HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2019 October 29.

Published in final edited form as:

Neuroimage. 2017 February 15; 147: 346–359. doi:10.1016/j.neuroimage.2016.12.010.

A multi-centre evaluation of eleven clinically feasible brain PET/MRI attenuation correction techniques using a large cohort of patients

Claes N. Ladefoged[a], **Ian Law**[a], **Udunna Anazodo**[b], **Keith St. Lawrence**[b], **David Izquierdo-Garcia**[c], **Ciprian Catana**[c], **Ninon Burgos**[d], **M. Jorge Cardoso**[d,e], **Sebastien Ourselin**[d,e], **Brian Hutton**[f], **Inés Mérida**[g,h], **Nicolas Costes**[g], **Alexander Hammers**[g,i], **Didier Benoit**[a], **Søren Holm**[a], **Meher Juttukonda**[j], **Hongyu An**[j], **Jorge Cabello**[k], **Mathias Lukas**[k], **Stephan Nekolla**[k], **Sibylle Ziegler**[k], **Matthias Fenchel**[l], **Bjoern Jakoby**[l,o], **Michael E. Casey**[m], **Tammie Benzinger**[n], **Liselotte Højgaard**[a], **Adam E. Hansen**[a], **Flemming L. Andersen**[a,*]

[a]Department of Clinical Physiology, Nuclear Medicine and PET, Rigshospitalet Copenhagen, Denmark

[b]Lawson Health Research Institute, London, ON, Canada

[c]Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA, USA

[d]Translational Imaging Group, Centre for Medical Image Computing, University College London, NW1 2HE, London, UK

[e]Dementia Research Centre, Institute of Neurology, University College London, WC1N 3AR, London, UK

[f]Institute of Nuclear Medicine, University College London, London, UK

[g]LILI-EQUIPEX – Lyon Integrated Life Imaging: hybrid MR-PET, CERMEP Imaging Centre, Lyon, France

[h]Siemens Healthcare France SAS, Saint-Denis, France

[i]King's College London & Guy's and St Thomas' PET Centre, Division of Imaging Sciences and Biomedical Engineering, King's College London, London, UK

[j]Joint Department of Biomedical Engineering, The University of North Carolina at Chapel Hill and North Carolina State University, Chapel Hill, NC 27599, USA

[k]Nuklearmedizinische Klinik und Poliklinik, Klinikum rechts der Isar, Technische Universitaet Muenchen, Munich, Germany

[l]Siemens Healthcare GmbH, Erlangen, Germany

*Corresponding author. fling@rh.dk (F.L. Andersen).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuroimage.2016.12.010.

[m]Siemens Medical Solutions USA, Inc., Knoxville, TN, USA

[n]Mallinckrodt Institute of Radiology, Washington University, St. Louis, MO 63130, USA

[o]University of Surrey, Guildford, Surrey, UK

Abstract

Aim: To accurately quantify the radioactivity concentration measured by PET, emission data need to be corrected for photon attenuation; however, the MRI signal cannot easily be converted into attenuation values, making attenuation correction (AC) in PET/MRI challenging. In order to further improve the current vendor-implemented MR-AC methods for absolute quantification, a number of prototype methods have been proposed in the literature. These can be categorized into three types: template/atlas-based, segmentation-based, and reconstruction-based. These proposed methods in general demonstrated improvements compared to vendor-implemented AC, and many studies report deviations in PET uptake after AC of only a few percent from a gold standard CT-AC. Using a unified quantitative evaluation with identical metrics, subject cohort, and common CT-based reference, the aims of this study were to evaluate a selection of novel methods proposed in the literature, and identify the ones suitable for clinical use.

Methods: In total, 11 AC methods were evaluated: two vendor-implemented (MR-AC_{DIXON} and MR-AC_{UTE}), five based on template/atlas information (MR-$AC_{SEGBONE}$ (Koesters et al., 2016), MR-$AC_{ONTARIO}$ (Anazodo et al., 2014), MR-AC_{BOSTON} (Izquierdo-Garcia et al., 2014), MR-AC_{UCL} (Burgos et al., 2014), and MR-$AC_{MAXPROB}$ (Merida et al., 2015)), one based on simultaneous reconstruction of attenuation and emission (MR-AC_{MLAA} (Benoit et al., 2015)), and three based on image-segmentation (MR-AC_{MUNICH} (Cabello et al., 2015), MR-$AC_{CAR\text{-}RiDR}$ (Juttukonda et al., 2015), and MR-$AC_{RESOLUTE}$ (Ladefoged et al., 2015)). We selected 359 subjects who were scanned using one of the following radiotracers: [^{18}F]FDG (210), [^{11}C]PiB (51), and [^{18}F] florbetapir (98). The comparison to AC with a gold standard CT was performed both globally and regionally, with a special focus on robustness and outlier analysis.

Results: The average performance in PET tracer uptake was within ± 5% of CT for all of the proposed methods, with the average ± SD global percentage bias in PET FDG uptake for each method being: MR-AC_{DIXON} (−11.3 ± 3.5)%, MR-AC_{UTE} (−5.7 ± 2.0)%, MR-$AC_{ONTARIO}$ (−4.3 ± 3.6)%, MR-AC_{MUNICH} (3.7 ± 2.1)%, MR-AC_{MLAA} (−1.9 ± 2.6)%, MR-$AC_{SEGBONE}$ (−1.7 ± 3.6)%, MR-AC_{UCL} (0.8 ± 1.2)%, MR-$AC_{CAR\text{-}RiDR}$ (−0.4 ± 1.9)%, MR-$AC_{MAXPROB}$ (−0.4 ± 1.6)%, MR-AC_{BOSTON} (−0.3 ± 1.8)%, and MR-$AC_{RESOLUTE}$ (0.3 ± 1.7)%, ordered by average bias. The overall best performing methods (MR-AC_{BOSTON}, MR-$AC_{MAXPROB}$, MR-$AC_{RESOLUTE}$ and MR-AC_{UCL}, ordered alphabetically) showed regional average errors within ± 3% of PET with CT-AC in all regions of the brain with FDG, and the same four methods, as well as MR-$AC_{CAR\text{-}RiDR}$, showed that for 95% of the patients, 95% of brain voxels had an uptake that deviated by less than 15% from the reference. Comparable performance was obtained with PiB and florbetapir.

