



Article scientifique

Article

2023

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

A cycle-consistent adversarial network for brain PET partial volume correction without prior anatomical information

Sanaat, Amirhossein; Shooli, Hossein; Böhringer, Andrew Stephen; Sadeghi, Maryam; Shiri Lord, Isaac; Salimi, Yazdan; Ginovart, Nathalie; Garibotto, Valentina; Arabi, Hossein; Zaidi, Habib

How to cite

SANAAT, Amirhossein et al. A cycle-consistent adversarial network for brain PET partial volume correction without prior anatomical information. In: European Journal of Nuclear Medicine and Molecular Imaging, 2023, vol. 50, n° 7, p. 1881–1896. doi: 10.1007/s00259-023-06152-0

This publication URL: <https://archive-ouverte.unige.ch/unige:167027>

Publication DOI: [10.1007/s00259-023-06152-0](https://doi.org/10.1007/s00259-023-06152-0)



A cycle-consistent adversarial network for brain PET partial volume correction without prior anatomical information

Amirhossein Sanaat¹ · Hossein Shooli² · Andrew Stephen Böhringer¹ · Maryam Sadeghi³ · Isaac Shiri¹ · Yazdan Salimi¹ · Nathalie Ginovart^{4,5,6} · Valentina Garibotto^{1,4} · Hossein Arabi¹ · Habib Zaidi^{1,4,7,8} 

Received: 25 November 2022 / Accepted: 12 February 2023
© The Author(s) 2023

Abstract

Purpose Partial volume effect (PVE) is a consequence of the limited spatial resolution of PET scanners. PVE can cause the intensity values of a particular voxel to be underestimated or overestimated due to the effect of surrounding tracer uptake. We propose a novel partial volume correction (PVC) technique to overcome the adverse effects of PVE on PET images.

Methods Two hundred and twelve clinical brain PET scans, including 50 ¹⁸F-Fluorodeoxyglucose (¹⁸F-FDG), 50 ¹⁸F-Flortaucipir, 36 ¹⁸F-Flutemetamol, and 76 ¹⁸F-Fluorodopa, and their corresponding T1-weighted MR images were enrolled in this study. The Iterative Yang technique was used for PVC as a reference or surrogate of the ground truth for evaluation. A cycle-consistent adversarial network (CycleGAN) was trained to directly map non-PVC PET images to PVC PET images. Quantitative analysis using various metrics, including structural similarity index (SSIM), root mean squared error (RMSE), and peak signal-to-noise ratio (PSNR), was performed. Furthermore, voxel-wise and region-wise-based correlations of activity concentration between the predicted and reference images were evaluated through joint histogram and Bland and Altman analysis. In addition, radiomic analysis was performed by calculating 20 radiomic features within 83 brain regions. Finally, a voxel-wise two-sample t-test was used to compare the predicted PVC PET images with reference PVC images for each radiotracer.

Results The Bland and Altman analysis showed the largest and smallest variance for ¹⁸F-FDG (95% CI: −0.29, +0.33 SUV, mean = 0.02 SUV) and ¹⁸F-Flutemetamol (95% CI: −0.26, +0.24 SUV, mean = −0.01 SUV), respectively. The PSNR was lowest (29.64 ± 1.13 dB) for ¹⁸F-FDG and highest (36.01 ± 3.26 dB) for ¹⁸F-Flutemetamol. The smallest and largest SSIM were achieved for ¹⁸F-FDG (0.93 ± 0.01) and ¹⁸F-Flutemetamol (0.97 ± 0.01), respectively. The average relative error for the kurtosis radiomic feature was 3.32%, 9.39%, 4.17%, and 4.55%, while it was 4.74%, 8.80%, 7.27%, and 6.81% for NGLDM_contrast feature for ¹⁸F-Flutemetamol, ¹⁸F-Fluorodopa, ¹⁸F-FDG, and ¹⁸F-Flortaucipir, respectively.

Conclusion An end-to-end CycleGAN PVC method was developed and evaluated. Our model generates PVC images from the original non-PVC PET images without requiring additional anatomical information, such as MRI or CT. Our model eliminates the need for accurate registration or segmentation or PET scanner system response characterization. In addition, no assumptions regarding anatomical structure size, homogeneity, boundary, or background level are required.

Keywords PET · Brain · Partial volume effect · Partial volume correction · Deep learning

Introduction

Over the recent decades, positron emission tomography (PET) imaging, among other molecular imaging modalities, has gained importance in preclinical, clinical, and research

fields. PET is widely used in the assessment of oncology patients, cardiac pathologies, and various neurological disorders, including Alzheimer's Disease (AD), Parkinson's Disease (PD), and epilepsy. PET provides functional information useful in the assessment of a variety of metabolic processes, such as tissue metabolism, protein accumulation, and neurotransmission pathways [1, 2]. Accurate and reliable quantification is a major strength of molecular PET imaging as it allows us to accurately assess molecular pathways and various diseases in their earliest phases. For instance, accurate localization and/or quantification of tracer uptake in malignant lesions is the basis for pre- and

This article is part of the Topical Collection on Advanced Image Analyses (Radiomics and Artificial Intelligence).

✉ Habib Zaidi
habib.zaidi@hcuge.ch

Extended author information available on the last page of the article

post-treatment evaluations in neurooncology. In addition, accurate delineation of tumor contours is crucial in monitoring treatment response and radiation therapy planning.

The limited spatial resolution and low signal-to-noise ratio are the main drawbacks of PET imaging, making accurate quantitative analysis a challenging task in clinical practice. The partial volume effect (PVE) results from the poor spatial resolution of PET scanners, typically in the range of 3.5 to 6 mm full-width-half-maximum (FWHM). As a result of PVE, the intensity of a particular voxel is affected not only by the tracer concentration of the tissue in which the voxel is located but also by the surrounding tissues/organs. In addition, the physical size and shape of the volume of interest (VOI) and its contrast relative to surrounding regions affect PVE. Therefore, correction for PVE is mandatory for reliable quantitative measurements of physiological parameters and image-derived metrics, such as the standardized uptake value (SUV) or tumor-to-background ratio (TBR) for specific VOIs. This is particularly relevant when the pathology itself affects the volume of the target regions, as is the case in neurodegenerative diseases which are typically associated with atrophy.

Partial volume correction (PVC) techniques can overcome the adverse effects of PVE on PET images. Studies have shown that PVC improves diagnostic accuracy and SUV quantification [3], estimation of tracer uptake in plaque in large vessels or in an atrophied gray matter [4], and measurement of ventricular mass [5], in addition to improving overall image quality for ^{18}F -Flortaucipir and amyloid PET tracers [6, 7]. Moreover, PVC PET images allow for the quantification of different physiologic processes in the brain, including cerebral blood flow, glucose metabolism, neuroreceptor binding, and tumor metabolism [8]. Applying PVC methods also proved to improve the statistical power in cross-sectional [9] and longitudinal [6] analyses in quantitative amyloid imaging. PVC can also eliminate confounding results in studies of aging [10] or atrophy effects in the brain [11, 12]. For instance, PVC prevents the underestimation of physiologic measurements due to the loss of cerebral volume resulting from healthy aging processes. A number of studies demonstrated that PVC improves clinical classification performance in AD [13] and PD [14] research. It can be concluded that PVC is necessary to ensure that measurements are truly quantitative for different regions within the brain. To this end, a number of PVC techniques have been developed and implemented with varying degrees of success [15–17].

Most popular PVC methods for brain PET imaging, such as Meltzer's method [15], Müller-Gärtner (MG) [16], or the geometric transfer matrix (GTM) method [17], typically require other imaging modalities, such as CT or MRI as a priori anatomical information. This dependence gives rise to a key drawback, namely the need for accurate co-registration of PET to CT or MR images. This dependency means that misregistration or inaccurate segmentation contributes to errors in PVC. Other

methods use the PET scanner's point spread function (PSF). The downside of these methods is that they require an accurate estimate of the spatially varying PSF, which might be difficult to measure [17]. Other methods require dedicated reconstruction software, which is readily not available for all PET/CT or PET/MRI systems. The mentioned downfalls of the current PVC methods highlight an unmet need for an end-to-end method to produce high-resolution PET images without the need for additional anatomical images and prior knowledge of PET scanner characteristics, tumor and VOI size, shape, or background level. Lu et al. assessed the impact of Müller-Gärtner (MG) and iterative Yang (IY) PVC on ^{11}C -UCB-J brain PET images for finding synaptic vesicle glycoprotein 2A (SV2A), which has been suggested as an indicator of synaptic density in Alzheimer's disease (AD) [18]. Onoue et al. compared CT and MRI-based PVC in brain ^{18}F -FDG PET and discussed the advantages of PVC using CT images [19]. An error propagation analysis was also performed for seven PVC methods by Oyama et al., where they showed around 30% bias in small and thin regions in AD patients with and without PVC [20].

Recently, machine learning (ML), especially deep learning (DL) as a subset of ML, has been increasingly used in various applications of PET imaging [21–23]. With advances in both DL algorithms and computational power, a paradigm shift favoring DL-based PVC approaches might be very promising toward the development of accurate and robust methods.

This work proposes a novel anatomical imaging-free DL-assisted PVC algorithm and evaluates its performance using clinical brain studies acquired with four PET neuroimaging radiotracers. The method is an end-to-end PVC pipeline, which inputs a low-resolution brain PET image to generate a high-quality PVC image, which does not require anatomical imaging and a priori knowledge of the PSF, VOI size, shape, or background level.

Materials and methods

PET/CT and MRI data acquisition

Patients undergoing a brain PET/CT/MRI scan collected between April 2017 and February 2020 at Geneva University Hospital were enrolled in this study. The study protocol was approved by the institution's ethics committee, and all patients gave written informed content. The two hundred and twelve patients dataset were acquired following injection of four different PET neuroimaging radiotracers (50 ^{18}F -FDG, 50 ^{18}F -Flortaucipir, 36 ^{18}F -Flutemetamol, and 76 ^{18}F -Fluorodopa). The corresponding CT and T1-weighted MR images were also used in this study. A combination of healthy patients and those diagnosed with different pathologies, such as neurodegenerative disease, cannabis use disorder, and internet gaming disorder, were considered for training the model to

increase the generalizability of our method. The corresponding demographic details are summarized in Table 1.

