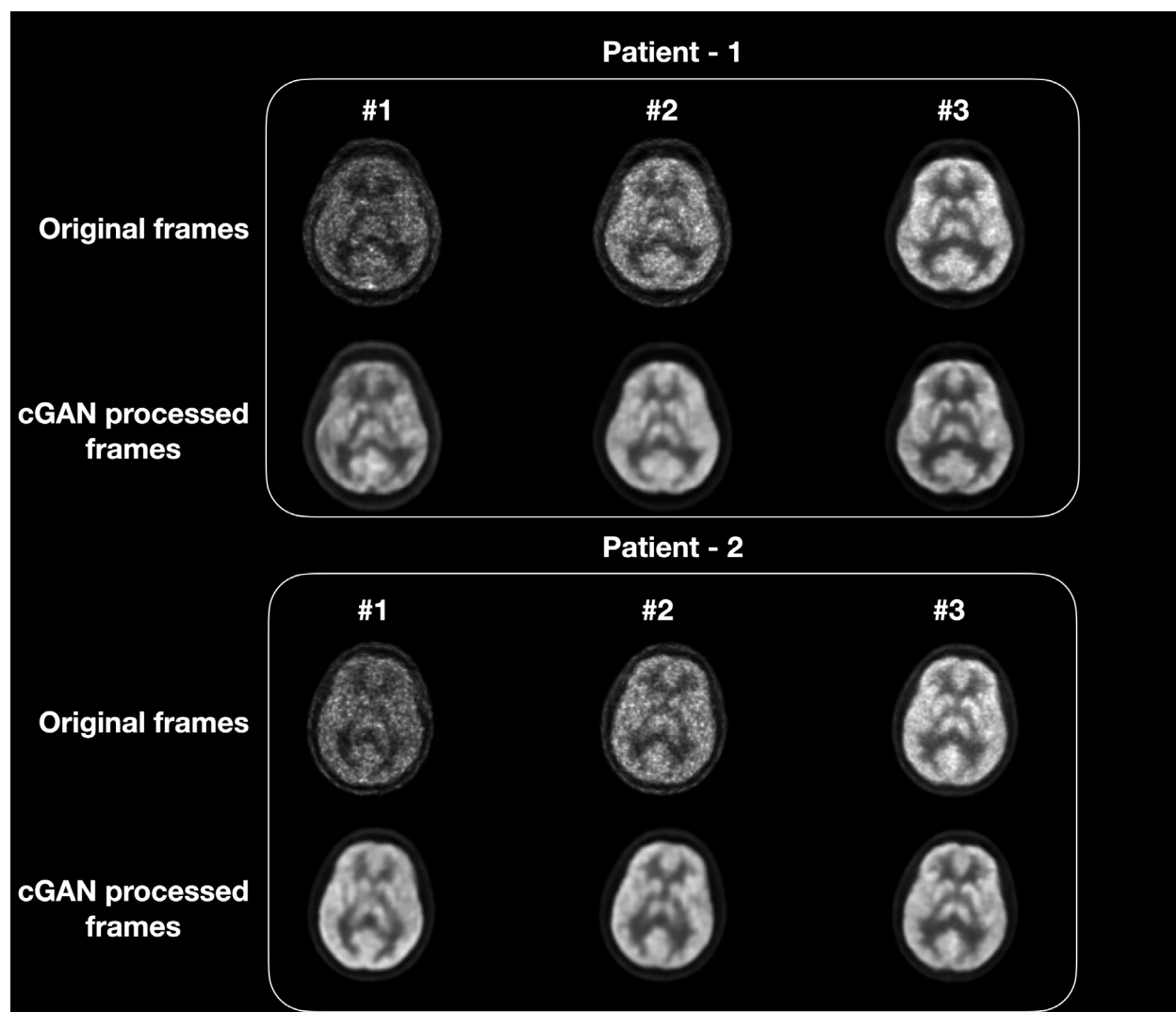


Supplementary Figure 5. 3D motion profile (3 translation and 3 rotation parameters) of representative subjects #1 and #2 from whom a non-invasive IDIF was derived (see also Suppl Fig 4).



Supplementary Figure 6. Comparison of the reference image histograms (blue) with the histogram of the low-count original image prior to cGAN (orange) and post cGAN (yellow) processing. It can be seen that in the cGAN processed histogram the frequency of high intensities in the image is overestimated (7.5 min p.i and 22.5 min p.i).





Supplementary Figure 7. The mappings (M-n) trained using  $^{18}\text{F}$ -FDG dynamic datasets from the Siemens Biograph mMR was applied to the dynamic datasets (2 scans) obtained from an external PET/CT system. Frames 1, 2 and 3 correspond to the following PET mid-times (2, 7, 35 min p.i). Although cGAN mappings were trained using data from the on-site PET/MR system, they were able to produce reasonable results with data obtained from an external PET/CT system, supporting the claim that cGAN mappings might be useable across imaging systems.