

---

---

# Dixon-VIBE Deep Learning (DIVIDE) Pseudo-CT Synthesis for Pelvis PET/MR Attenuation Correction

Angel Torrado-Carvajal<sup>1</sup>, Javier Vera-Olmos<sup>2</sup>, David Izquierdo-Garcia<sup>1</sup>, Onofrio A. Catalano<sup>1</sup>, Manuel A. Morales<sup>1,3</sup>, Justin Margolin<sup>1,4</sup>, Andrea Soricelli<sup>5,6</sup>, Marco Salvatore<sup>5</sup>, Norberto Malpica<sup>2</sup>, and Ciprian Catana<sup>1</sup>

<sup>1</sup>Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Charlestown, Massachusetts; <sup>2</sup>Medical Image Analysis and Biometry Lab, Universidad Rey Juan Carlos, Madrid, Spain; <sup>3</sup>Department of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts; <sup>4</sup>Northeastern University, Boston, Massachusetts; <sup>5</sup>SDN—Istituto di Ricerca Diagnostica e Nucleare, IRCCS, Naples, Italy; and <sup>6</sup>Department of Motor Sciences and Wellness, University of Naples “Parthenope,” Naples, Italy

Whole-body attenuation correction (AC) is still challenging in combined PET/MR scanners. We describe Dixon-VIBE Deep Learning (DIVIDE), a deep-learning network that allows synthesizing pelvis pseudo-CT maps based only on the standard Dixon volumetric interpolated breath-hold examination (Dixon-VIBE) images currently acquired for AC in some commercial scanners. **Methods:** We propose a network that maps between the four 2-dimensional (2D) Dixon MR images (water, fat, in-phase, and out-of-phase) and their corresponding 2D CT image. In contrast to previous methods, we used transposed convolutions to learn the up-sampling parameters, we used whole 2D slices to provide context information, and we pretrained the network with brain images. Twenty-eight datasets obtained from 19 patients who underwent PET/CT and PET/MR examinations were used to evaluate the proposed method. We assessed the accuracy of the  $\mu$ -maps and reconstructed PET images by performing voxel- and region-based analysis comparing the SUVs (in g/mL) obtained after AC using the Dixon-VIBE (PET<sub>Dixon</sub>), DIVIDE (PET<sub>DIVIDE</sub>), and CT-based (PET<sub>CT</sub>) methods. Additionally, the bias in quantification was estimated in synthetic lesions defined in the prostate, rectum, pelvis, and spine. **Results:** Absolute mean relative change values relative to CT AC were lower than 2% on average for the DIVIDE method in every region of interest except for bone tissue, where it was lower than 4% and 6.75 times smaller than the relative change of the Dixon method. There was an excellent voxel-by-voxel correlation between PET<sub>CT</sub> and PET<sub>DIVIDE</sub> ( $R^2 = 0.9998$ ,  $P < 0.01$ ). The Bland-Altman plot between PET<sub>CT</sub> and PET<sub>DIVIDE</sub> showed that the average of the differences and the variability were lower (mean PET<sub>CT</sub> - PET<sub>DIVIDE</sub> SUV, 0.0003; PET<sub>CT</sub> - PET<sub>DIVIDE</sub> SD, 0.0094; 95% confidence interval, [-0.0180, 0.0188]) than the average of differences between PET<sub>CT</sub> and PET<sub>Dixon</sub> (mean PET<sub>CT</sub> - PET<sub>Dixon</sub> SUV, 0.0006; PET<sub>CT</sub> - PET<sub>Dixon</sub> SD, 0.0264; 95% confidence interval, [-0.0510, 0.0524]). Statistically significant changes in PET data quantification were observed between the 2 methods in the synthetic lesions, with the largest improvement in femur and spine lesions. **Conclusion:** The DIVIDE method can accurately synthesize a pelvis pseudo-CT scan from standard Dixon-VIBE images, allowing for accurate AC in combined PET/MR scanners. Additionally, our implementation allows rapid pseudo-CT synthesis, making it suitable for routine applications and even allowing retrospective processing of Dixon-VIBE data.

**Key Words:** pseudo-CT; image synthesis; PET/MR; attenuation correction; deep learning

**J Nucl Med 2019; 60:429–435**  
DOI: 10.2967/jnumed.118.209288

**T**he evolution of PET and MRI into complementary in vivo molecular imaging techniques has generated increased interest in the development of combined PET/MRI systems (1). However, many challenges have slowed the adoption of the recently developed simultaneous PET/MR scanners (2). The most important methodologic challenge for whole-body applications is attenuation correction (AC), which is critical for quantitative studies and even qualitative data interpretation.

In the case of the commercially available PET/MR scanners, attenuation maps ( $\mu$ -maps) are typically generated from the MR data acquired using a Dixon volumetric interpolated breath-hold examination (Dixon-VIBE) sequence or a liver acquisition with volume acceleration-flexible (LAVA-Flex) sequence. Four tissue classes (background, fat, lung, and soft tissue) are segmented from these data, and discrete linear attenuation coefficients are assigned to each class (3). The major limitation of this method is that it misclassifies bone as fat tissue, leading to substantial bias in the PET measurements (4), particularly in the pelvis.

Most of the early efforts in MR-based AC have focused on the head (5), differing mainly in the type of semantic representation used to describe the image data (i.e., templates or atlases (6,7), segmentation (8,9), or reconstruction (10)). It has recently been shown that several of the template- or atlas-based methods are accurate (11). However, their translation to whole-body applications is still challenging. For example, the large intersubject variability (in terms of anatomy, size, body mass index, pathology) makes the generation of atlases and templates that adequately describe the population difficult. Furthermore, using these methods in other body regions (e.g., pelvis) requires a combination of rigid and nonrigid registration. Consequently, these methods showed limited accuracy when applied to whole-body AC (12). In this context, Leynes et al. presented a hybrid zero-echo-time Dixon MR-based AC method that included bone density estimates obtained by converting the zero-echo-time signal intensity to Hounsfield units using a continuous 2-segment piecewise linear

---

Received Feb. 6, 2018; revision accepted Aug. 23, 2018.  
For correspondence or reprints contact: Ciprian Catana, A.A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, 149 13th St., Charlestown, MA 02129.  
E-mail: ccatana@mgh.harvard.edu  
Published online Aug. 30, 2018.  
COPYRIGHT © 2019 by the Society of Nuclear Medicine and Molecular Imaging.

model (13). However, this method relies on a manual correction step during bone segmentation.