Conclusions: All of the proposed novel methods have an average global performance within likely acceptable limits (± 5% of CT-based reference), and the main difference among the methods was found in the robustness, outlier analysis, and clinical feasibility. Overall, the best performing methods were MR-ACBOSTON, MR-ACMAXPROB, MR-ACRESOLUTE and MR-

ACUCL, ordered alphabetically. These methods all minimized the number of outliers, standard deviation, and average global and local error. The methods MR-ACMUNICH and MR-ACCAR-RiDR were both within acceptable quantitative limits, so these methods should be considered if processing time is a factor. The method MR-ACSEGBONE also demonstrates promising results, and performs well within the likely acceptable quantitative limits. For clinical routine scans where processing time can be a key factor, this vendor-provided solution currently outperforms most methods. With the performance of the methods presented here, it may be concluded that the challenge of improving the accuracy of MR-AC in adult brains with normal anatomy has been solved to a quantitatively acceptable degree, which is smaller than the quantification reproducibility in PET imaging.

Keywords

Attenuation correction; PET/MRI; Brain

1. Introduction

Positron emission tomography/magnetic resonance imaging (PET/MRI) combines the powerful and functional imaging of PET for assessment of patients in oncology (Hillner et al., 2008) and neurology (Heiss, 2009) with MRI's excellent soft-tissue characterization, contrast, and power to provide additional functional information like perfusion and diffusion. A likely key clinical application of simultaneous PET/MRI is in neurological disorders, in particular in dementia (Bailey et al., 2015; Drzezga et al., 2014; Dukart et al., 2011). In order to accurately quantify the radioactivity concentration measured by PET, emission data need to be corrected for photon attenuation.

Currently, two fully integrated simultaneous whole-body human PET/MRI systems are commercially available: the Biograph mMR (Siemens Healthcare GmbH, Erlangen, Germany) and the Signa PET/MRI (GE Healthcare, Waukesha WI, USA). The vendor-implemented attenuation correction (AC) method in the Signa PET/MRI system is based on an atlas of MRI/CT pairs used to derive the approximate size and location of bones and air cavities (Wollenweber et al., 2013) but is yet to be fully evaluated on a large patient cohort (Sekine et al., 2016a). Two AC-methods are implemented in the Biograph mMR system. They are both segmenting the MR image into tissue classes: the two-point Dixon-Water-Fat sequence (Dixon) based on Martinez-Möller et al. (2009), which can lead to underestimation of PET tracer uptake near the skull (Andersen et al., 2014), and the ultrashort echo time sequence (UTE), based on a similar acquisition method as Catana et al. (2010) and a different segmentation method based on a voxel-based classifier, which is challenged by remaining segmentation errors in the skull base and near the frontal sinuses (Dickson et al., 2014). A third, atlas-based AC method is currently being developed (Koesters et al., 2016; Paulus et al., 2015).

To further improve the quantitative accuracy, a number of methods have been proposed in the literature. These can be categorized into three types: template/atlas-based, segmentation-based, and reconstruction-based. The template/atlas-based methods create a pseudo-CT image by co-registering database-subjects to a patient's MR image. The atlas or template

can be based on pairs of CT and dual UTE (Delso et al., 2014b; Roy et al., 2014), CT and anatomical MRI (Dixon, T1, and/or T2-weighted sequences) (Anazodo et al., 2014; Andreasen et al., 2015; Burgos et al., 2014; Izquierdo-Garcia et al., 2014; Merida et al., 2015; Schreibmann et al., 2010; Sekine et al., 2016a; Torrado-Carvajal et al., 2016). Some methods also use probabilistic measures or machine learning techniques to compare with an atlas or template (Chan et al., 2013; Chen et al., 2015; Hofmann et al., 2008; Johansson et al., 2011; Navalpakkam et al., 2013; Poynton et al., 2014; Yang et al., 2013). Alternatives to template/atlas-based methods are the segmentation-based methods, which derive attenuation maps purely from MR images. These methods usually segment the T1 (Fei et al., 2012; Zaidi et al., 2003) or UTE images (Cabello et al., 2015; Catana et al., 2010; Juttukonda et al., 2015; Ladefoged et al., 2015) into multiple tissue classes, each assigned a mean linear attenuation coefficient (LAC). Furthermore, a number of UTE-based methods calculate a patient specific bone density (Cabello et al., 2015; Juttukonda et al., 2015; Ladefoged et al., 2015) using the normalized logarithmic difference between the UTE images (R_2* map) (Keereman et al., 2010). A number of methods also exist employing special MRI sequences with modified sampling schemes using either short echo time (STE)/Dixon (Khateri et al., 2015) or improved UTE sequences, using point-wise encoding time reduction with radial acquisition (PETRA) (Grodzki et al., 2012), using zero TE (ZTE) (Wiesinger et al., 2016), or using triple UTE (Aitken et al., 2014; Berker et al., 2012; Su et al., 2015). Finally, maximum-likelihood reconstruction-based methods, where activity and attenuation are simultaneously reconstructed based on maximum-likelihood (Nuyts et al., 1999), using either time-of-flight information (TOF) (Mehranian and Zaidi, 2015; Rezaei et al., 2012, 2014; Salomon et al., 2011) or non-TOF information using MR-priors (Benoit et al., 2015), are yet to be evaluated for PET/MRI-brain applications. The methods using TOF would currently be limited to the Signa PET/MRI system or the sequential Ingenuity TF PET/MRI system (Philips Healthcare, Cleveland, OH) (Zaidi et al., 2011). However, recently, a TOF-based MLAA implementation failed to show additional benefit for attenuation correction in PET/MR imaging compared to a robust atlas-based method (Mehranian et al., 2016).