Attenuation and scatter-corrected PET images as well as T1-weighted MR images were acquired on the Biograph mCT scanner and 3 T MAGNETOM Skyra scanner (Siemens Healthcare, Erlangen, Germany), respectively. The PET scanning protocol for the different radiotracers, including injected activities, scan time durations, and delay times between injection and PET scanning, is summarized in Table 1. MRI data acquisition protocol was similar for the various radiotracers. The PET/CT/MRI scanning protocol details were summarized in Supplementary Table 1.

Data processing and image registration

After cropping PET and MR images, they were coregistered to the corresponding standard brain template defined into Montreal Neurological Institute (MNI) (Montreal Neurological Institute, McGill University) standard stereotactic space [24] using the 3D Slicer software [25]. An affine registration method with 12 degrees of freedom was employed for all images [26]. Because PET and CT images acquired on the PET/CT scanner were already registered, PET images were registered to the MNI

template, and the resulting registration matrix was applied to CT images. Subsequently, T1-weighted MRI was registered to CT images. All images were visually assessed to ensure accurate registration between PET, CT, and MR images.

Data augmentation

Since the number of cases for each radiotracer was not similar, the effect of dataset size on model performance was minimized using a previously developed augmentation method using the Laplacian blending (LB) technique, referred to as Robust-Deep [27], to increase the dataset size to a fixed number of 100 per radiotracer. The Robust-Deep technique increases the number of brain images by combining images of two different cases through a predefined mask to create a semi-realistic image, which can significantly enhance the robustness of the deep learning models.

Partial volume correction

The Iterative Yang (IY) technique [4] was selected from the PET-PVC toolbox [28] for PVC. Unlike region-based PVC

Table 1 Demographics of the patient population included in this study protocol

	Dataset (5 cross-validation)	Scanning protocol
¹⁸ F-FDG		
Number	100 (50 actual + 50 LB)	Delay between injection and scanning: 32 ± 6 min
Male/female	38/12	Injected activity: 208 ± 14 MBq
Age (mean ± SD)	69 ± 5	Scan duration: 20 min
Indication/diagnosis	Cognitive symptoms of possible neurodegenerative etiology	
¹⁸ F-FluoroDOPA		
Number	100 (50 actual + 50 LB)	Delay between injection and scanning: 0 min
Male/female	31/19	Injected activity: 185 ± 12 MBq
Age (mean ± SD)	25 ± 4	Scan duration: 90 min
Indication/diagnosis	Healthy patients and patients with a psychiatric disorder	
¹⁸ F-Flortaucipir		
Number	100 (36 actual + 64 LB)	Delay between injection and scanning: 76 ± 8 min
Male/Female	20/16	Injected activity: 205 ± 10 MBq
Age (mean ± SD)	53 ± 9	Scan duration: 30 min
Indication/diagnosis	Cognitive symptoms of possible neurodegenerative etiology	
¹⁸ F-Flutemetamol		
Number	100 (76 actual + 24 LB)	Delay between injection and scanning: 91 ± 6 min
Male/female	42/28	Injected activity: 199 ± 11 MBq
Age (mean ± SD)	66 ± 8	Scan duration: 20 min
Indication/diagnosis	Cognitive symptoms of possible neurodegenerative etiology	

Data augmentation using the Laplacian blending (LB) technique was used to increase the sample size

methods, where the corrections are only valid for voxels within a selected region to provide regional mean values (e.g., GTM, MGM), a voxel-by-voxel correction is applied to the whole image in the IY method. As such, the PVC image $f_{PVC}^{itr}(x)$ is estimated from the multiplication of the uncorrected PET image $f(x)$ and the ratio of artificial PET images $f_a^{itr}(x)$ and a blurred/smoothed version of this image (achieved by convolving $f_a^{itr}(x)$ with the PSF of the PET scanner):

$$f_{PVC}^{itr+1}(x) = f(x) \cdot \left[\frac{f_a^{itr}(x)}{f_a^{itr}(x) \otimes PSF} \right] \quad (1)$$

where the artificial PET images $f_a^{itr}(x)$ is renewed at each iteration by multiplying the average value of the artificial PET $f_a^{itr}(x)$ at j -th regions ($A_{j, f_{PVC}^{itr}(x)}$) and anatomical probability of j -th regions at location x $P_j(x)$, which is extracted from MR images:

$$f_a^{itr}(x) = \sum_{j=1}^{\#Regions} \left[A_{j, f_{PVC}^{itr}(x)} \cdot P_j(x) \right] \quad (2)$$

We initially considered the first PVC PET images as equal to the uncorrected PET images:

$$f_{PVC}^0(x) = f(x) \quad (3)$$

Ten iterations were used for PVC in this work. The FWHM of the 3D Gaussian convolution kernel was set to $3.0 \times 3.0 \times 3.0$ mm.

Network architecture

A Cycle-Consistent Generative Adversarial Network (CycleGAN), which learns a function to translate non-PVC PET images to PVC PET images (Fig. 1), was used in this work. The model consists of two GANs, including

Fig. 1 Schematic architecture of the cycle-consistent generative adversarial network (CycleGAN) model used for PET-PVC synthesis. The top panel depicts the training process, whereas the down panel shows the testing process

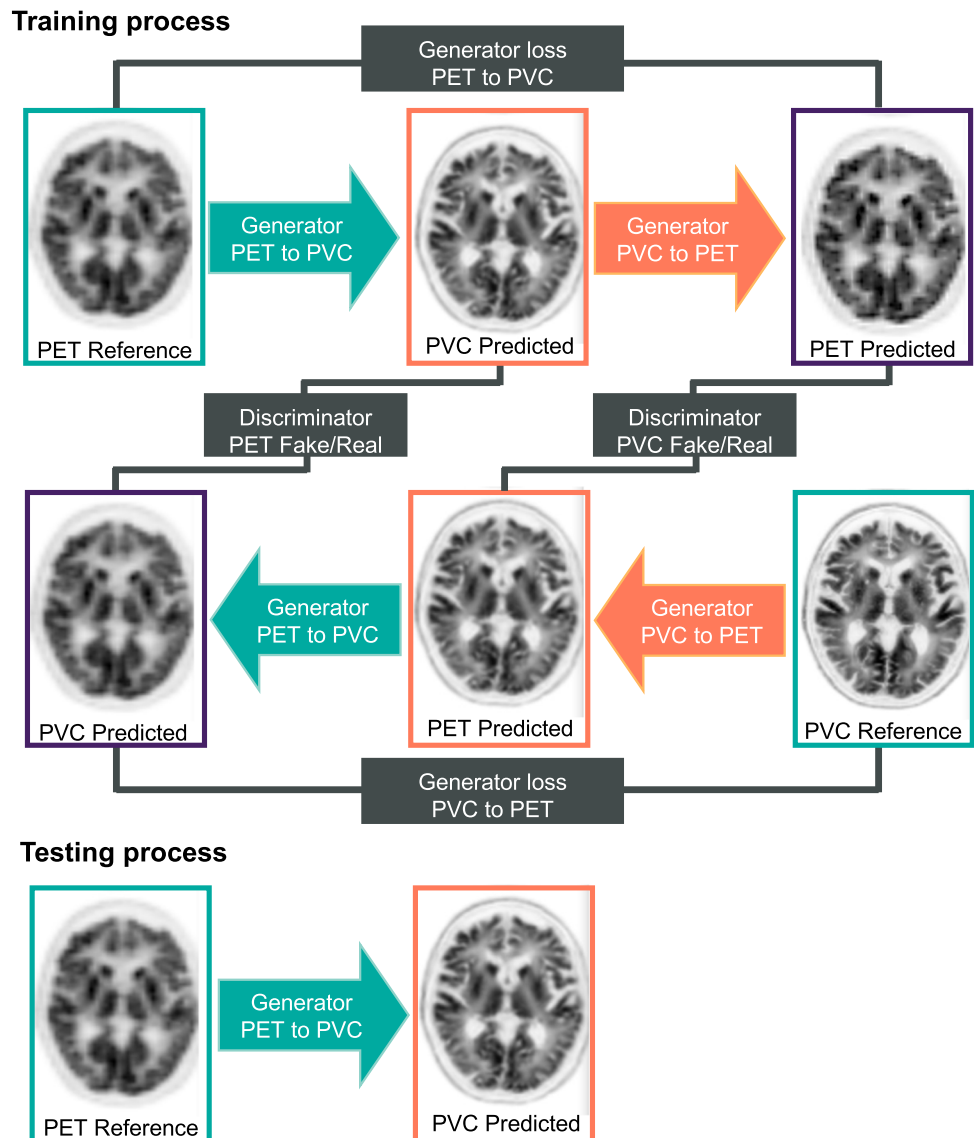


Table 2 Summary of the 20 radiomic features belonging to the 7 main categories calculated for the 83 brain regions

Radiomic feature category	Radiomic feature names
Conventional indices	SUV _{mean} SUV _{std} SUV _{max} SUV Q1 SUV Q2 SUV Q3 TLG (mL)
First-order features – histogram	Kurtosis Entropy_log10 Entropy_log2 Uniformity
Area under the curve of cumulative SUV histogram	AUC_CSH
Gray-level co-occurrence matrix (GLCM)	Homogeneity Energy Dissimilarity
Gray-level run length matrix (GLRLM)	Run percentage (RP)
Neighborhood gray-level difference matrix (NGLDM)	Contrast
Gray-level zone length matrix (GLZLM)	Short-zone emphasis (SZE) Low gray-level zone emphasis (LGZE) High gray-level zone emphasis (HGZE)

four main model architectures – two generators and two discriminators – as described in detail in Supplementary Table 2. The model training and evaluation were performed on an NVIDIA 2080Ti GPU with 11 GB memory running under Windows 10 operating system. We trained four different models with five-fold cross-validation for each radiotracer.

Visual and quantitative evaluation for the test dataset

All images, namely original PVC and DL-predicted PVC images, were visually inspected to assess overall image quality and the presence of potential alterations and artifacts in tracer distribution.

Quantitative analysis was performed by calculating well-established metrics, such as structural similarity index metrics (SSIM), root mean squared error (RMSE), and peak signal-to-noise ratio (PSNR), showing geometric similarity between the DL-predicted and ground truth images, the level of error/noise, and the strength of the signal-to-noise ratio, respectively. Voxel-wise and region-wise activity concentration correlations between the DL-predicted and reference PET images were evaluated through joint histogram and Bland and Altman analysis. For region-wise analysis, 20 radiomic features from 83 brain regions were extracted through registering the reference and predicted images to the Hammers N30R83 brain atlas [29].