As an alternative, deep-learning approaches can be used to generate  $\mu$ -maps from the MR data. Deep-learning approaches based on convolutional neural networks have already shown tremendous potential for image classification (14), image segmentation (15–17), and image synthesis (18,19), becoming popular in medical imaging (20–24).

Several deep-learning–based approaches have already been proposed for MR-based AC. Ribeiro et al. proposed a feed-forward neural network to directly output a continuous-valued head  $\mu$ -map by non-linear regression of several ultrashort-echo-time images and the template-based AC map (25). Very recently, Leynes et al. also proposed a multiparametric MRI model to generate pelvis pseudo-CT maps based on Dixon MRI and proton-density–weighted zero-echo-time MR images (26). Although the preliminary results obtained with these methods are encouraging, they have the disadvantage of relying on specialized MR sequences that require additional acquisition times.

In this paper, we describe a Dixon-VIBE Deep Learning (DIVIDE) method to synthesize pelvis pseudo-CT maps based only on the standard Dixon-VIBE images routinely acquired for AC purposes on commercial scanners. Our network maps between the 2-dimensional (2D) Dixon MRI slices and their corresponding CT slices. The network architecture is similar to one previously proposed (15,16,21), but instead of using unpooling to up-sample the features, we use transposed convolutions to learn the up-sampling parameters. In contrast to previously proposed patch-based methods, we used the whole 2D slice to provide context information. We showed the PET quantification bias to be reduced when using DIVIDE compared with the standard Dixon-VIBE approach.

## MATERIALS AND METHODS

### Data Acquisition

CT and PET/MR images of 13 colorectal and 6 prostate cancer patients (mean age,  $61.42 \pm 10.63$  y; range, 44–80 y; 12 men and 7 women; mean body mass index,  $22.30 \pm 2.88$ ; range, 17.43–29.73) were retrospectively analyzed. Additionally, follow-up images for 9 of the colorectal cancer patients were also included (substantial anatomy and body mass index changes were noted after treatment/surgery). All patients underwent same-day PET/CT and PET/MR studies following the routine protocols. No obvious artifacts were observed in the CT or MR images. All patients gave written informed consent, and the local Institutional Review Board approved the study.

**CT Data.** Low-dose CT images were acquired on a Discovery PET/CT 710 scanner (GE Healthcare) with a matrix of  $512 \times 512$ , resolution of

$1.37 \times 1.37$  mm, slice thickness of 3.75 mm, pitch of 0.94 mm, acquisition angle of 0, voltage of 120 kV, and tube current of 150 mA.

**PET/MR Data.** The PET/MR examinations were performed 66.60  $\pm$  22.83 min after the PET/CT scans using the Biograph mMR scanner (Siemens Healthineers). PET emission data were acquired for  $16.1 \pm 10.0$  min (range, 7–35 min) and a variable number of bed positions (1–5)  $120.0 \pm 36.8$  min after administration of  $378.1 \pm 64.1$  MBq of  $^{18}\text{F}$ -FDG (16 subjects, 25 scans) or  $291.0 \pm 50.5$  MBq of  $^{18}\text{F}$ -choline (3 subjects, 3 scans; Supplemental Table 1, available at <http://jnm.snmjournals.org>).

The standard Dixon-VIBE sequence (repetition time, 3.6 ms; first echo time, 1.23 ms; second echo time, 2.46 ms; acquisition matrix,  $192 \times 128$ ; 128 slices per bed position; voxel size,  $2.6 \times 2.6 \times 3.1$  mm [in-plane resolution  $\times$  slice thickness]; flip angle,  $10^\circ$ ; acquisition time, 18 s) was run for AC purposes. Other MR sequences were additionally run as part of the clinical protocols for prostate and colorectal cancer, respectively.

### Data Preprocessing

Images were preprocessed as follows. First, MRI bias was corrected using the N4ITK MRI bias correction module in 3D Slicer 4 (27). Given the large field of view, we set the advanced parameters to 3 levels of resolution with 500, 400, and 300 iterations, respectively, and a convergence threshold of 0.00001 to ensure proper modeling of the inhomogeneity bias. Then, intrasubject rigid and nonrigid registration of the MR and CT data was performed using SPM8 software (28) and the Elastix package (29). Finally, reslicing and cropping of the data to a fixed  $50 \times 50$  cm field of view with  $2 \times 2$  mm pixels and a 1-mm slice thickness was performed to ensure spatial homogeneity among subjects.

### DIVIDE Pseudo-CT Synthesis

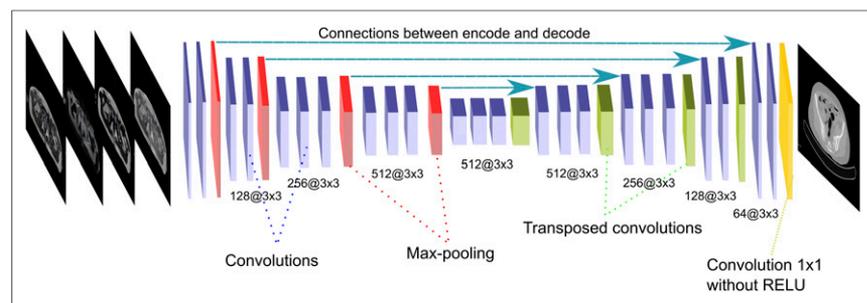
Figure 1 provides a schematic representation of the convolutional-deconvolutional neural network implemented in TensorFlow (30) that takes four 2D images as input—water, fat, in-phase, and out-of-phase Dixon-VIBE images—and computes the corresponding pseudo-CT slice.