The proposed methods demonstrate improvements compared to the currently vendor-implemented AC, and many studies report deviations in PET uptake after AC of only a few percent from a gold standard CT-AC; or conclude that the methods are ready for clinical use on the basis of quantitative accuracy (Burgos et al., 2014; Izquierdo-Garcia et al., 2014; Ladefoged et al., 2015). There are publications indicating the desire for more accuracy in PET brain attenuation correction (Andersen et al., 2014; Dickson et al., 2014; Hitz et al., 2014), but whether the problem of AC has been solved fully is still debatable (Bailey et al., 2015).

There is currently no consensus on a method for summarizing evaluations, and this lack of standardization of metrics makes direct comparisons across the published AC methods difficult. The methods rely on different gold/silver standards or make use of cohorts with different diseases or imaging tracers. The cohorts are of varying sizes, from 5 (Navalpakkam et al., 2013) to 92 subjects (Juttukonda et al., 2015), which might not be enough to ensure robustness and capture potential outliers in a clinical setting.

In this work, we present the performance of prototype methods within a multi-centre study, which include:

- A large patient cohort (n=359) who were scanned with three different radiotracers ([^{18}F]fluoro-D-glucose (FDG), [^{11}C] Pittsburgh-compound-B (PiB), and [^{18}F]florbetapir), allowing for detailed evaluation of outliers and robustness.
- Patients recruited from clinical populations with established or emerging clinical indications for both PET and MRI, and therefore well-suited for evaluating regional effects of brain MR-AC.
- A unified comparison between 11 PET/MRI brain AC methods using identical metrics and a common CT-based reference: two vendor-implemented (Dixon and UTE), five based on template/atlas-based information (Anazodo et al., 2014; Burgos et al., 2014; Izquierdo-Garcia et al., 2014; Koesters et al., 2016; Merida et al., 2015), one based on simultaneous reconstruction of attenuation and emission (Benoit et al., 2015), and three based on image-segmentation (Cabello et al., 2015; Juttukonda et al., 2015; Ladefoged et al., 2015).

The methods were assessed both globally and regionally using multiple metrics, including a detailed robustness and outlier analysis across the whole patient group. We also compared the methods by assessing the time that was required to execute the methods, referring to the actual run-time and scan-time of the corresponding sequences. While time may not be of high importance in a research setting, a clinical setup may demand high requirements for this factor in order to provide a timely diagnosis and acceptable patient management and throughput. The aim of this study was to evaluate the proposed methods and identify the ones suitable for clinical use by performing a unified quantitative evaluation.

2. Materials and methods

All patient studies were performed using a fully-integrated PET/MRI system (Siemens Biograph mMR, Siemens Healthcare, Erlangen, Germany) (Delso et al., 2011). For the purpose of obtaining a reference low-dose CT image of the head, whole-body PET/CT systems were used (Biograph TruePoint 40 and Biograph TruePoint 64, Siemens Healthcare) (Jakoby et al., 2009).

2.1. Patients

Data sets were obtained retrospectively from two different centers; Rigshospitalet, University Hospital Copenhagen, Denmark, and Washington University, St. Louis, United States of America. The Rigshospitalet data were comprised from data sets from the complete cohort of patients referred for a clinical FDG PET/MRI brain examination, either in neuro-oncology or dementia, as well as the complete cohort of subjects having undergone PiB PET/MRI examinations. These cohorts included healthy subjects and patients with mild cognitive impairment (MCI), Alzheimer’s disease (AD), or clinical dementia. In total, 210 subjects with FDG PET/MRI scans (5 healthy volunteers, 4 neuro-oncological patients, 201 MCI, AD or clinical dementia patients) and 51 subjects scanned with PiB (1 healthy volunteer, 50 MCI, AD or clinical dementia patients) were consecutively selected between

November 2013 and April 2015. From Washington University, data sets were obtained from 98 subjects referred for a florbetapir examination (3 very mild dementia, 95 MCI or clinical dementia patients).

Of a total of 359 PET/MRI brain studies from the two centers, 22 studies were excluded (9 FDG, 3 PiB and 10 florbetapir). Exclusion was due to metal implant-induced artifacts in the brain images (3 subjects) or data errors due to corruption or missing raw-data (19 subjects) leaving 337 studies available for analysis. The local ethics committees had approved the original studies and all patients gave informed consent.

2.2. Imaging protocol

2.2.1. PET—The PET scan information is given in Table 1. Patients were positioned head first with their arms down on the fully-integrated PET/MRI system. Data were acquired over a single bed position of 25.8 cm covering the head and neck for 10/20/15 min for FDG/PiB/florbetapir respectively (Table 1 for initiation of scan). For the purpose of this study, the PET data from the PET/MR acquisition were reconstructed using 3D Ordinary Poisson-Ordered Subset Expectation Maximization (OP-OSEM) with 4/3/4 iterations, 21 subsets, and 3/5/5 mm Gaussian post filtering for FDG/PiB/florbetapir, respectively, on 344×344 matrices (2.1×2.1×2.0 mm^3 voxels) in line with the respective clinical protocols.

2.2.2. MRI—The scan protocol included two-point Dixon images using the vendor-provided DIXON-VIBE AC sequence with repetition time (TR)/echo time 1 (TE1)/ echo time 2 (TE2)=2300/1.23/2.46 ms, flip angle 10°, coronal orientation, voxel size of 2.6×2.6×3.12 mm^3, 19 s acquisition time; a UTE AC sequence with TR/TE1/ TE2=11.94/0.07/2.46 ms, a flip angle of 10°, axial orientation, field of view (FOV) of 300 mm^2, reconstructed on 192×192×192 matrices (1.6×1.6×1.6 mm^3 voxels), 100 s acquisition time; and a T1w MPRAGE with TR/TE=1900/2.44 ms (FDG/PiB) and 2300/2.95 ms (florbetapir), inversion time=900 ms, flip angle=9°, sagittal orientation, reconstructed on 512×512×192 matrices (0.49×0.49×1 mm^3 voxels) (FDG/PiB) and 256×240×176 matrices (1.05×1.05×1.2 mm^3 voxels) (florbetapir), 300 s acquisition time. The patients injected with FDG or PiB were all imaged using the software version VB20P, whereas those injected with florbetapir were imaged with VB18P, which included a work-inprogress version of the UTE sequence. Due to the difference in version numbering, mainly expressed in the UTE sequences, the tracers were evaluated separately.