Radiomics analysis

The image biomarker standardization initiative (IBSI) [30] compliant LIFEx software [31] was used for the extraction of the radiomic features. The list of the extracted radiomic features and their related categories are presented in Table 2. The relative bias between radiomic features extracted from the reference and DL-predicted PVC PET images were calculated over all radiotracers.

Voxel-based statistical analysis

All T1-weighted, original non-PVC PVC, and DL-predicted PVC images for all PET neuroimaging tracers were pre-processed using FSL (FMRIB Software Library v6.0.1, Analysis Group, FMRIB, Oxford, UK). In each step, we initially pre-processed T1-weighted images and then applied transformation matrices to the original and DL-predicted PVC images. Therefore, the original non-PVC PVC and DL-predicted PVC PET images were identically pre-processed for each patient.

First, brain tissue was extracted from T1-weighted images using the BET function implemented within FSL (Brain Extraction Tool, FSL). Subsequently, skull-stripped T1-weighted images were used as a mask to extract brain tissue both from the original non-PVC PVC and DL-predicted PVC PET images for each patient. Afterward, T1-weighted images were registered to MNI standard space using the FLIRT function (FMRIB's Linear Image Registration Tool, FSL). Then, the original non-PVC PVC

and DL-predicted PVC PET images of each patient were registered to MNI space via FLIRT using the same transformation matrix employed for registering the T1-weighted image of that subject. We applied a linear image registration method that does not change the voxels' values without smoothing to minimize the effect of pre-processing on the results. In each step, the outcome of pre-processing procedures was manually checked for potential errors, and appropriate corrections were performed when needed.

After these pre-processing steps, a mass univariate methodology of Statistical Parametric Mapping (SPM12; Wellcome Centre for Human Neuroimaging, UCL, UK) was used to perform a voxel-wise two-sample t-test that compared the DL-predicted PVC with reference PVC PET images for each tracer dataset [32]. This analysis identifies voxel clusters with statistically significant differences in the DL-predicted PVC images compared to the reference PVC PET images. Statistical significance was determined at a voxel-wise threshold of $p < 0.05$ (family-wise error corrected), and no voxel clusters exceeding the threshold were determined.

Results

All DL-predicted PVC PET images were considered visually adequate and comparable to the corresponding original PVC PET images, as exemplified in Figs. 2 and 3. In particular, Fig. 2 illustrates three different transaxial slices of MRI, non-PVC PET, reference MRI-based PVC PET, and the DL-predicted PVC PET images as well as the corresponding bias maps for the four different patients/radiotracers. The effectiveness of our model in terms of highlighting and enhancing the contours of the anatomical information in the DL-predicted PVC PET images is observable. It is worth noting that the DL-predicted PVC PET images are synthesized from only PET images as opposed to reference PVC PET which is generated from both MR and PET images. Figure 3 presents four abnormal cases depicting some artifacts and anatomical information loss in MR images, likely because of probable patient motion and the existence of metallic objects, such as a dental crown or a ventriculoperitoneal shunt or post-operative changes, causing artifacts in MR images. The reference PVC PET generated from MR

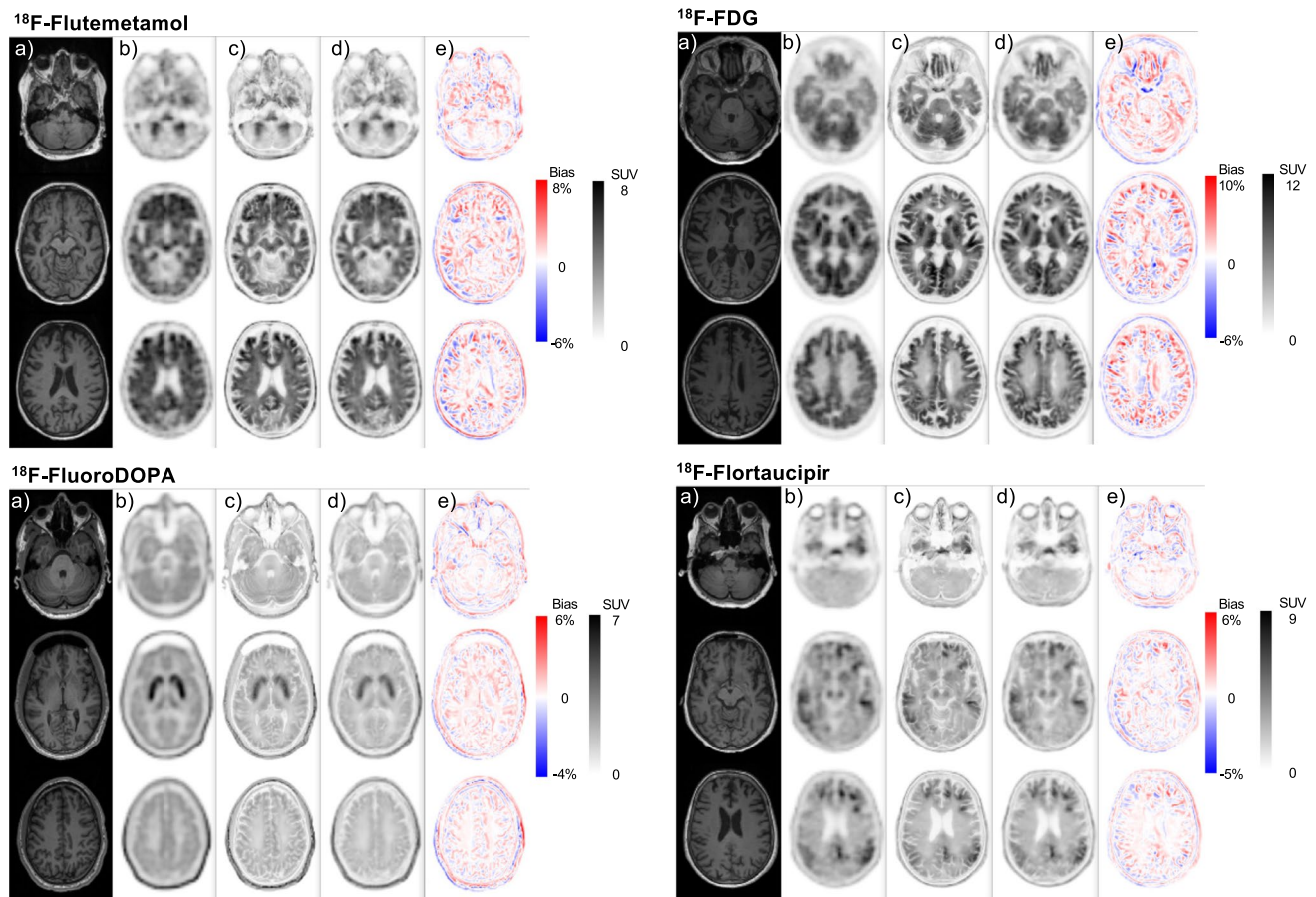


Fig. 2 Representative slices of multi-tracer brain PET images of various patients showing (a) coregistered T1-weighted MRI, (b) non-PVC PET images, (c) reference corresponding MRI-guided PVC PET images, (d) DL-predicted PVC PET images, and (e) the corresponding bias maps

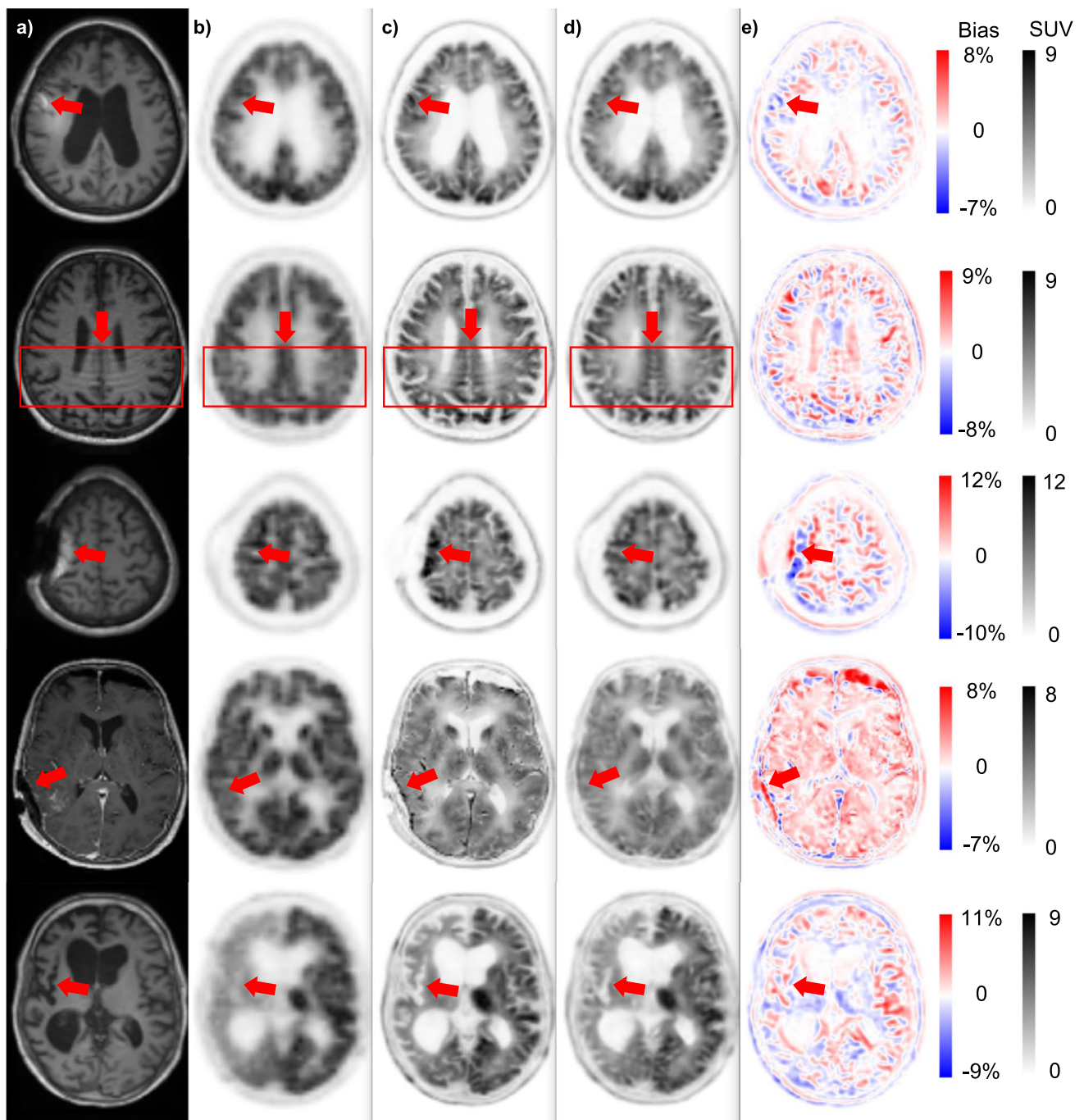


Fig. 3 Five special cases of patients presenting with artifacts in MR images (first and second row), anatomical abnormalities or existence of external objects (third and fourth row), and ununiform activity distribution in PET images (last row): (a) coregistered T1-weighted

MRI, (b) non-PVC PET images, (c) reference corresponding MRI-guided PVC PET images, (d) DL-predicted PVC PET images, and (e) the corresponding bias maps

and PET images highlights the propagation of MR artifacts/abnormalities into PVC PET images, while the DL-based PVC images are immune to these artifacts.