The encoder network consists of 13 convolutional layers, corresponding to the first 13 convolutions of the Visual Geometry Group network (31), that convert the input MR images into a multidimensional feature representation. This is followed by a decoder network, which performs several up-samplings of the features to generate the final pseudo-CT slice. Shortcuts between the layers of the encoder and the decoder were created to use features from different scales in the decoding step.

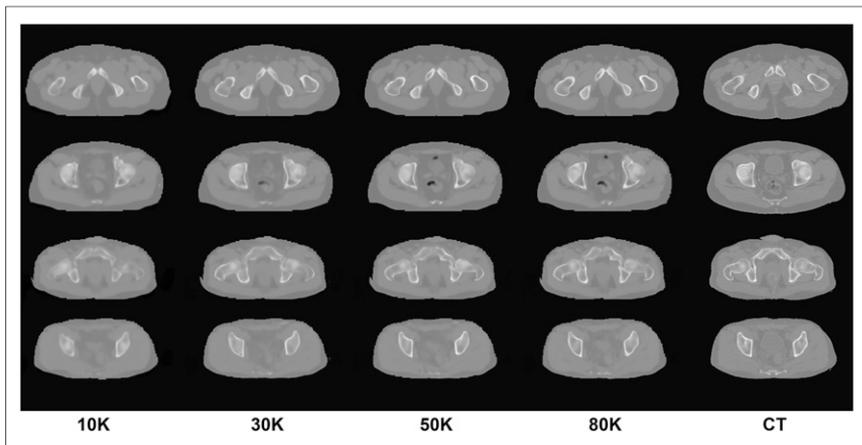
Each layer of the encoder performed several convolutions with a  $3 \times 3$  filter bank to produce a set of feature maps. After every convolution, a batch normalization was performed (32) and an elementwise rectified linear unit was applied. At the end of each encoding layer, a max-pooling with a  $2 \times 2$  window and stride 2 was performed, resulting in a subsampling by a factor of 2 (33). After every max-pooling, the number of filters was doubled. This step introduced larger context information in the resulting feature map. Increasing this context information resulted in a loss of spatial information, which is the reason for adding the shortcuts between the layers of the encoder and the decoder.

The decoder network comprised 10 convolutions and 4 transposed convolutions. Transposed convolutions (33) with a kernel of  $3 \times 3$  and stride 2 were used to up-sample the feature map by a factor of 2. After every convolution, a batch normalization was performed and a rectified linear unit was applied. A final  $1 \times 1$  convolution was performed to obtain the Hounsfield units.

The mean absolute error between CT and pseudo-CT slices was used as loss function.



**FIGURE 1.** Schematic representation of our network architecture. Our implementation in TensorFlow processes in-phase and out-of-phase fat and water 2D Dixon slices to generate pseudo-CT slice. RELU = rectified linear unit.



**FIGURE 2.** Examples of progression of pseudo-CT generation during training. Columns represent output of network after a certain number of training iterations. Corresponding real CT used as ground truth is shown in rightmost column.

The mean absolute error was calculated only for the voxels within the body using a mask generated from the Dixon images.

To avoid overfitting, we performed data augmentation by applying random displacements of 5 voxels and a random flip in the slices, similarly to Han (21). Additionally, we used a dropout layer (32) with

linear transformation (34). The corresponding  $\mu$ -maps were denoted  $\mu_{\text{DIVIDE}}$  and  $\mu_{\text{CT}}$ . Gaussian smoothing with a kernel of 4 mm in full width at half maximum was applied to match the PET spatial resolution. Additionally, the Dixon-VIBE-based  $\mu$ -maps ( $\mu_{\text{Dixon}}$ ) (9) were available. Air pockets in all the resulting  $\mu$ -maps ( $\mu_{\text{CT}}$ ,  $\mu_{\text{Dixon}}$ , and  $\mu_{\text{DIVIDE}}$ ) were filled as soft tissue to avoid any bias that might result from differences caused by their displacement between the acquisitions.

### $\mu$ -Map Generation and PET Image Reconstruction

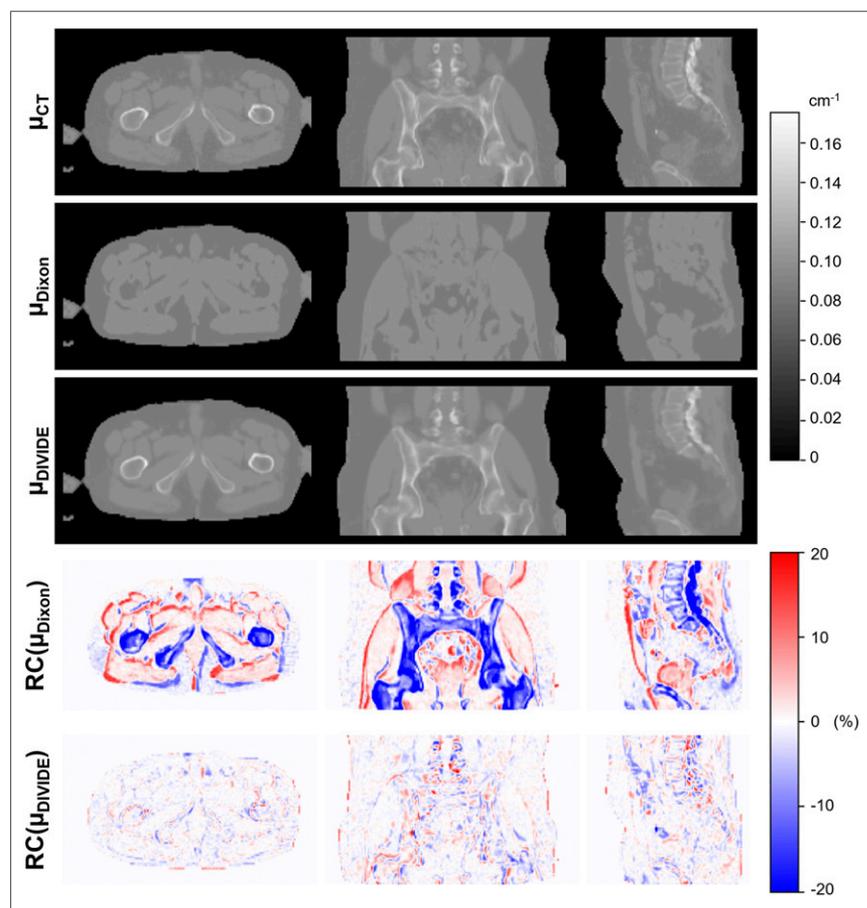
The pseudo-CT images were converted to linear attenuation coefficients using the bilinear transformation (34). The corresponding  $\mu$ -maps were denoted  $\mu_{\text{DIVIDE}}$  and  $\mu_{\text{CT}}$ . Gaussian smoothing with a kernel of 4 mm in full width at half maximum was applied to match the PET spatial resolution. Additionally, the Dixon-VIBE-based  $\mu$ -maps ( $\mu_{\text{Dixon}}$ ) (9) were available. Air pockets in all the resulting  $\mu$ -maps ( $\mu_{\text{CT}}$ ,  $\mu_{\text{Dixon}}$ , and  $\mu_{\text{DIVIDE}}$ ) were filled as soft tissue to avoid any bias that might result from differences caused by their displacement between the acquisitions.