2.2.3. CT—A low dose CT (120 kVp, 36 mAs, 74 slices, 0.6×0.6×3 mm^3 voxels) was acquired from each patient to be used as gold standard. The CT images were acquired either on the same day as the PET/MRI examination (FDG and PiB) or within 8.3 ± 6 days of each other with no surgical procedures in between (florbetapir).

2.3. Creating the attenuation maps

In all, nine proposed MR-AC methods were selected for evaluation. These were selected to be representative of the three major types of MR-AC categories (template/atlas-based, segmentation-based, and reconstruction-based), and had to be applicable to the Siemens mMR data available. Furthermore, they should be implemented on a central site

(Copenhagen). Twelve attenuation maps were created. First, vendor-provided MR-based attenuation maps (MR-AC_{DIXON} and MR-AC_{UTE}) were derived using the DIXON VIBE sequence and the UTE MR sequence. Next, the nine proposed brain MR-AC methods were used. A brief introduction of each method is given below. We refer to the original publications for detailed explanations. Finally, for each subject, the CT image was co-registered to the T1 image using a 6-parameter rigid alignment procedure (minctracc, McConnell Imaging Center, Montreal, Canada) with normalized mutual information as objective function. The patient bed and head-holder were extracted manually from the CT images using an oval ROI (OsiriX software) applied to each transverse CT slice. The attenuation values were then converted from HUs to LACs at 511 keV by using a standard bi-linear scaling approach (Carney et al., 2006) as implemented in the Siemens PET/CT systems used. The co-registered CT attenuation map was substituted into the UTE attenuation file (MR-AC_{CT}) to facilitate use of routine reconstruction. We blurred the resulting image with a 4 mm Gaussian filter to simulate PET/CT reconstructions. Due to the limited coverage in the neck region by the acquired CT, we replaced the missing area by the values from MR-AC_{DIXON}. To ensure a fair comparison, this replacement was also performed in all the other attenuation maps.

All attenuation maps, except for MR-AC_{DIXON}, were created in the image space defined by the method, and post-resampled and placed into the UTE file just as the reference CT. Since the CT image was co-registered to the T1 image, we also co-registered the Dixon and UTE images to the T1 image, and aligned the MR-AC maps of the methods defined using either of these.

All of the methods presented in this study, except for two, had been made available and were processed in Copenhagen. The MR-$AC_{SEGBONE}$ attenuation maps were created by Siemens, and MR-AC_{UCL} attenuation maps were created in London, as the CT/MRI database could not be shared at the time of this study.

2.3.1. Template/Atlas-based methods (Segbone, Ontario, Boston, UCL, MaxProb)

- ***Siemens Healthcare GmbH - Segbone:*** The prototype method by Siemens Healthcare (Koesters et al., 2016; Paulus et al., 2015) is a template-based method, where an MR model image of the skull region is non-rigidly registered to an individual subject exclusively on the Dixon input images. A pre-aligned skull mask is then brought to subject space, and attenuation values greater than soft-tissue are superimposed on the original Dixon attenuation map. This method is still a work-in-progress. The attenuation map is denoted MR-$AC_{SEGBONE}$.
- ***London Ontario:*** The method by Anazodo et al. (2014) uses the SPM8 function ‘new segment’ to extract a bone probability map from each individual’s T1 MPRAGE image. The bone map defined by voxels with probability above 80% are, following a post-processing procedure, superimposed on MR-AC_{DIXON} and assigned the constant value 0.143 cm^{-1}. The resulting attenuation map is denoted MR-$AC_{ONTARIO}$.

- ***Boston MGH:*** The method by Izquierdo-Garcia et al. (2014) also uses SPM8 to extract patient specific tissue probability maps from the individual's intensity normalized T1 MPRAGE image and register these to a template of co-aligned probability maps of 15 subjects. The averaged corresponding CT template is then back-warped to patient space, converted to LACs, blurred with a 4 mm Gaussian filter, and used as the template-based AC map MR-AC_{BOSTON}.
- ***London UCL:*** The method by Burgos et al. (2014) non-linearly aligns a database of 41 T1-CT pairs to the intensity-normalized T1 image of a patient. At each voxel, a patch is extracted and the local normalized correlation coefficient is calculated between the patient T1 and each of the T1 images in the database. A weighting vector is calculated from the coefficients ranking each T1-CT pair relative to the other pairs, and used to obtain a target CT voxel value. The resulting synthesized CT image is converted to LACs, blurred with a 4 mm Gaussian filter, and denoted MR-AC_{UCL}.
- ***Lyon - MaxProb:*** The method by Merida et al. (2015) is similar to Burgos et al. (2014) by also aligning a database of 27 T1-CT pairs to an intensity normalized T1 MPRAGE image of a patient, but deviates in the method used to combine the database into a target CT image. The authors discretize the CT images into air, soft tissue and bone voxels, calculate the modal tissue type across the database, and assign the mean of the CT values from the database subjects with that tissue type to the voxel. The final fused image is converted to LACs, blurred with a 4 mm Gaussian filter, and denoted MR-$AC_{MaxProb}$.

2.3.2. Maximum-likelihood reconstruction-based

- ***Copenhagen - MLAA:*** The method by Benoit et al. (2015) is a reconstruction-based method aimed at optimizing the existing non-TOF MLAA technique. The method uses the individual's MR-AC_{UTE} and a discretized version of the individual's T1 image as prior information to the MLTR algorithm. A heuristically determined schema assigns the value for α_j based on the two prior images, as well as the update term calculated in each MLTR iteration step. The final image of the MLAA algorithm is denoted MR-AC_{MLAA}.