The scatter and Bland and Altman plots for 83 brain regions over the test dataset for each radiotracer are illustrated in Fig. 4. For all radiotracers, the scatter plots show

high correlations between SUVs calculated on DL-based PVC PET images and those on reference MRI-based PVC PET images, with a correlation coefficient (R^2) larger than 0.98 and RMSE smaller than 0.15 SUV. The Bland and Altman plots show that the largest variance in terms of mean error and confidence interval (CI) was achieved for ^{18}F -FDG

Fig. 4 The Bland–Altman plots (right panel) and scatter plots (left panel) of SUV_{mean} differences in the 83 brain regions for various tracers. In the Bland–Altman plots, the black solid and dashed lines denote the mean and 95% confidence interval (CI) of the SUV differences, respectively. In the scatter plots, the black solid and dashed lines denote the linear regression line and identity line, respectively

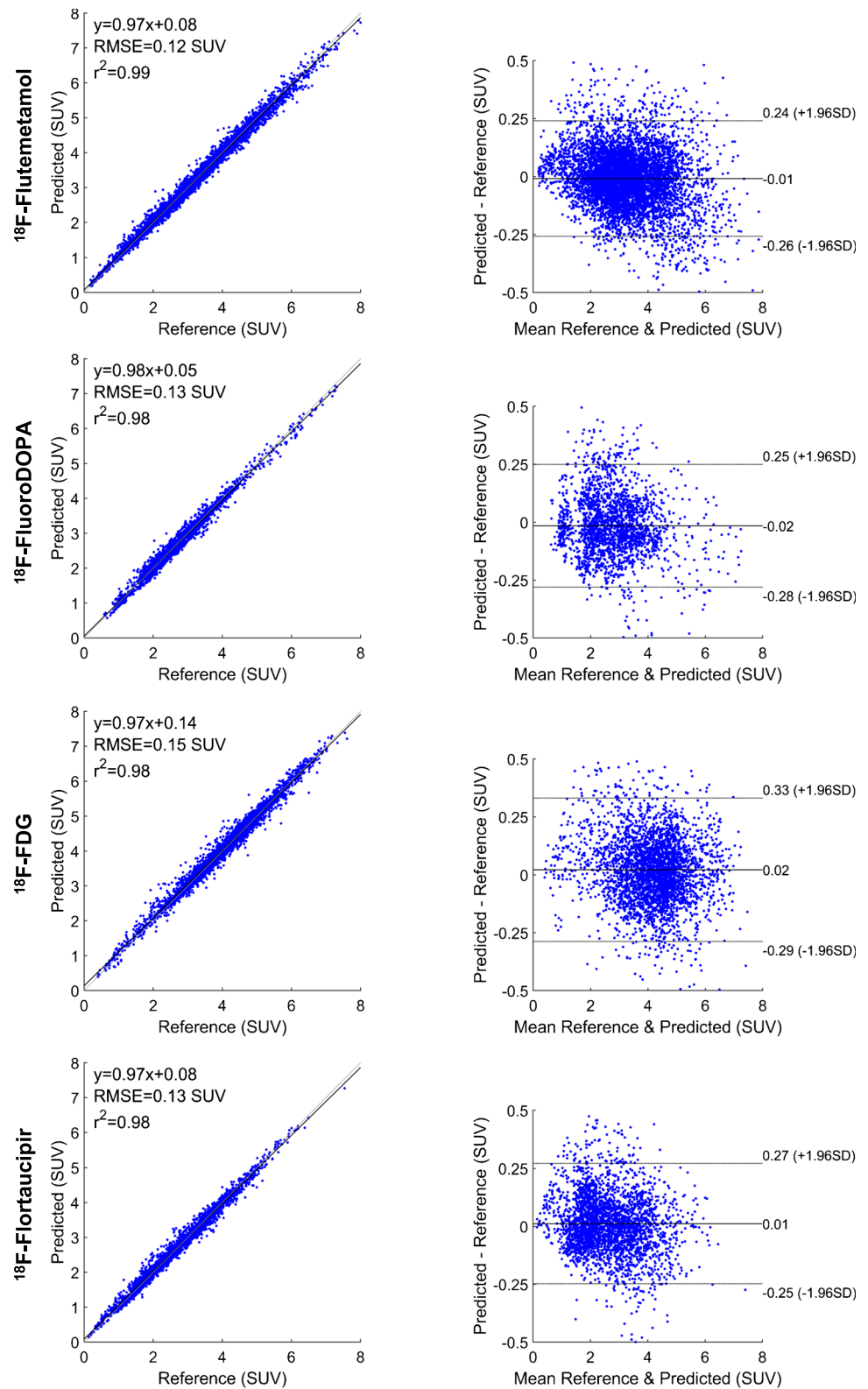


Table 3 Comparison of the results obtained from the analysis of image quality in reference and DL-predicted PVC PET images for the different tracers in the test dataset

Tracer	SSIM	PSNR (dB)	RMSE
^{18}F -FDG	0.93 ± 0.01	29.64 ± 1.13	$6.16 \times 10^{-6} \pm 1.48 \times 10^{-6}$
^{18}F -Flortaucipir	0.97 ± 0.02	34.97 ± 1.97	$2.31 \times 10^{-6} \pm 1.13 \times 10^{-6}$
^{18}F -Flutemetamol	0.97 ± 0.01	33.42 ± 1.92	$2.82 \times 10^{-6} \pm 1.31 \times 10^{-6}$
^{18}F -FluoroDOPA	0.94 ± 0.02	36.01 ± 3.26	$1.72 \times 10^{-6} \pm 1.55 \times 10^{-6}$

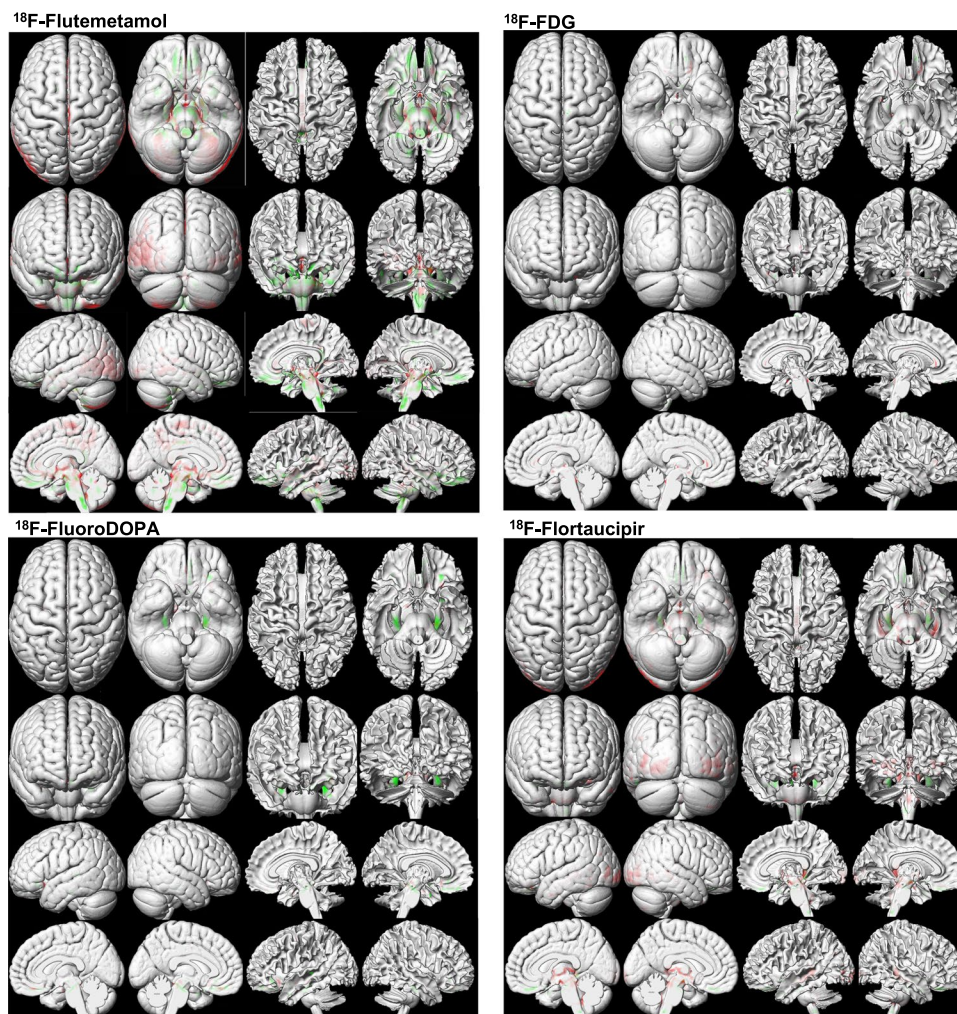
SSIM, structural similarity index metrics; PSNR, peak signal-to-noise ratio; RMSE, root mean squared error

(95% CI: $-0.29, +0.33$ SUV, mean = 0.02 SUV), whereas the smallest variance was obtained for ^{18}F -Flutemetamol (95% CI: $-0.26, +0.24$ SUV, mean = -0.01 SUV).