The PET volumes were reconstructed using the standard 3-dimensional ordinary Poisson ordered-subset expectation-maximization algorithm with 21 subsets and 3 iterations, correcting for normalization, random prompts, background, dead time, photon attenuation, and scatter using the *e7tools* software provided by the manufacturer. Spatial smoothing was performed after image reconstruction using a gaussian filter of 4 mm in full width at half maximum. The final volumes were reconstructed into a matrix of  $344 \times 344$  voxels of  $2.0863 \times 2.0863$  mm in plane and a 2.0313-mm slice thickness and were denoted  $\text{PET}_{\text{DIVIDE}}$ ,  $\text{PET}_{\text{CT}}$ , and  $\text{PET}_{\text{Dixon}}$ .

A k-fold cross-validation ( $k = 4$ ) was performed to evaluate the proposed method using 28 Dixon and CT volume pairs from 19 subjects (using the subjects instead of the pairs, meaning the training set contained 15 subjects and the remaining 4 were part of the test set in each iteration). All pairs from the same subject were always in the same collection.

### Data Analysis

Voxel- and region-based analyses were performed to assess the accuracy of the  $\mu$ -maps and reconstructed PET images. Only voxels included in the pelvic area where the different  $\mu$ -maps overlapped were used for comparison. Bland-Altman plots, bias, and variability Pearson correlation coefficients were calculated for all comparisons. Absolute relative changes (RCs) were defined as follows:



**FIGURE 3.** Sagittal, coronal, and axial  $\mu$ -maps of subject 02\_01 (Supplemental Table 1) obtained using 3 different methods— $\mu_{\text{CT}}$ ,  $\mu_{\text{Dixon}}$ , and  $\mu_{\text{DIVIDE}}$ —as well as distribution of RC of  $\mu_{\text{Dixon}}$  and  $\mu_{\text{DIVIDE}}$  compared with  $\mu_{\text{CT}}$ .

**TABLE 1**

Voxel- and ROI-Based RCs for  $\mu$ -Maps in Whole-Pelvis Area

RC	$\mu_{\text{DIVIDE}}$	$\mu_{\text{Dixon}}$
<b>Voxel-based</b>		
Absolute	2.36 $\pm$ 3.15	4.73 $\pm$ 6.31
Nonabsolute	0.04 $\pm$ 3.94	-0.03 $\pm$ 7.89
<b>ROI-based</b>		
<b>Fat</b>		
Absolute	2.19 $\pm$ 2.71	4.05 $\pm$ 5.03
Nonabsolute	0.88 $\pm$ 3.36	1.90 $\pm$ 6.17
<b>Soft-tissue</b>		
Absolute	2.25 $\pm$ 2.96	4.18 $\pm$ 4.76
Nonabsolute	-0.39 $\pm$ 3.70	-0.01 $\pm$ 6.36
<b>Bone</b>		
Absolute	7.08 $\pm$ 6.07	24.05 $\pm$ 8.66
Nonabsolute	-4.14 $\pm$ 8.32	-24.05 $\pm$ 8.66

Data are mean percentage  $\pm$  SD for all subjects.

$$RC I_X(\%) = 100 \frac{I_X - I_{GS}}{I_{GS}}, \quad \text{Eq. 1}$$

where  $X$  is the AC method (DIVIDE or Dixon);  $I_X$  corresponds to either the  $\mu$ -maps ( $\mu_{\text{DIVIDE}}$  or  $\mu_{\text{Dixon}}$ ) or the PET images ( $\text{PET}_{\text{DIVIDE}}$  or  $\text{PET}_{\text{Dixon}}$ ); and  $I_{GS}$  corresponds to the gold standard image ( $\mu_{\text{CT}}$  or  $\text{PET}_{\text{CT}}$  for  $\mu$ -map and PET comparisons, respectively). Nonabsolute RCs were also calculated.

Regions of interest (ROIs) were defined as segmented tissues (fat, soft tissue, and bone) obtained by thresholding the Hounsfield units from the patient-specific CT image. Mean and maximum SUV normalized by weight was calculated for all the patients; all SUVs reported in this article are in g/mL.

Additionally, we simulated synthetic lesions using the PETSTEP approach (35). Briefly, 1-cm-diameter spheric lesions were defined in the prostate, rectum, pelvis, and spine. Using the original reconstructed PET images, new lesion images were generated by increasing the activity within the lesion 5-fold and setting the background to zero. Lesion sinograms were generated from these images as described by Berthon et al. (35), except that random and scatter events due to the lesion were neglected. The lesion and preexisting patient sinograms were combined and reconstructed as previously described.