2.3.3. Segmentation-based (Munich, CAR-RiDR, RESOLUTE)

- ***Munich:*** The method by Cabello et al. (2015) uses the individual's UTE TE1 and TE2 images to extract air, soft tissue as well as the bone signal, which is extracted using R_2* and scaled to LACs by normalizing to the maximum intensity and thresholding using an empirically chosen value. The intensity of the bone voxels from the R_2* map are equalized by the mean and maximum values from a CT-based database. The resulting image is blurred with a 3 mm Gaussian filter, and denoted MR-AC_{MUNICH}.
- ***University of North Carolina - CAR-RiDR:*** The method by Juttukonda et al. (2015) also uses the individual's UTE TE images to extract air, soft tissue and R_2* signal. By fitting a sigmoid-best-fit to a set of 98 R_2*-CT relationship pairs,

the authors convert the measured R_2* signal to a continuous CT value. The resulting image is denoted MR-$AC_{CAR\text{-}RiDR}$.

- ***Copenhagen - RESOLUTE:*** The method by Ladefoged et al. (2015) segments brain, CSF, soft tissue, and air from the UTE TE images, and also uses R_2* to extract a bone signal, which is converted to LACs using a set of R_2*-CT relationship pairs from 10 training patients. To limit possible bias from regions with known R_2* noise (Delso et al., 2014a), and because the amount of included bone is higher compared to Cabello et al. (2015), Juttukonda et al. (2015), the threshold for included bone was regionally varied within the head. The resulting image is blurred with a 4 mm Gaussian filter and denoted MR-$AC_{RESOLUTE}$.

2.4. Image analysis

As Dixon-based methods may suffer from fat/water inversion (Ladefoged et al., 2014), we excluded the patients with inversion from further analysis in most of the metrics. This was the case for nine FDG, one PiB, and three florbetapir patients. To avoid bias in the results, the patients with fat/water inversion are excluded in most figures. When included, it will be clearly stated. The methods directly affected by the inversion are MR-$AC_{SEGBONE}$, MR-$AC_{ONTARIO}$ and MR-$AC_{CAR\text{-}RiDR}$.

2.4.1. Global attenuation map performance—To evaluate the performance precision of the methods based on the attenuation maps, we measured the Jaccard-index for the bone compared to CT-derived attenuation maps. We defined bone as being greater than 0.101 cm^{-1}, representing values above soft-tissue.

2.4.2. Global PET performance—As a prior step to evaluating the global PET performance, we first created individual brain masks. We moved all patient data to common MNI space using ANTs (Avants et al., 2011) by non-rigidly registering the patients' T1 images to the ICBM 152 2009a template (Fonov et al., 2009). The MNI brain mask was back-warped to patient space, and the PET image reconstructed with the CT attenuation map was thresholded at > 20% of the maximum intensity value. Finally, the intersection between the two masks was calculated and used as the final mask of the brain, and used throughout this manuscript when referring to the brain region. Next, the voxel-wise percent difference relative to PET_{CT}, defined as:

$$Rel_{\%} = \frac{PET_x - PET_{CT}}{PET_{CT}} \times 100, \quad (1)$$

as well as the absolute percent difference, defined as:

$$Abs_{\%} = \frac{|PET_x - PET_{CT}|}{PET_{CT}} \times 100, \quad (2)$$

were calculated for the PET images corrected with each of the evaluated methods (PET_X). We calculated the averaged performance of $Rel_{\%}$ and $Abs_{\%}$ for the full brain across all patients for each of the methods. We computed the distribution of $Rel_{\%}$ errors for each of the methods using a histogram analysis. Finally, we computed the voxel-wise correlation

between PET_{CT} and each of the proposed methods for all voxels within the brain. We plotted the correlations in a joint histogram, and computed the goodness-of-fit value (R^2) to the identity line for each of the patients.

2.4.3. Regional PET performance—To evaluate the regional PET performance, we back-warped the anatomical predefined regions from MNI space (Collins et al., 1999; Fonov et al., 2009) to each patient, and extracted the mean $Rel_{\%}$ for each region in the template. We furthermore calculated the average and standard deviation $Rel_{\%}$-images across all patients for each of the methods for visual evaluation.

2.4.4. Analysis of robustness and outliers—To estimate the number of outliers for each method, we introduced a novel metric. The metric calculates the percentage of patients with at least x% of the brain within ± y% in the $Rel_{\%}$ images, where x varies from 0–100 and y={5,10,15}. A perfect method therefore has 100% of the patients within ± 0% in the $Rel_{\%}$ images. To understand the reason for the worst outliers, we then analyzed the three FDG patients with the worst performance based on the ± 10% threshold and categorized the error. We present the worst outlier of each method.

3. Results

3.1. Global attenuation map evaluation

Fig. 1 shows the axial, sagittal and coronal views for each attenuation method for a single sample patient. The representative sample patient was chosen objectively to be that which minimizes the difference to the median $Rel_{\%}$ error (Eq. (1)) across all methods.

The Jaccard similarity coefficient for bone for each tracer-group was comparable. For the largest group (FDG) the averaged scores ranged between 0.50–0.81 for the entire head, 0.40–0.77 in the bottom of the head covering from the eye socket and downwards, and 0.51–0.85 for the top of the head. The overall best performance was seen for $MR\text{-}AC_{BOSTON}$, $MR\text{-}AC_{MAXPROB}$, $MR\text{-}AC_{MUNICH}$, $MR\text{-}AC_{RESOLUTE}$ and $MR\text{-}AC_{UCL}$, in alphabetical order, with Jaccard indices of 0.68–0.81 in the entire head, and 0.79–0.85 in the top of the head. The individual result for all tracer-groups and methods are shown in Table 2. Note that $MR\text{-}AC_{SEGBONE}$ was not evaluated on florbetapir-data, and thus no results are available.