Table 3 summarizes the outcome of quantitative evaluation metrics, including SSIM, PSNR, and RMSE for the different radiotracers. The PSNR varies from 29.64 ± 1.13 dB for ^{18}F -FDG to 36.01 ± 3.26 dB for ^{18}F -Flutemetamol. The smallest SSIM was achieved for ^{18}F -FDG (0.93 ± 0.01), whereas the largest SSIM was obtained for ^{18}F -Flutemetamol (0.97 ± 0.01). 3D-rendered views of voxel-wise statistical analysis of reference and DL-predicted PVC PET images for each PET tracer are

shown in Fig. 5. The red and green regions represent voxels with statistically significant overestimation and underestimation of tracer uptake, respectively. In Fig. 6, clusters presenting with statistically significant differences between the DL-predicted and reference PVC PET images are depicted. By comparing the DL-based images with the original images, we have classified errors into two categories, namely overestimation and underestimation. The first describes the DL-predicted PVC PET voxels with a significantly lower value compared with the reference PVC PET voxels, while the latter describes voxels with a significantly higher value compared with the reference value.

Fig. 5 3D-rendered views of voxel-wise analysis of DL-predicted PVC images compared with reference PVC PET images for the different PET tracers. The images highlight voxel clusters with statistically significant differences compared with reference PVC PET images. In comparison to reference PVC images, the red regions represent voxels with statistically significant overestimation, while the green regions indicate voxels with statistically significant underestimation of tracer uptake



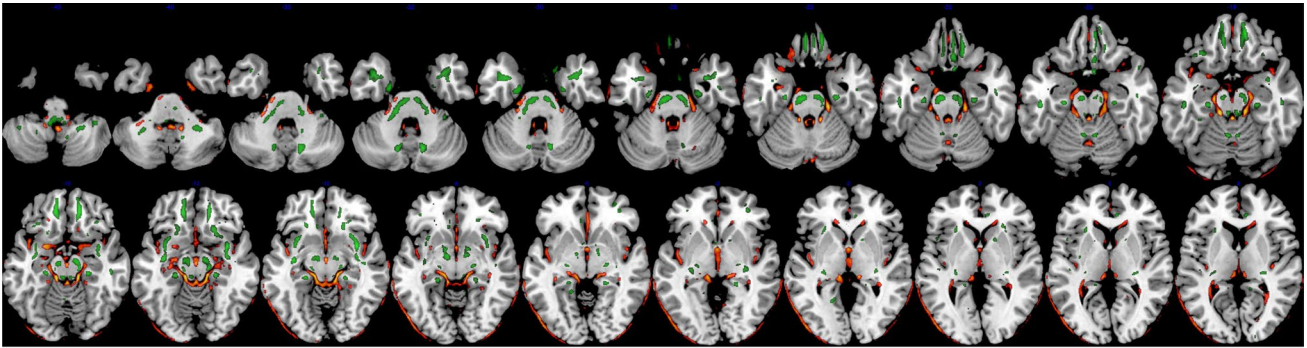
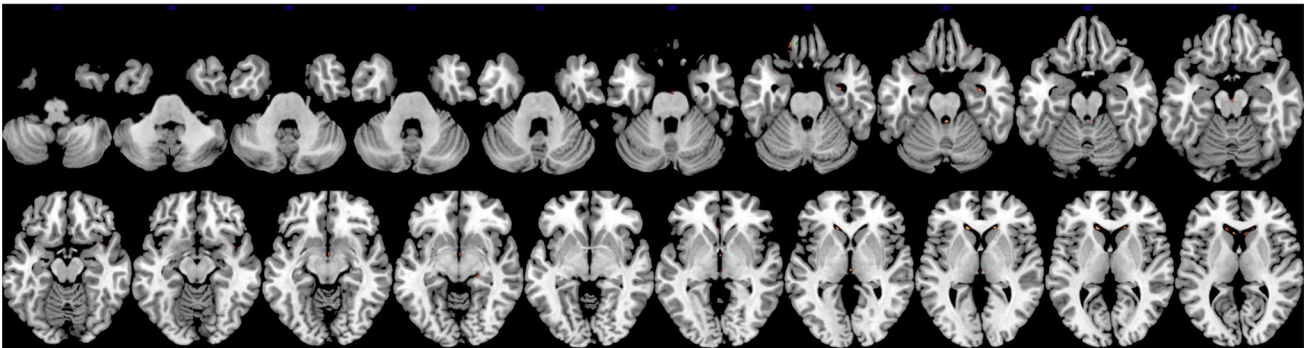
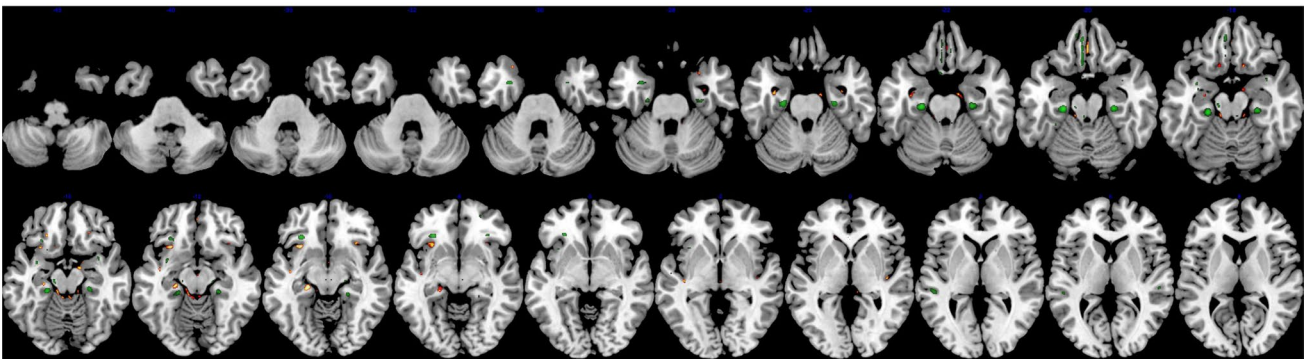
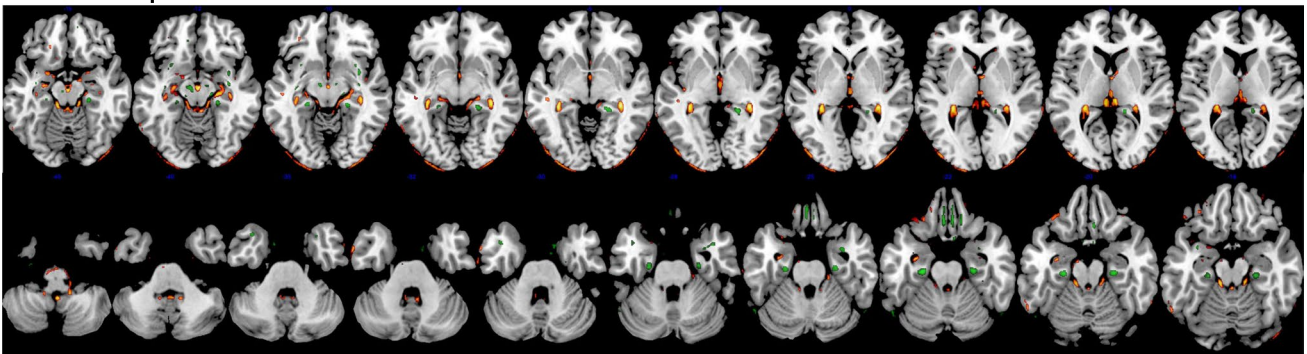
^{18}F -Flutemetamol **^{18}F -FluoroDOPA** **^{18}F -FDG** **^{18}F -Flortaucipir**

Fig. 6 Multi-slice views of voxel-wise analysis of DL-predicted PVC PET images compared with reference PVC PET images for the different neuroimaging tracers. These images show voxel clusters with statistically significant differences in DL-predicted PVC PET images compared with reference PVC PET images at different slices of

the brain. In comparison to reference PVC PET images, red/yellow regions represent voxel clusters with statistically significant overestimation, while green regions indicate voxel clusters with statistically significant underestimation of tracer uptake

Table 4 Voxel-based statistical analysis between DL-predicted and original MRI-guided PVC for the different PET tracer

Tracer	Total number of voxels	Voxel number (%)	Mean T-value (\pm SD)
Underestimation*			
^{18}F -Flortaucipir	1,827,095	1998 (0.10%)	6.11 (\pm 0.95)
^{18}F -Flutemetamol	1,827,095	17,846 (0.97%)	6.33 (\pm 0.81)
^{18}F -FDG	1,820,714	1730 (0.09%)	7.03 (\pm 1.39)
^{18}F -FluoroDOPA	1,824,363	80 (0.00%)	6.27 (\pm 0.27)
Overestimation**			
^{18}F -Flortaucipir	1,827,095	10,258 (0.56%)	6.09 (\pm 1.09)
^{18}F -Flutemetamol	1,827,095	27,419 (1.50%)	6.49 (\pm 1.06)
^{18}F -FDG	1,820,714	878 (0.04%)	6.31 (\pm 0.58)
^{18}F -FluoroDOPA	1,824,363	400 (0.02%)	6.48 (\pm 0.40)

The models using ^{18}F -FDG and ^{18}F -FluoroDOPA images for predicting corresponding PVC images had fewer voxels with statistically significant differences, yielding better performance. Conversely, models using ^{18}F -Flutemetamol and ^{18}F -Flortaucipir images had more voxels with statistically significant differences, demonstrating worse prediction compared to ^{18}F -FDG or ^{18}F -FluoroDOPA. Here, “voxel number” represents the extent of a difference with statistical significance, and “T-values” represent the degree of a difference with statistical significance

FEW, family-wise error

*Voxels with significantly lower values (model underestimation); *p*-value (0.05, FWE corrected)

**Voxels with significantly higher values (model underestimation); *p*-value (0.05, FWE corrected)

The DL model for ^{18}F -FDG and ^{18}F -FluoroDOPA datasets yielded a lower number of voxels with statistically significant differences compared with model performance in ^{18}F -Flutemetamol and ^{18}F -Flortaucipir datasets (Table 4).

The joint voxel-wise histogram analysis between reference and DL-predicted PVC PET images are depicted in Supplementary Fig. 1. The results are in good agreement with region-wise scatter plots. Figure 7 shows the relative error heat maps for 20 radiomic features and 83 regions for the different radiotracers. For a more concise presentation of the heat map, we reported the average of the left and right regions. The complete heat map for the 83 regions is depicted in Supplementary Figs. 2 and 3 to highlight abnormal cases where the left and right regions have different significantly different errors. The maximum underestimation and overestimation errors for each radiotracer can be appreciated from their corresponding color bar. It can be seen that the largest underestimation and overestimation is around 10% for ^{18}F -FluoroDOPA. With this radiotracer, the SUV was mostly underestimated in the DL-predicted PVC PET images for all radiomic features, except gray-level zone length matrix low gray-level zone emphasis. The average relative error for the kurtosis radiomic feature was 3.32%, 9.39%, 4.17%, and 4.55%, whereas it was 4.74%,

8.80%, 7.27%, and 6.81% for NGLDM_contrast feature for ^{18}F -Flutemetamol, ^{18}F -FluoroDOPA, ^{18}F -FDG, and ^{18}F -Flortaucipir, respectively. The average relative error of HISTO_energy_Uniformity, a feature depicting the strength of the signal, varied from 2.81%, 5.93%, 4.30%, and 3.93% for ^{18}F -Flutemetamol, ^{18}F -FluoroDOPA, ^{18}F -FDG, and ^{18}F -Flortaucipir, respectively.