**Statistical Tests**

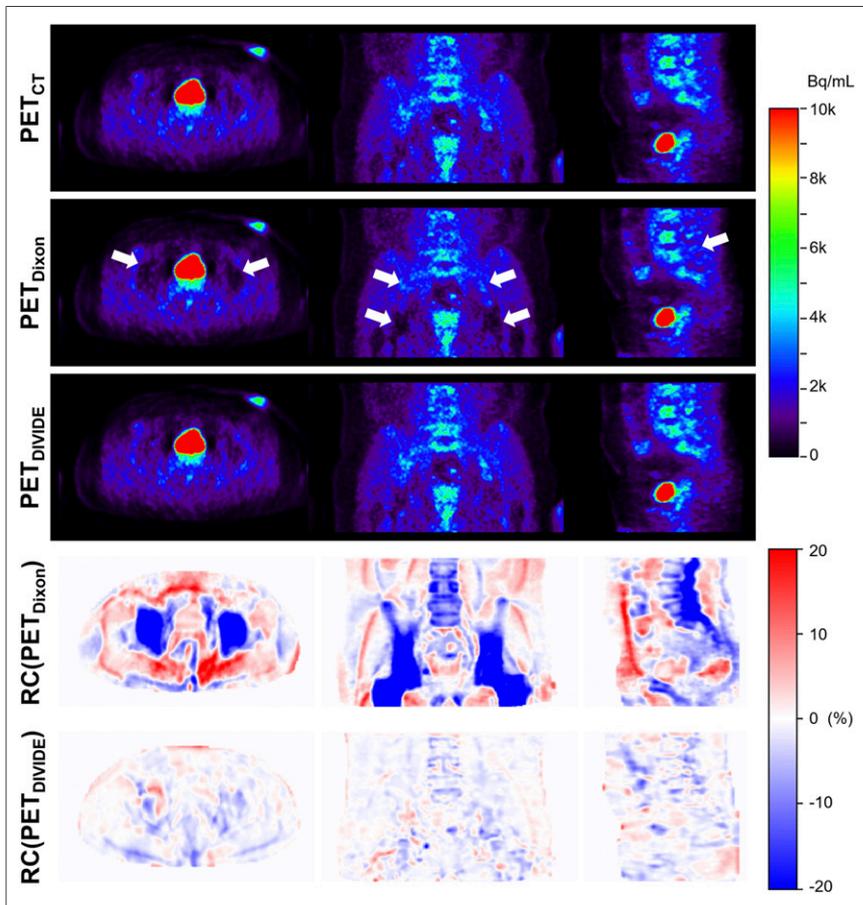
Mean  $\pm$  SD for  $\mu$ -maps and reconstructed PET images were computed for all subjects. Paired-sample Wilcoxon signed rank-tests were performed to assess whether there were differences when correcting for attenuation using the CT-based ( $\text{PET}_{\text{CT}}$ ) and the MR-based ( $\text{PET}_{\text{DIVIDE}}$  and  $\text{PET}_{\text{Dixon}}$ ) methods. The same tests were performed for the analysis of the synthetic lesions. Statistical significance was considered when the  $P$  value was lower than 0.05.

**RESULTS**

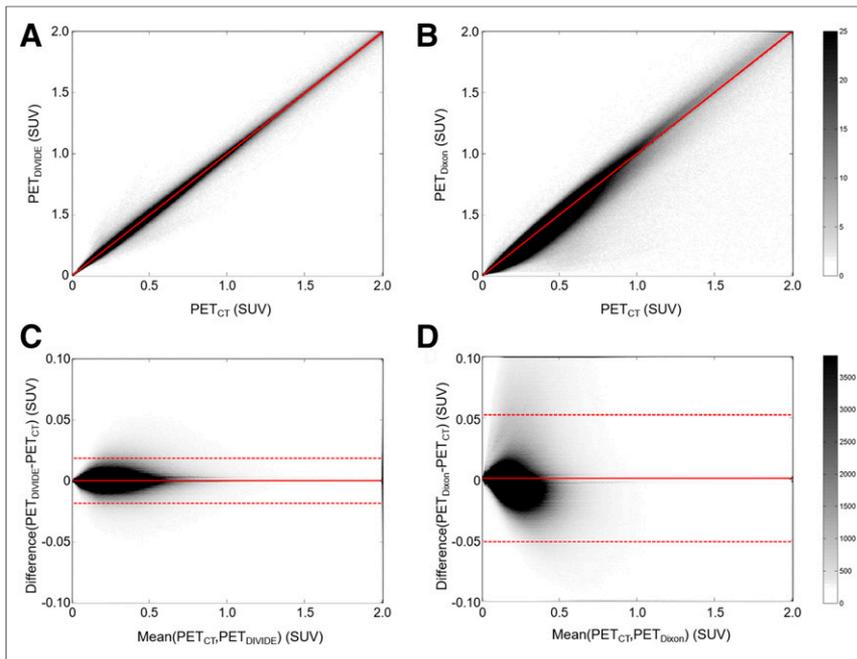
The Hounsfield unit error was lower when using the pretrained weights from the brain database as opposed to initializing them from noise, showing a faster learning in the initial epochs and a slightly smaller final error (Supplemental Fig. 1). The number of training iterations also had an effect on the progressive pseudo-CT improvement (e.g., better bone delineation) at different stages (Fig. 2). Once the network is trained, the synthesis of a complete volume takes about 2 min using a NVIDIA Tesla K40 GPU.

Figure 3 shows the  $\mu$ -maps obtained using the 3 methods and the corresponding RC maps for a representative subject, demonstrating excellent correlation between  $\mu_{\text{DIVIDE}}$  and  $\mu_{\text{CT}}$  and noticeable misclassification of bone tissues in  $\mu_{\text{Dixon}}$ . The bone contours were accurately visualized in  $\mu_{\text{DIVIDE}}$ , and the shape of the pelvis was estimated well despite the anatomic complexity and large interpatient variability.

The results of the quantitative voxel- and ROI-based analyses for all subjects are summarized in Table 1. Absolute mean RCs were lower for DIVIDE than for the Dixon-VIBE-based method in every ROI.