3.2. Global PET performance

The reconstructed PET images for the sample patient in Fig. 1 are shown for the same slice in Fig. 2. Visually comparing the maximum uptake areas in PET_{CT}, it is apparent that PET_{DIXON} and PET_{UTE} underestimate the PET uptake, whereas all the alternative methods improve the performance. For this patient, the average error for the entire brain was between −3.3% (Ontario) and 3.9% (Munich) for the proposed methods, and −10% and −5.3% for Dixon and UTE, respectively. The global $Rel_{\%}$ performance for the entire brain is shown for all 201 FDG patients and 11 methods in the boxplot in Fig. 3. The average performance was within ± 5% of PET_{CT} for all of the proposed methods, with the best performance seen for $MR\text{-}AC_{BOSTON}$, $MR\text{-}AC_{CAR\text{-}RiDR}$, $MR\text{-}AC_{MAXPROB}$, $MR\text{-}AC_{RESOLUTE}$, and $MR\text{-}AC_{UCL}$,

ordered alphabetically. MR-AC_{DIXON} showed −11% and MR-AC_{UTE} showed −7%. The global $Abs_{\%}$ performance for the entire brain is summarized for each method and tracer in Supplementary Table 1. The global histogram analysis is shown in Supplemental Figs. 1–3 for the three tracers. The averaged voxel-wise joint histograms are shown for PET_{CT} versus each of the proposed methods in Fig. 4 (FDG) and Supplemental Figs. 4–5 (PiB and Florbetapir). The highest correlation with the unity line was for FDG obtained by MR-$AC_{RESOLUTE}$ (0.96 ± 0.02), MR-AC_{UCL} (0.96 ± 0.02), MR-$AC_{MAXPROB}$ (0.96 ± 0.04), MR-AC_{BOSTON} (0.95 ± 0.03), and MR-$AC_{CAR\text{-}RiDR}$ (0.94 ± 0.03).

3.3. Regional PET performance

The regional $Rel_{\%}$ performance for the anatomical regions defined in the MNI template is shown in Fig. 5 for FDG and Supplemental Figs. 6–7 for PiB and Florbetapir. The best performance across all regions was again found with MR-AC_{BOSTON}, MR-$AC_{CAR\text{-}RiDR}$, MR-$AC_{MAXPROB}$, MR-$AC_{RESOLUTE}$ and MR-AC_{UCL}, ordered alphabetically. Only MR-$AC_{MAXPROB}$ and MR-$AC_{RESOLUTE}$, ordered alphabetically, had an average ± 1 SD error across all tracers below ± 5% in all regions. The averaged absolute difference, $Abs_{\%}$, is shown in Supplemental Figs. 8–9 for each of the regions and tracers. The average and standard deviation images of the $Rel_{\%}$ calculated for each FDG patient are shown in Figs. 6 and 7, respectively, and in Supplemental Figs. 10–13 for PiB and florbetapir. Considering all regions in the brain and all tracers, the overall best performance in terms of error and robustness was seen for MR-AC_{BOSTON}, MR-$AC_{MAXPROB}$, MR-$AC_{RESOLUTE}$ and MR-AC_{UCL}, ordered alphabetically.

3.4. Analysis of robustness and outliers

The result of the outlier analysis is shown in Fig. 8 for errors within ± 5%, ± 10%, and ± 15% (FDG) and Supplemental Fig. 14 (PiB and florbetapir). The proposed methods exceed the performance of the current Dixon and UTE methods. The cutoff where all of the patients are within ± 10% of PET_{CT} is achieved for MR-AC_{BOSTON}, MR-$AC_{MAXPROB}$, MR-$AC_{RESOLUTE}$ and MR-AC_{UCL}, ordered alphabetically, and at 90% of the brain, and between 60% and 80% for the other methods. At ± 15%, there is a clear separation between five of the methods (MR-AC_{BOSTON}, MR-$AC_{CAR\text{-}RiDR}$, MR-$AC_{MAXPROB}$, MR-$AC_{RESOLUTE}$ and MR-AC_{UCL}, ordered alphabetically) and the remaining methods. The robustness of each method can be assessed in the standard deviation images (Fig. 7). The best performance was obtained by MR-AC_{UCL} in the center of the brain, shortly followed by MR-$AC_{MAXPROB}$, MR-AC_{BOSTON} and MR-$AC_{RESOLUTE}$. In the cortical regions, the best performance was obtained by MR-$AC_{RESOLUTE}$ and MR-$AC_{CAR\text{-}RiDR}$. The three patients with the worst performance are categorized in Table 3, and the worst outlier is shown for each method in Supplemental Fig. 15.

In general, the worst outliers in the template/atlas-based methods (MR-$AC_{ONTARIO}$, MR-AC_{BOSTON}, MR-AC_{UCL}, MR-$AC_{MAXPROB}$) all have the same error category – the bone density is over or underestimated. Only a single case is due to registration errors (#1 in MR-$AC_{MAXPROB}$). The effect of the error is greatest in the cortical regions, with errors up to 30% locally.

The general errors in the proposed segmentation-based methods (MR-AC_{MUNICH}, MR-$AC_{CAR\text{-}RiDR}$, MR-$AC_{RESOLUTE}$) are over/under estimation of bone in air/tissue interface areas, resulting in local errors of up to 25% in the medulla and cerebellum. For the vendor-provided methods, the worst outliers were due to missing bone (MR-AC_{DIXON}) and underestimated bone density (MR-AC_{UTE}).

4. Discussion

This study facilitates a framework for comparing a large number of state-of-the-art methods using multiple metrics, and does so across a large patient cohort. We aimed to represent a broad spectrum of published methods, showing promises of clinical feasibility, and selected nine methods that were applicable to our large cohort of Siemens mMR examinations. Adding the vendor-provided methods, Dixon and UTE, we evaluated 11 methods for obtaining an MR-AC image. The key findings for the proposed methods were: almost all achieved relative errors below ± 5% globally and in all regions of the brain; the methods essentially differed in robustness, clinical feasibility, and outliers. Dixon and UTE have been included as current vendor references and confirm earlier results from Andersen et al. (2014), Dickson et al. (2014) with a tendency to global underestimation of activity.