Discussion

There is a growing interest in applying PVC for PET image interpretation and for quantifying various physiological parameters of interest in clinical and research settings. A variety of PVC algorithms have been developed; however, they are not yet widely applied in the clinical setting. One possible explanation for this fact could be that most available algorithms rely on certain assumptions that introduce uncertainty in the computation and ensuing quantification and require extra-anatomical images, such as CT and MRI. Moreover, additional imaging modalities are not always available; the radiation dose burden from CT and the acquisition time and cost of MRI considerably limit the clinical adoption of these techniques.

Two of the most popular PVC algorithms, namely MG and GTM, rely on anatomical/structural information provided by other imaging modalities, such as CT or MRI. Anatomically based methods assume perfect registration and segmentation of multimodal images prior to the application of PVC. In previous studies, the deleterious effect of co-registration errors [33] and segmentation errors [34, 35] on PVC implementation have been investigated and reported, specifically in the context of brain imaging [17, 36–38]. Quarantelli et al. [37] showed that, of all possible sources of error, misregistration errors demonstrated the most substantial impact on the accuracy of PVC in brain PET imaging.

An alternative to these strategies is iterative deconvolution methods [39, 40], which do not require anatomical information or assumptions regarding surrounding structures, tumor size, homogeneity, or background. One drawback of deconvolution-based methods is that they can amplify the high-frequency content of images, thus resulting in increased image noise [41]. As a result, ideal/perfect PVC algorithms appear problematic to achieve [11]. In addition, similar to other PVC methods, deconvolution-based methods still need to incorporate the scanner's PSF in the reconstruction process [42–44]. As mentioned earlier, accurate characterization of the scanner's response function could be challenging as it is spatially variable, object-dependent, and can be affected by reconstruction parameters [7]. It has been shown that any PSF mismatch might be critical [28, 45].

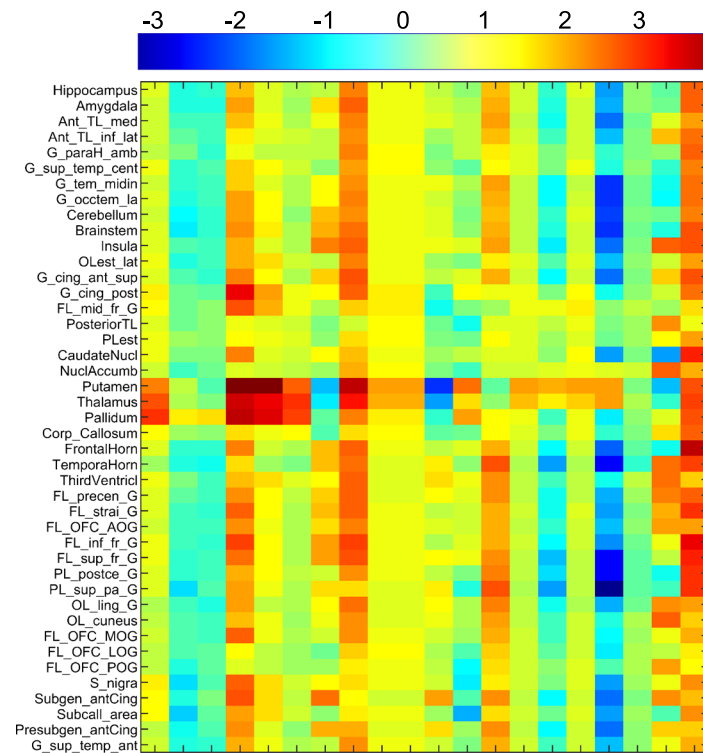
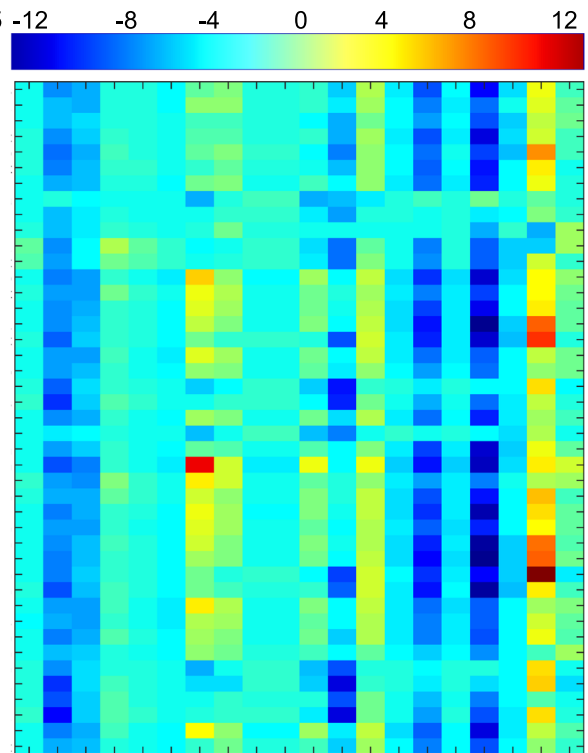
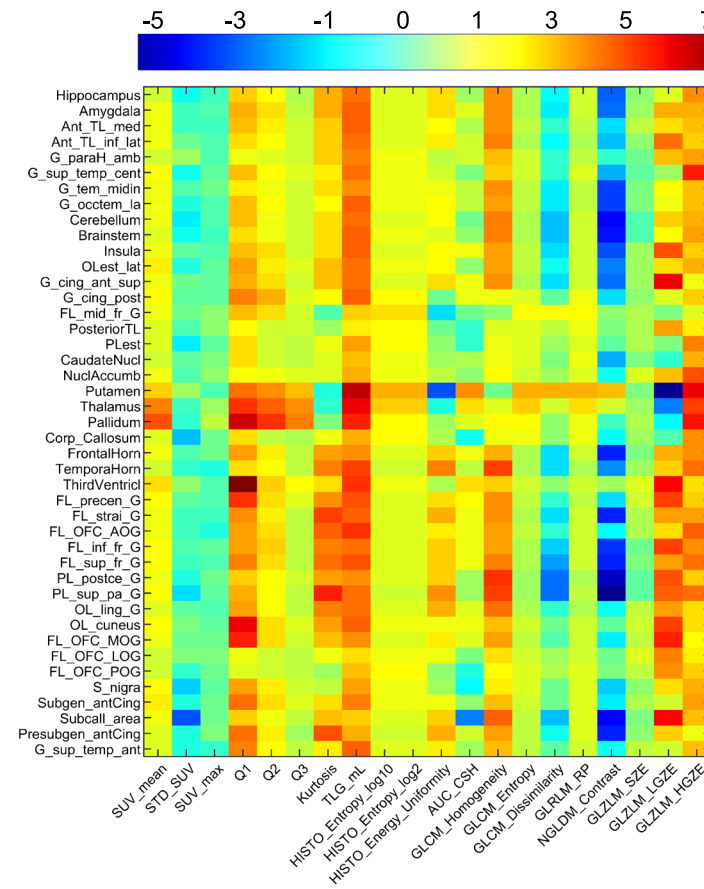
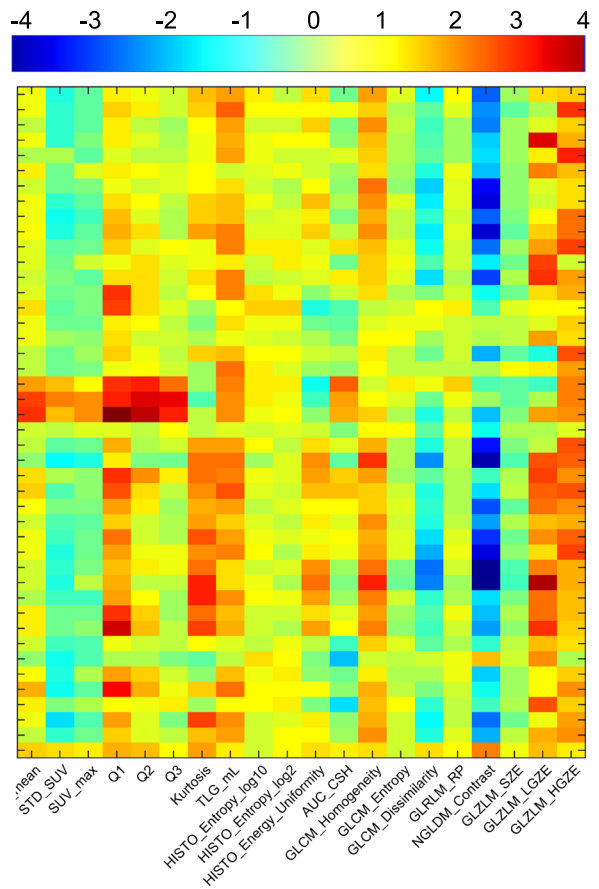
^{18}F -Flutemetamol **^{18}F -FluoroDOPA** **^{18}F -FDG** **^{18}F -Flortaucipir**

Fig. 7 Heat maps of the relative error of 20 radiomics features calculated across 43 brain regions (for better presentation, the average of right and left regions were reported) for DL-predicted PVC PET images with respect to reference PVC PET images

Radiomic features analysis evaluates the consistency and robustness of existing patterns in DL-predicted and reference PVC PET images. Considering the relatively poor spatial resolution of clinical PET systems and the importance of PVE in brain PET, conventional radiomic features, such as SUV_{max} , SUV_{mean} , and total lesion glycolysis (TLG), are expected to be significantly impacted by PVC. Furthermore, high-order features, such as GLZLM which represent small regions/patterns with low gray levels, are essential to evaluate the impact of PVC since PVE can lead to higher bias in small structures. Although our results highlight the importance of radiomic features for the assessment of PVC methods, separate studies are necessary to further understand the relevance of radiomics analysis.