**FIGURE 4.** Sagittal, coronal, and axial PET images of subject 02\_01 (Supplemental Table 1) obtained using 3 different AC methods— $\text{PET}_{\text{CT}}$ ,  $\text{PET}_{\text{Dixon}}$  and  $\text{PET}_{\text{DIVIDE}}$ —as well as distribution of RC of  $\text{PET}_{\text{Dixon}}$  and  $\text{PET}_{\text{DIVIDE}}$  compared with  $\text{PET}_{\text{CT}}$ . Arrows point to areas where  $\text{PET}_{\text{Dixon}}$  exhibits noticeable PET signal underestimation.



**FIGURE 5.** Voxel-by-voxel correlation plots between PET<sub>CT</sub> and PET<sub>DIVIDE</sub> (A) and between PET<sub>CT</sub> and PET<sub>Dixon</sub> (B), as well as Bland–Altman plots between PET<sub>CT</sub> and PET<sub>DIVIDE</sub> (C) and between PET<sub>CT</sub> and PET<sub>Dixon</sub> (D), for all subjects. Gray scale bars show number of voxels within each element of 2D histogram. SUV units are g/mL.

Particularly, bone tissue demonstrated a statistically significant 3.4 times reduction of bias for DIVIDE. Nonabsolute values showed a similar trend. The SDs were almost 2 times smaller for DIVIDE than for the Dixon-VIBE-based method.

The PET<sub>DIVIDE</sub> images more closely resembled the PET<sub>CT</sub> images than the PET<sub>Dixon</sub> images (Fig. 4). Some deviations could still be appreciated in the PET<sub>DIVIDE</sub> images, with the highest errors in

**TABLE 2**

Voxel- and ROI-Based RCs for PET Data in Whole-Pelvis Area

RC	PET <sub>DIVIDE</sub>	PET <sub>Dixon</sub>	<i>P</i> <sup>*</sup>
<b>Voxel-based</b>			
Absolute	1.83 ± 2.36	6.33 ± 7.88	<0.01
Nonabsolute	0.06 ± 2.93	-0.41 ± 10.06	0.1866
<b>ROI-based</b>			
<b>Fat</b>			
Absolute	1.70 ± 2.03	4.49 ± 5.05	<0.01
Nonabsolute	0.27 ± 2.59	1.48 ± 6.51	<0.01
<b>Soft-tissue</b>			
Absolute	1.83 ± 2.42	6.71 ± 7.48	<0.01
Nonabsolute	-0.03 ± 2.98	-0.34 ± 10.00	0.6653
<b>Bone</b>			
Absolute	3.75 ± 3.90	25.34 ± 12.51	<0.01
Nonabsolute	-0.95 ± 5.09	-25.11 ± 12.71	<0.01

\*Paired-sample Wilcoxon signed-rank test.

Data are mean percentage ± SD for all subjects.

the pelvis being near bone boundaries. However, the errors were lower than 2% on average.

There was an excellent correlation between PET<sub>CT</sub> and PET<sub>DIVIDE</sub> ( $R^2 = 0.9998$ ,  $P < 0.01$ ; Fig. 5). The Bland–Altman plot between PET<sub>CT</sub> and PET<sub>DIVIDE</sub> showed that the mean differences and variability were lower (mean PET<sub>CT</sub>–PET<sub>DIVIDE</sub> SUV, 0.0003; PET<sub>CT</sub>–PET<sub>DIVIDE</sub> SD, 0.0094; 95% confidence interval, [–0.0180,0.0188]) than the mean differences and variability between PET<sub>CT</sub> and PET<sub>Dixon</sub> (mean PET<sub>CT</sub>–PET<sub>Dixon</sub> SUV, 0.0006; PET<sub>CT</sub>–PET<sub>Dixon</sub> SD, 0.0264; 95% confidence interval, [–0.0510,0.0524]); the difference between the proposed method and the ground truth tended to decrease as the average increased. The comparison of the Bland–Altman plots for the Dixon-VIBE-based and DIVIDE approaches showed that the proposed method is less biased.

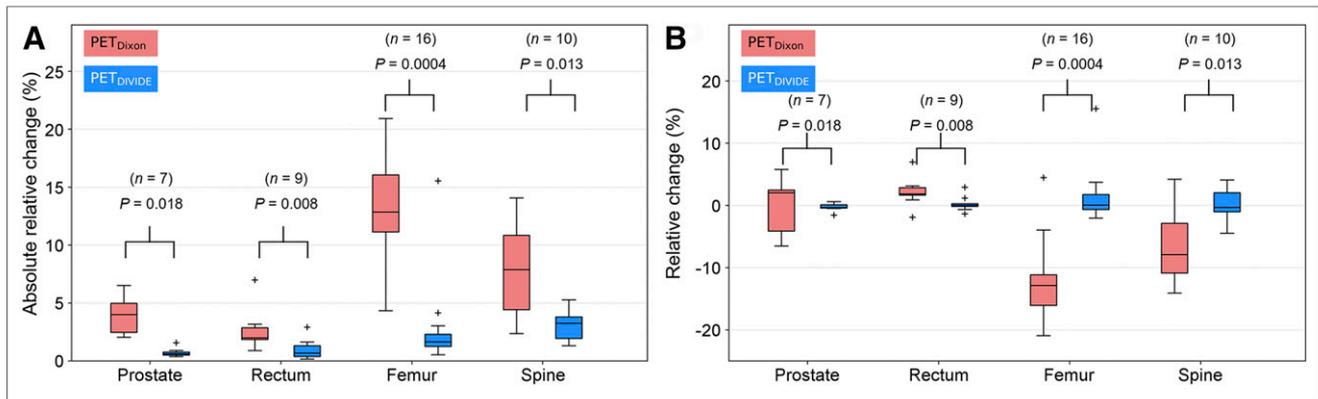
ROI analysis showed that absolute mean RCs were lower than 2% for the DIVIDE method in every nonbone ROI (Table 2). A 6.75-times bias reduction in bone tissue was observed for DIVIDE compared with the Dixon-VIBE-based method, a difference that was statistically significant. Again, the nonabsolute values showed a similar trend, and a factor of 3.5 reduction in variability was observed in all cases considered, suggesting that the method is more precise.