The cerebellum is a region that is often used for normalization purposes, especially in kinetic modeling, and therefore requires extra attention. Considering the average ± 1 SD, the best performance is obtained with MR-AC_{BOSTON} and MR-$AC_{MAXPROB}$, shortly followed by most of the other methods; the only methods exceeding ± 5% were MR-AC_{MLAA} and MR-AC_{MUNICH} (Fig. 5). The maximum error in the cerebellum was 6–9% from the reference for most of the proposed methods (Supplemental Fig. 16), except for MR-$AC_{CAR\text{-}RiDR}$ (14%), MR-AC_{MLAA} and MR-AC_{MUNICH} (17%).

The result of the attenuation map evaluation (Table 2) illustrates that the multi-atlas and multi-registration based methods (MR-AC_{UCL}, MR-$AC_{MAXPROB}$) have the highest overall bone accuracy, with equally good results achieved in the upper part of the head by R_2*-based methods (MR-AC_{MUNICH}, MR-$AC_{RESOLUTE}$). This area-dependent difference could be explained by the challenges reported in air/tissue interface areas originating from the UTE sequence, such as the mouth and esophagus (Delso et al., 2014a). Metal implant-induced artifacts resulting in overestimation of bone in the mouth area especially affects the accuracy of the R_2*-based methods, but obviously also the remaining methods. Of all methods, two of the R_2*-based methods (MR-$AC_{CAR\text{-}RiDR}$ and MR-$AC_{RESOLUTE}$) had the lowest standard deviation in the cortical regions near the bone (Fig. 7). This can be explained by the fact that these methods extract patient specific bone density, whereas the template/atlas-based methods use an average CT value. This can in some cases lead to local under- or overestimation of the PET signal of up to 20% (Table 3, Supplemental Fig. 15). It should be noted that for the subjects where the CT include the dental region, the CTs were also affected by metal implants resulting in streak artifacts, which could bias the results further.

The template/atlas-based methods generally produced attenuation maps very similar to the CT-based attenuation maps, but they are based on the assumption that any new subject

can be represented by a template, a local combination of atlas patients, or by a non-rigid registration to a template. The template/atlas-based methods are therefore also potentially challenged by abnormal anatomy and patient surgery. The template/atlas-based methods are, furthermore, currently only optimized and tested for adults, as the application to children might require a purpose specific MRI/CT database. We did not include children, or any patients who have undergone anatomy-changing surgery in this study.

Most subjects chosen came from dementia studies, as this disease has the potential to be a key clinical application of PET/MRI (Bailey et al., 2015). The disease is clinically imaged with FDG, or amyloid binding radiotracers such as PiB or florbetapir, with the majority of patients having relatively widely distributed uptake in the brain, making the tracers well suited for a large-scale evaluation of MR-AC methods. The overall performance was, generally, similar across the three tracers for all methods (Supplemental Figs. 1–3).

When considering patient management, the acquisition time and number of steps required for generating AC maps has also to be taken into account. A potential challenge for the template/atlas-based methods is the long computation time due to multiple registrations for each patient in the database (Table 4). The MR-$AC_{MAXPROB}$ and MR-AC_{UCL} methods currently require at least an hour of processing time for each attenuation map, whereas the MR-AC_{MUNICH} and MR-$AC_{CAR\text{-}RiDR}$ methods only take a few seconds. The shorter processing time is usually a key argument for choosing the segmentation-based methods in fast-paced clinical services, whereas accuracy is much more important than speed in research applications. The MR-$AC_{SEGBONE}$ prototype requires less than 2 min of processing time, with further potential for speed-up. The processing time for multi-atlas methods could be shortened by code optimization and by reducing the number of database studies using pre-selection of the candidates more likely to match the patient (e.g. using gender, age, etc.) (Aljabar et al., 2009) but this requires further evaluation. An alternative is to use the newly proposed one-registration multi-atlas approach, which, however, comes at the cost of a lower accuracy in bone extraction compared to the multi-atlas approach (Arabi and Zaidi, 2016). In line with this, recent result on GE Signa system has shown, that compared to the clinical single-atlas, multi-atlas (MR-AC_{UCL}) is more accurate especially in regions close to the skull base (Sekine et al., 2016b). The emission-based method, MLAA, is currently computationally intensive, and relies on multiple MR sequences, which might complicate its use in a clinical setting.

We excluded a single machine-learning-based AC-method (Chen et al., 2015) from this study, as it had excessively long processing time for each patient (> 4 days), making it infeasible both in clinical routine and in this study in its current state.

4.1. Limitations

The use of a CT as gold standard is debatable. The CT image does not fully cover the FOV of the MR and PET area, and is further subject to streak artifacts from dental implants. Since the CT images are not simultaneously acquired with the PET/MRI, they are also subject to non-rigid movement, e.g. jaw movement and swallowing, not accounted for during the rigid registration. In lack of a true gold standard, such as a transmission scan with a rotating source, we considered the CT to be the reference.

The cohort, mainly consisting of patients with dementia and age-matched controls, represents an ageing population. This study did not include data to test the performance of these MR-AC methods on brain scans from younger patients. Furthermore, the study cohort did not include patients with focal PET lesions. Therefore, the performance of these methods in brain tumor patients has yet to be evaluated in most cases, although evaluation of individual methods (e.g. MR-AC_{BOSTON}) has already been performed in challenging brain tumors as in Izquierdo-Garcia et al. (2014).

The data from Washington University (florbetapir) were used to train the mapping of bone values from HU to LACs in the MR-$AC_{CAR\text{-}RiDR}$ method (Juttukonda et al., 2015). This could potentially bias the results.