Other assumptions include homogeneity of tracer distribution in a region or tissue component or homogeneous VOI [46, 47]. However, since the VOIs can be very heterogeneous in practice, the homogeneity assumption can introduce uncertainty and bias in parameter estimates [48]. In most voxel-based methods, the correction is valid only for voxels within the target region and requires initial information about the mean or relative mean values in various regions [46]. Region-based methods [42, 49] require manual VOI definition, which suffers from inter- and intra-observer variability. This might potentially lead to different VOI definitions for the same target [50, 51], where the difference in delineation can go up to 15 mm in diameter [52, 53]. In addition, some PVC algorithms require dedicated reconstruction software [42, 54] or extensive parametrization [7, 40, 55].

Research and development efforts are still being spent to tackle the limitations of currently available PVC algorithms. To encourage the clinical community to adopt PVC methods as part of standard processing procedures, more robust and straightforward methods must be developed and made available. It is essential to develop techniques that can be easily integrated, take as few assumptions as possible, and require as little parameter setting as possible.

Similar to other application fields, especially computer vision, DL can be helpful in tackling different problems encountered in PET imaging [56–59]. However, to the best of our knowledge, no DL-based method has been proposed to address the PVE problem in brain PET to date. Application in other body regions, e.g., in clinical oncology, is very sparse, with only a few studies so far [60]. We proposed a method that consists of an end-to-end DL-based pipeline to generate PVC PET images without the need for additional anatomical imaging modality. In addition, it does not depend on any aforementioned underlying assumptions

and eliminates the need for prior information, such as VOI size, homogeneity, or regional mean value. We trained and evaluated our proposed model in 83 brain regions defined on a template for various PET neuroimaging radiotracers. The evaluation demonstrated excellent quantitative and qualitative performance. In addition, our method is not affected by the limitations or artifacts present in other imaging modalities or the registration and segmentation inaccuracies commonly existing in alternative methods. One limitation of the current study is that the data were not multi-institutional and were instead collected from a single site. Related to and as a consequence of this, the images were also acquired on the same PET and MRI scanner models. This might affect the generalizability of the model that needs to be addressed in future studies through the use of a more diverse dataset from multiple institutions to further enhance the robustness of the model. Using images acquired on different PET scanners and using different acquisition and reconstruction protocols might improve the robustness and reproducibility of the model, thus leading to better performance. In addition, due to the differing sizes of the datasets for each radiotracer, data augmentation was required. Though this was beneficial in reducing the effect of sample size and increasing the robustness of the model, it may introduce some additional bias. Eliminating the need for additional imaging modalities might be particularly useful in cases where these other modalities are not available or are available but have been acquired in other conditions (e.g., post-operative) or with an important time delay or harbor artifacts that could then be transferred to PET images, as exemplified in Fig. 3. We hope that such end-to-end approaches will facilitate the implementation of PVC in routine clinical setting owing to ease of implementation on different systems. Another limitation of the current study is the absence of an ideal ground truth for the assessment of the proposed PVC technique. The MRI-based PVC method used in this work as a surrogate of the ground truth does not reflect ideal PVC PET images. Despite the advantages of simulations where the ground truth is available for evaluation [60], no simulations/phantoms are capable of perfectly mimicking clinical scenarios. Our model performed better if it was fed with PET images in MNI space. The normalization to MNI space can be automated through simple coding to transfer the images from native space to standard space. This will enable the user to feed the model with images in the native space directly.

PVC has been shown to improve diagnostic accuracy in conditions associated with atrophy and in small brain regions [61]. An added clinical value is also expected in the evaluation of small focal abnormalities, namely the localization of epileptic foci or in the detection of small malignant lesions [62]. Our results demonstrated that the proposed approach provides quantitative accuracy equivalent to alternative approaches without the need for anatomical images.

Conclusion

This work presents an end-to-end anatomical imaging-free DL-based PVC algorithm to correct for PVE in brain PET imaging. The technique is efficient because it eliminates the need for accurate registration or segmentation or PET scanner response function characterization. In addition, no assumptions regarding VOI size, homogeneity, boundary, or background level are required. The proposed approach fits most situations encountered in the clinical setting and provides sufficient training data. Moreover, it is relatively less sensitive to minor errors that may affect intersubject comparisons and thus is more robust. Given the post-reconstruction nature of the technique, it can be used on existing clinical PET scanners to improve PET's quantitative accuracy. The qualitative and quantitative performance of the proposed method demonstrated its potential in clinical brain PET studies using various neuroimaging molecular imaging probes. The achieved performance and robustness might make the proposed approach a good candidate for the incorporation of PVC in routine clinical practice.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00259-023-06152-0>.

Acknowledgements This work was supported by the Swiss National Science Foundation under Grants No. SNSF 320030_176052, 185028, 188355, 169876, and 31003A_179373, the Louis-Jeantet Foundation with contributions of the Clinical Research Center, University Hospital and Faculty of Medicine, University of Geneva, the Velux Foundation, and the Schmidheiny Foundation. VG received research/teaching support through her institution from Siemens Healthineers, GE Healthcare, Roche, Merck, Cerveau Technologies, and Life Molecular Imaging. Avid radiopharmaceuticals provided access to the ^{18}F -Flortaucipir radiotracer but were not involved in data analysis or interpretation.

Funding Open access funding provided by University of Geneva.

Data availability Data used in this work are not available owing to privacy/ethical restrictions.

Declarations

Ethics approval and consent to participate All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in

the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Langbein T, Weber WA, Eiber M. Future of theranostics: an outlook on precision oncology in nuclear medicine. *J Nucl Med*. 2019;60:13s–s19.
- Villemagne VL, Barkhof F, Garibotto V, Landau SM, Nordberg A, van Berckel BNM. Molecular imaging approaches in dementia. *Radiology*. 2021;298:517–30.
- Frouin V, Comtat C, Reilhac A, Gregoire MC. Correction of partial-volume effect for PET striatal imaging: fast implementation and study of robustness. *J Nucl Med*. 2002;43:1715–26.
- Erlandsson K, Buvat I, Pretorius PH, Thomas BA, Hutton BF. A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Phys Med Biol*. 2012;57:R119–59.
- Espe EKS, Bendiksen BA, Zhang L, Sjaastad I. Analysis of right ventricular mass from magnetic resonance imaging data: a simple post-processing algorithm for correction of partial-volume effects. *Am J Physiol Heart Circ Physiol*. 2021;320:H912–22.
- Su Y, Blazey TM, Snyder AZ, Raichle ME, Marcus DS, Ances BM, et al. Partial volume correction in quantitative amyloid imaging. *Neuroimage*. 2015;107:55–64.
- Thomas BA, Erlandsson K, Modat M, Thurfjell L, Vandenberghe R, Ourselin S, et al. The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease. *Eur J Nucl Med Mol Imaging*. 2011;38:1104–19.
- Rousset O, Rahmim A, Alavi A, Zaidi H. Partial volume correction strategies in PET. *PET Clin*. 2007;2:235–49.
- Rullmann M, Dukart J, Hoffmann KT, Luthardt J, Tiepolt S, Patt M, et al. Partial-Volume Effect correction improves quantitative analysis of 18F-Florbetaben beta-amyloid PET scans. *J Nucl Med*. 2016;57:198–203.
- Meltzer CC, Cantwell MN, Greer PJ, Ben-Eliezer D, Smith G, Frank G, et al. Does cerebral blood flow decline in healthy aging? A PET study with partial-volume correction. *J Nucl Med*. 2000;41:1842–8.
- Ibáñez V, Pietrini P, Alexander GE, Furey ML, Teichberg D, Rajapakse JC, et al. Regional glucose metabolic abnormalities are not the result of atrophy in Alzheimer's disease. *Neurology*. 1998;50:1585–93.
- Knowlton RC, Laxer KD, Klein G, Sawrie S, Ende G, Hawkins RA, et al. In vivo hippocampal glucose metabolism in mesial temporal lobe epilepsy. *Neurology*. 2001;57:1184–90.
- Yang J, Hu C, Guo N, Dutta J, Vaina LM, Johnson KA, et al. Partial volume correction for PET quantification and its impact on brain network in Alzheimer's disease. *Sci Rep*. 2017;7:13035.
- Rousset OG, Deep P, Kuwabara H, Evans AC, Gjedde AH, Cumming P. Effect of partial volume correction on estimates of the influx and cerebral metabolism of 6-[(18)F]fluoro-L-dopa studied with PET in normal control and Parkinson's disease subjects. *Synapse*. 2000;37:81–9.
- Meltzer CC, Leal JP, Mayberg HS, Wagner HN Jr, Frost JJ. Correction of PET data for partial volume effects in human cerebral cortex by MR imaging. *J Comput Assist Tomogr*. 1990;14:561–70.
- Müller-Gärtner HW, Links JM, Prince JL, Bryan RN, McVeigh E, Leal JP, et al. Measurement of radiotracer concentration in brain gray matter using positron emission tomography:

- MRI-based correction for partial volume effects. *J Cereb Blood Flow Metab.* 1992;12:571–83.
17. Rousset OG, Ma Y, Evans AC. Correction for partial volume effects in PET: principle and validation. *J Nucl Med.* 1998;39:904–11.
 18. Lu Y, Toyonaga T, Naganawa M, Gallezot JD, Chen MK, Mecca AP, et al. Partial volume correction analysis for (11)C-UCB-J PET studies of Alzheimer's disease. *Neuroimage.* 2021;238:118248.
 19. Onoue F, Yamamoto S, Uozumi H, Kamezaki R, Nakamura Y, Ikeda R, et al. Correction of partial volume effect using CT images in brain (18)F-FDG PET. *Nihon Hoshasen Gijyutsu Gakkai zasshi.* 2022;78:741–9.
 20. Oyama S, Hosoi A, Ibaraki M, McGinnity CJ, Matsubara K, Watanuki S, et al. Error propagation analysis of seven partial volume correction algorithms for [(18)F]THK-5351 brain PET imaging. *EJNMMI physics.* 2020;7:57.
 21. Sanaat A, Arabi H, Mainta I, Garibotto V, Zaidi H. Projection space implementation of deep learning-guided low-dose brain PET imaging improves performance over implementation in image space. *J Nucl Med.* 2020;61:1388–96.
 22. Zaidi H, El Naqa I. Quantitative molecular positron emission tomography imaging using advanced deep learning techniques. *Annu Rev Biomed Eng.* 2021;23:249–76.
 23. Bradshaw TJ, Boellaard R, Dutta J, Jha AK, Jacobs P, Li Q, et al. Nuclear medicine and artificial intelligence: best practices for algorithm development. *J Nucl Med.* 2022;63:500–10.
 24. Della Rosa PA, Cerami C, Gallivanone F, Prestia A, Caroli A, Castiglioni I, et al. A standardized [18F]-FDG-PET template for spatial normalization in statistical parametric mapping of dementia. *Neuroinformatics.* 2014;12:575–93.
 25. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D Slicer as an image computing platform for the quantitative imaging network. *Magn Reson Imaging.* 2012;30:1323–41.
 26. Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging.* 2010;29:196–205.
 27. Sanaat A, Shiri I, Ferdowsi S, Arabi H, Zaidi H. Robust-Deep: a method for increasing brain imaging datasets to improve deep learning models' performance and robustness. *J Dig Imaging.* 2022;35:469–81.
 28. Thomas BA, Cuplov V, Bousse A, Mendes A, Thielemans K, Hutton BF, et al. PETPVC: a toolbox for performing partial volume correction techniques in positron emission tomography. *Phys Med Biol.* 2016;61:7975–93.
 29. Hammers A, Allom R, Koeppe MJ, Free SL, Myers R, Lemieux L, et al. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp.* 2003;19:224–47.
 30. Zwanenburg A, Vallieres M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The Image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology.* 2020;295:328–38.
 31. Nioche C, Orlhac F, Boughdad S, Reuze S, Goya-Outi J, Robert C, et al. LIFEX: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res.* 2018;78:4786–9.
 32. Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp.* 1994;2:189–210.
 33. Matsubara K, Ibaraki M, Shidahara M, Kinoshita T. Iterative framework for image registration and partial volume correction in brain positron emission tomography. *Radiol Phys Technol.* 2020;13:348–57.
 34. Zaidi H, Ruest T, Schoenahl F, Montandon ML. Comparative assessment of statistical brain MR image segmentation algorithms and their impact on partial volume correction in PET. *Neuroimage.* 2006;32:1591–607.
 35. Gutierrez D, Montandon ML, Assal F, Allaoua M, Ratib O, Lovblad KO, et al. Anatomically guided voxel-based partial volume effect correction in brain PET: impact of MRI segmentation. *Comput Med Imaging Graph.* 2012;36:610–9.
 36. Meltzer CC, Zubietta JK, Links JM, Brakeman P, Stumpf MJ, Frost JJ. MR-based correction of brain PET measurements for heterogeneous gray matter radioactivity distribution. *J Cereb Blood Flow Metab.* 1996;16:650–8.
 37. Quarantelli M, Berkouk K, Prinster A, Landeau B, Svarer C, Balkay L, et al. Integrated software for the analysis of brain PET/SPECT studies with partial-volume-effect correction. *J Nucl Med.* 2004;45:192–201.
 38. Strul D, Bendriem B. Robustness of anatomically guided pixel-by-pixel algorithms for partial volume effect correction in positron emission tomography. *J Cereb Blood Flow Metab.* 1999;19:547–59.
 39. Cysouw MCF, Golla SVS, Frings V, Smit EF, Hoekstra OS, Kramer GM, et al. Partial-volume correction in dynamic PET-CT: effect on tumor kinetic parameter estimation and validation of simplified metrics. *EJNMMI Res.* 2019;9:12.
 40. Teo BK, Seo Y, Bacharach SL, Carrasquillo JA, Libutti SK, Shukla H, et al. Partial-volume correction in PET: validation of an iterative postreconstruction method with phantom and patient data. *J Nucl Med.* 2007;48:802–10.
 41. Mignotte M, Meunier J. Three-dimensional blind deconvolution of SPECT images. *IEEE Trans Biomed Eng.* 2000;47:274–80.
 42. Hoetjes NJ, van Velden FH, Hoekstra OS, Hoekstra CJ, Krak NC, Lammertsma AA, et al. Partial volume correction strategies for quantitative FDG PET in oncology. *Eur J Nucl Med Mol Imaging.* 2010;37:1679–87.
 43. Kuhn FP, Warnock GI, Burger C, Ledermann K, Martin-Soelch C, Buck A. Comparison of PET template-based and MRI-based image processing in the quantitative analysis of C11-raclopride PET. *EJNMMI Res.* 2014;4:7.
 44. Zhu Y, Bilgel M, Gao Y, Rousset OG, Resnick SM, Wong DF, Rahmim A. Deconvolution-based partial volume correction of PET images with parallel level set regularization. *Phys Med Biol.* 2021;66(14):145003. <https://doi.org/10.1088/1361-6560/ac0d8f>.
 45. Lehnert W, Gregoire M-C, Reilhac A, Meikle SR. Comparative study of partial volume correction methods in small animal positron emission tomography (PET) of the rat brain. In: 2011 IEEE Nuclear Science Symposium and Medical Imaging Conference Record, Valencia, Spain. 2011. pp. 3807–11. <https://doi.org/10.1109/NSSMIC.2011.6153722>.
 46. Aston JA, Cunningham VJ, Asselin MC, Hammers A, Evans AC, Gunn RN. Positron emission tomography partial volume correction: estimation and algorithms. *J Cereb Blood Flow Metab.* 2002;22:1019–34.
 47. Shidahara M, Thomas BA, Okamura N, Ibaraki M, Matsubara K, Oyama S, et al. A comparison of five partial volume correction methods for Tau and Amyloid PET imaging with [(18)F]THK5351 and [(11)C]PIB. *Ann Nucl Med.* 2017;31:563–9.
 48. Du Y, Tsui BM, Frey EC. Partial volume effect compensation for quantitative brain SPECT imaging. *IEEE Trans Med Imaging.* 2005;24:969–76.
 49. Gao Y, Zhu Y, Bilgel M, Ashrafinia S, Lu L, Rahmim A. Voxel-based partial volume correction of PET images via subtle MRI guided non-local means regularization. *Phys Med.* 2021;89:129–39.
 50. Caldwell CB, Mah K, Ung YC, Danjoux CE, Balogh JM, Ganguli SN, et al. Observer variation in contouring gross tumor volume in

- patients with poorly defined non-small-cell lung tumors on CT: the impact of 18FDG-hybrid PET fusion. *Int J Radiat Oncol Biol Phys.* 2001;51:923–31.
51. Steenbakkers RJ, Duppen JC, Fitton I, Deurloo KE, Zijp L, Uitterhoeve AL, et al. Observer variation in target volume delineation of lung cancer related to radiation oncologist-computer interaction: a “Big Brother” evaluation. *Radiother Oncol.* 2005;77:182–90.
 52. Lavelly WC, Scarfone C, Cevikalp H, Li R, Byrne DW, Cmelak AJ, et al. Phantom validation of coregistration of PET and CT for image-guided radiotherapy. *Med Phys.* 2004;31:1083–92.
 53. Nömayr A, Römer W, Hothorn T, Pfahlberg A, Hornegger J, Bautz W, et al. Anatomical accuracy of lesion localization. Retrospective interactive rigid image registration between 18F-FDG-PET and X-ray CT. *Nuklearmedizin.* 2005;44:149–55.
 54. Ibaraki M, Matsubara K, Shinohara Y, Shidahara M, Sato K, Yamamoto H, et al. Brain partial volume correction with point spreading function reconstruction in high-resolution digital PET: comparison with an MR-based method in FDG imaging. *Ann Nucl Med.* 2022;36:717–27.
 55. Tohka J, Reilhac A. Deconvolution-based partial volume correction in Raclopride-PET and Monte Carlo comparison to MR-based method. *Neuroimage.* 2008;39:1570–84.
 56. Wang Y, Zhou L, Yu B, Wang L, Zu C, Lalush DS, et al. 3D Auto-context-based locality adaptive multi-modality GANs for PET synthesis. *IEEE Trans Med Imaging.* 2019;38:1328–39.
 57. Xiang L, Qiao Y, Nie D, An L, Wang Q, Shen D. Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI. *Neurocomputing.* 2017;267:406–16.
 58. Kuang G, Jiahui G, Kyungsang K, Xuezhu Z, Jaewon Y, Youngho S, et al. Iterative PET image reconstruction using convolutional neural network representation. *IEEE Trans Med Imaging.* 2019;38:675–85.
 59. Dong X, Lei Y, Wang T, Higgins K, Liu T, Curran WJ, et al. Deep learning-based attenuation correction in the absence of structural information for whole-body positron emission tomography imaging. *Phys Med Biol.* 2020;65:055011.
 60. Dal Toso L, Chalampalakakis Z, Buvat I, Comtat C, Cook G, Goh V, et al. Improved 3D tumour definition and quantification of uptake in simulated lung tumours using deep learning. *Phys Med Biol.* 2022;67:095013. <https://doi.org/10.1088/1361-6560/ac65d6>.
 61. Zhao Q, Liu M, Ha L, Zhou Y. Alzheimer’s Disease Neuroimaging I. Quantitative (18)F-AV1451 brain Tau PET imaging in cognitively normal older adults, mild cognitive impairment, and Alzheimer’s disease patients. *Front Neurol.* 2019;10:486.
 62. Goffin K, Van Paesschen W, Dupont P, Baete K, Palmini A, Nuyts J, et al. Anatomy-based reconstruction of FDG-PET images with implicit partial volume correction improves detection of hypometabolic regions in patients with epilepsy due to focal cortical dysplasia diagnosed on MRI. *Eur J Nucl Med Mol Imaging.* 2010;37:1148–55.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Amirhossein Sanaat¹ · Hossein Shooli² · Andrew Stephen Böhringer¹ · Maryam Sadeghi³ · Isaac Shiri¹ · Yazdan Salimi¹ · Nathalie Ginovart^{4,5,6} · Valentina Garibotto^{1,4} · Hossein Arabi¹ · Habib Zaidi^{1,4,7,8} 

¹ Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211 Geneva, Switzerland

² Persian Gulf Nuclear Medicine Research Center, Department of Molecular Imaging and Radionuclide Therapy (MIRT), Bushehr Medical University Hospital, Faculty of Medicine, Bushehr University of Medical Sciences, Bushehr, Iran

³ Department of Medical Statistics, Informatics and Health Economics, Medical University of Innsbruck, Schoepfstr. 41, Innsbruck, Austria

⁴ Geneva University Neurocenter, University of Geneva, Geneva, Switzerland

⁵ Department of Psychiatry, Geneva University, Geneva, Switzerland

⁶ Department of Basic Neuroscience, Geneva University, Geneva, Switzerland

⁷ Department of Nuclear Medicine and Molecular Imaging, University of Groningen, Groningen, Netherlands

⁸ Department of Nuclear Medicine, University of Southern Denmark, Odense, Denmark