Figure 6 shows the results of the quantitative ROI-based analyses for the synthetic lesions. Statistically significant differences between the Dixon-VIBE-based and DIVIDE methods were observed when comparing the absolute mean RCs (Fig. 6A). Prostate lesions demonstrated a 6.9-times bias reduction for DIVIDE compared with the Dixon-VIBE-based method, whereas rectum lesions showed a 2.9-times bias reduction. As expected, the largest reduction in bias was observed for DIVIDE in the case of bone lesions. Femur and spine lesions revealed a 7.8- and a 2.4-times bias reduction, respectively. The improved accuracy compared with the Dixon-VIBE-based method was also evident when comparing the nonabsolute mean RCs (Fig. 6B).

## DISCUSSION

We have developed DIVIDE, a pelvis deep-learning pseudo-CT synthesis method, as a first step in a “DIVIDE and conquer” approach that could eventually prove successful in addressing the whole-body PET/MR AC challenge. DIVIDE avoids the oversimplification of most MR segmentation-based approaches and the limitations of atlas-based techniques. The qualitative (Figs. 4 and 5) and quantitative (Tables 1 and 2) analyses provided results similar to those obtained using the current standard (CT). The low RCs indicated that DIVIDE allows estimation of the patient-specific CT substantially better than the Dixon-VIBE-based one.

Some of the remaining differences observed at the bone boundaries could be due to inaccuracies in the MR–CT intrasubject registration, which is complicated by several factors: differences in patient positioning between the CT and MR scans, the



**FIGURE 6.** Box-and-whisker plots of impact of Dixon-VIBE-based and DIVIDE-based AC on PET data quantification in synthetic lesions: absolute (A) and nonabsolute (B) mean RC. For each box, central mark is median, edges of box are 25th and 75th percentiles, whiskers extend to most extreme data points ( $1.5 \times$  interquartile range), and outliers are plotted individually.

complex anatomy of the pelvis, nonrigid displacement of the organs between acquisitions, and high freedom of movement of the bone structures. Additionally, there were mismatches between the MR and CT air-pocket locations; however, air pockets generated in our approach showed an acceptable correlation with the anatomy in the MR image, which is sometimes impaired by the information present in the CT image (Supplemental Fig. 2). Bland-Altman plots for the PET images also suggest that DIVIDE is closer to the CT image than the Dixon-VIBE-based method. The narrow 95% confidence interval for DIVIDE suggests it is more accurate than the Dixon-VIBE-based approach.

Several other deep-learning methods have previously been described for head (21,36) and, very recently, pelvis AC (26). Our findings agree with those reported by Leynes et al. (26). Their method (PET<sub>ZeDD</sub> [zero-echo-time and Dixon deep pseudo-CT]) largely corrected underestimation of uptake within and around bony regions compared with the Dixon-based method (PET<sub>Dixon</sub>), in line with the improvement we showed in this paper. Also, they reported a reduction of root-mean-square error from 6.10% for PET<sub>Dixon</sub> to 2.85% for PET<sub>ZeDD</sub> within soft-tissue regions, whereas we report a change in RC from 6.71% for PET<sub>Dixon</sub> to 1.83% for PET<sub>DIVIDE</sub> in the same region. The underestimation bias was reduced by a factor of 3 and the SD was reduced by a factor of 2 in the PET<sub>ZeDD</sub>; we reduced that bias by a factor of 11.3 and the SD by a factor of 3.4 in the PET<sub>DIVIDE</sub>.

Compared with the ZeDD-CT approach, the DIVIDE method presents several advantages: relying on only the standard Dixon-VIBE images currently used for AC, decreasing the scanning time, and allowing for retrospective generation of  $\mu$ -maps from already acquired Dixon-VIBE data. Moreover, DIVIDE is completely automated and uses the whole image (as opposed to patches) to reduce potential boundary artifacts. On the other hand, the ZeDD-CT approach (26) might be more accurate because it can extract more information about air pockets from the zero-echo-time image.

Our study had several limitations. First, our training set was relatively small and comprised only subjects above 44 y old. Further improvement of the current model could be achieved by increasing the number of training sets and fine-tuning the bone and weight parameters or training the network with the new higher-resolution CAIPIRINHA-3-dimensional-Dixon-VIBE sequence. Second, as our subjects were above 44 y old, the direct application of the proposed method to pediatric patients is difficult because of differences in

anatomy and bone densities. Third, we have assessed the impact of our method on the quantification of only a limited number of synthetic pelvic lesions. Finally, although we obtained encouraging results for the subjects who underwent repeated examinations, the reproducibility of our method needs to be properly assessed.

## CONCLUSION

We described the development and initial validation of DIVIDE, a deep-learning approach to estimate pelvis pseudo-CT from standard Dixon-VIBE images to be used for AC in integrated PET/MRI scanners. We showed that the proposed method introduces only minimal bias compared with the CT-based approach, the current standard for AC.

## DISCLOSURE

This project was partially supported by Real Colegio Complutense at Harvard University Research Fellowships, project DPI2015-68664-C4-2-R of the Spanish Ministry of Economy and Innovation, and NIH-NCI grant 1R01CA218187-01A1. No other potential conflict of interest relevant to this article was reported.

## ACKNOWLEDGMENT

We thank NVIDIA Corporation for donating the Tesla K40 and the Titan X Pascal GPUs used for this research.

## REFERENCES

- Shao Y, Cherry S, Farahani K, et al. Simultaneous PET and MR imaging. *Phys Med Biol.* 1997;42:1965–1970.
- Herzog H, Pietrzyk U, Shah NJ, Ziemons K. The current state, challenges and perspectives of MR-PET. *Neuroimage.* 2010;49:2072–2082.
- Beyer T, Lassen ML, Boellaard R, et al. Investigating the state-of-the-art in whole-body MR-based attenuation correction: an intra-individual, inter-system, inventory study on three clinical PET/MR systems. *MAGMA.* 2016;29:75–87.
- Andersen FL, Ladefoged CN, Beyer T, et al. Combined PET/MR imaging in neurology: MR-based attenuation correction implies a strong spatial bias when ignoring bone. *Neuroimage.* 2014;84:206–216.
- Hofmann M, Pichler B, Schölkopf B, Beyer T. Towards quantitative PET/MRI: a review of MR-based attenuation correction techniques. *Eur J Nucl Med Mol Imaging.* 2009;36(suppl 1):S93–S104.
- Torrado-Carvajal A, Herraiz JL, Alcaín E, et al. Fast patch-based pseudo-CT synthesis from T1-weighted MR images for PET/MR attenuation correction in brain studies. *J Nucl Med.* 2016;57:136–143.

7. Izquierdo-Garcia D, Hansen AE, Förster S, et al. An SPM8-based approach for attenuation correction combining segmentation and nonrigid template formation: application to simultaneous PET/MR brain imaging. *J Nucl Med*. 2014;55:1825–1830.
8. Torrado-Carvajal A, Herraiz JL, Hernandez-Tamames JA, et al. Multi-atlas and label fusion approach for patient-specific MRI based skull estimation. *Magn Reson Med*. 2016;75:1797–1807.
9. Martinez-Möller A, Souvatzoglou M, Delso G, et al. Tissue classification as a potential approach for attenuation correction in whole-body PET/MRI: evaluation with PET/CT data. *J Nucl Med*. 2009;50:520–526.
10. Bowen SL, Fuin N, Levine MA, Catana C. Transmission imaging for integrated PET-MR systems. *Phys Med Biol*. 2016;61:5547–5568.
11. Ladefoged CN, Law I, Anazodo U, et al. A multi-centre evaluation of eleven clinically feasible brain PET/MRI attenuation correction techniques using a large cohort of patients. *Neuroimage*. 2017;147:346–359.
12. Izquierdo-Garcia D, Catana C. MR imaging-guided attenuation correction of PET data in PET/MR imaging. *PET Clin*. 2016;11:129–149.
13. Leynes AP, Yang J, Shanbhag DD, et al. Hybrid ZTE/Dixon MR-based attenuation correction for quantitative uptake estimation of pelvic lesions in PET/MRI. *Med Phys*. 2017;44:902–913.
14. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, Inception-ResNet and the impact of residual connections on learning. arXiv.org website. <https://arxiv.org/pdf/1602.07261.pdf>. Published August 23, 2016. Accessed October 15, 2018.
15. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Lecture Notes in Computer Science, vol. 9351. Cham, Switzerland: Springer; 2015:234–241.
16. Badrinarayanan V, Kendall A, Cipolla R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation. arXiv.org website. <https://arxiv.org/pdf/1511.00561.pdf>. Published October 10, 2016. Accessed October 15, 2018.
17. Chen L-C, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation. arXiv.org website. <https://arxiv.org/abs/1706.05587>. Submitted June 17, 2017. Revised December 5, 2017. Accessed October 15, 2018.
18. Gregor K, Danihelka I, Graves A, Rezende DJ, Wierstra D. DRAW: a recurrent neural network for image generation. arXiv.org website. <https://arxiv.org/abs/1502.04623>. Submitted February 16, 2015. Revised May 20, 2015. Accessed October 15, 2018.
19. van den Oord A, Kalchbrenner N, Espeholt L, Vinyals O, Graves A. Conditional image generation with pixelcnn decoders. In Proceedings of the 30th International Conference on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc., USA; 2016:4797–4805.
20. Milletari F, Navab N, Ahmadi S-A. V-net: fully convolutional neural networks for volumetric medical image segmentation. In Proceeding of the 2016 Fourth International Conference on 3D Vision (3DV). New York, NY: IEEE, USA; 2016:565–571.
21. Han X. MR-based synthetic CT generation using a deep convolutional neural network method. *Med Phys*. 2017;44:1408–1419.
22. Kamnitsas K, Ledig C, Newcombe VFJ, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal*. 2017;36:61–78.
23. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88.
24. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev Biomed Eng*. 2017;19:221–248.
25. Ribeiro A, Mops E, Herzog H, Almeida P. Hybrid approach for attenuation correction in PET/MR scanners. *Nucl Instrum Methods Phys Res A*. 2014;734:166–170.
26. Leynes AP, Yang J, Wiesinger F, et al. Zero-echo-time and Dixon deep pseudo-CT (ZeDD CT): direct generation of pseudo-CT images for pelvic PET/MRI attenuation correction using deep convolutional neural networks with multiparametric MRI. *J Nucl Med*. 2018;59:858–858.
27. Kikinis R, Pieper SD, Vosburgh KG. 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support. In: Jolesz FA, ed. *Intraoperative Imaging and Image-Guided Therapy*. New York, NY: Springer; 2014:277–289.
28. Penny WD, Friston KJ, Ashburner JT, Kiebel SJ, Nichols TE. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. London, U.K.: Academic Press; 2011.
29. 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.
30. Abadi M, Barham P, Chen J, et al. TensorFlow: a system for large-scale machine learning. usenix website. <https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf>. Accessed January 16, 2019.
31. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv.org website. <https://arxiv.org/abs/1409.1556>. Submitted September 4, 2014. Revised April 10, 2015. Accessed October 15, 2018.
32. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv.org website. <https://arxiv.org/pdf/1502.03167v3.pdf>. Accessed January 16, 2019.
33. Dumoulin V, Visin F. A guide to convolution arithmetic for deep learning. <https://arxiv.org/abs/1603.07285>. arXiv.org website. Submitted March 23, 2016. Revised January 11, 2018. Accessed October 15, 2018.
34. Burger C, Goerres G, Schoenes S, Buck A, Lonn AH, Von Schulthess GK. PET attenuation coefficients from CT images: experimental evaluation of the transformation of CT into PET 511-keV attenuation coefficients. *Eur J Nucl Med Mol Imaging*. 2002;29:922–927.
35. Berthon B, Håggström I, Apte A, et al. PETSTEP: generation of synthetic PET lesions for fast evaluation of segmentation methods. *Phys Med*. 2015;31:969–980.
36. Liu F, Jang H, Kijowski R, Bradshaw T, McMillan AB. Deep learning MR imaging-based attenuation correction for PET/MR imaging. *Radiology*. 2018;286:676–684.