5. Conclusion

All of the proposed methods have an average global performance within likely acceptable limits (± 5% of CT-based reference), and differences between methods were found in terms of the robustness, clinical feasibility and outlier analysis. The overall best performance was seen in the MR-AC_{BOSTON}, MR-$AC_{MAXPROB}$, MR-$AC_{RESOLUTE}$ and MR-AC_{UCL} methods, ordered alphabetically, which all minimize the number of outliers, standard deviation, and average global and local error. The methods MR-$AC_{CAR\text{-}RiDR}$ and MR-AC_{MUNICH} were both within acceptable quantitative limits, so these methods should be considered if processing time is a factor. The method MR-$AC_{SEGBONE}$ also demonstrates promising results, and performs well within the likely acceptable quantitative limits. For a clinical routine setting where time is critical, this vendor-provided prototype solution currently outperforms most methods, especially when acquisition time of the fast DIXON is included. With the performance of the methods presented here, it may be concluded that the challenge of improving the accuracy of MR-AC in adult brains with normal anatomy has been solved to an acceptable degree, with errors smaller than the quantification reproducibility in PET imaging.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

A special thanks to biomedical engineering student David Nyrnberg and computer science student Casper Hansen for aiding with the large data processing tasks involved in this study. The PET/MR system at Rigshospitalet was kindly provided by the John and Birthe Meyer Foundation, Denmark. At UCL this work was supported by an IMPACT studentship funded jointly by Siemens and the UCL Faculty of Engineering Sciences. Funding was also received from the National Institute for Health Research University College London Hospitals Biomedical Research Centre (NIHR BRC UCLH/UCL High Impact Initiative BW.mn.BRC10269) and the EPSRC (EP/K005278/1). In Lyon, this work was supported by a CIFRE studentship jointly funded by Siemens and the French national agency for research and Technology (ANRT). Funding was also received from the French national ‘invest for the future’ programs (LILI – Lyon Integrated Life Imaging: hybrid MR-PET ANR-11-EQPX-0026) and the hospital University Institute CESAME (Brain and Mental Health ANR-10-IBHU-0003). In Munich, this work was supported by the European Union Seventh Framework Program (FP7) under Grant Agreement n° 602621 Trimage and n° 294582-MUMI. In Boston, this work was supported by the NIH National Institute of Biomedical Imaging and Bioengineering (NIBIB) grant number 1R01EB014894.

References

Aitken AP, Giese D, Tsoumpas C, Schleyer P, Kozerke S, et al. , 2014. Improved UTE-based attenuation correction for cranial PET-MR using dynamic magnetic field monitoring. Med. Phys 41, 012302. [PubMed: 24387523]

Aljabar P, Heckemann RA, Hammers A, Hajnal JV, Rueckert D, 2009. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. Neuroimage 46 (3), 726–738. [PubMed: 19245840]

Anazodo UC, Thiessen JD, Ssali T, Mandel J, Gunther M, et al. , 2014. Feasibility of simultaneous whole-brain imaging on an integrated PET-MRI system using an enhanced 2-point Dixon attenuation correction method. Front. Neurosci 8, 434. [PubMed: 25601825]

Andersen FL, Ladefoged CN, Beyer T, Keller SH, Hansen AE, et al. , 2014. Combined PET/MR imaging in neurology: MR-based attenuation correction implies a strong spatial bias when ignoring bone. Neuroimage 84, 206–216. [PubMed: 23994317]

Andreasen D, Van Leemput K, Hansen RH, Andersen JA, Edmund JM, 2015. Patch-based generation of a pseudo CT from conventional MRI sequences for MRI- only radiotherapy of the brain. Med. Phys 42, 1596–1605. [PubMed: 25832050]

Arabi H, Zaidi H, 2016. One registration multi-atlas-based pseudo-CT generation for attenuation correction in PET/MRI. Eur. J. Nucl. Med. Mol. Imaging 43, 2021–2035. [PubMed: 27260522]

Avants BB, Tustison NJ, Song G, Cook PA, Klein A, et al. , 2011. A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage 54, 2033–2044. [PubMed: 20851191]

Bailey DL, Pichler BJ, Guckel B, Barthel H, Beer AJ, et al. , 2015. Combined PET/ MRI: Multi-modality Multi-parametric Imaging Is Here: Summary Report of the 4th International Workshop on PET/MR Imaging, February 23–27, 2015, Tubingen, Germany. Mol. Imaging Biol vol. 17, pp. 595–608. [PubMed: 26286794]

Benoit D, Ladefoged CN, Rezaei A, Keller SH, Andersen FL, Højgaard L, Hansen AE, Holm S, Nuyts J, 2016. Optimized MLAA for quantitative non-TOF PET/MR of the brain. Phys. Med. Biol 61 (24), 8854. [PubMed: 27910823]

Berker Y, Franke J, Salomon A, Palmowski M, Donker HC, et al. , 2012. MRI-based attenuation correction for hybrid PET/MRI systems: a 4-class tissue segmentation technique using a combined ultrashort-echo-time/Dixon MRI sequence. J. Nucl. Med 53, 796–804. [PubMed: 22505568]

Burgos N, Cardoso MJ, Thielemans K, Modat M, Pedemonte S, et al. , 2014. Attenuation correction synthesis for hybrid PET-MR scanners: application to brain studies. IEEE Trans. Med. Imaging 33, 2332–2341. [PubMed: 25055381]

Cabello J, Lukas M, Forster S, Pyka T, Nekolla SG, et al. , 2015. MR-based attenuation correction using ultrashort-echo-time pulse sequences in dementia patients. J. Nucl. Med 56, 423–429. [PubMed: 25678486]

Carney JP, Townsend DW, Rappoport V, Bendriem B, 2006. Method for transforming CT images for attenuation correction in PET/CT imaging. Med. Phys 33, 976–983. [PubMed: 16696474]

Catana C, van der Kouwe A, Benner T, Michel CJ, Hamm M, et al. , 2010. Toward implementing an MRI-based PET attenuation-correction method for neurologic studies on the MR-PET brain prototype. J. Nucl. Med 51, 1431–1438. [PubMed: 20810759]

Chan S-LS, Jeffree RL, Fay M, Crozier S, Yang Z, et al. , 2013. Automated Classification of Bone and Air Volumes for Hybrid PET-MRI Brain Imaging. International Conference on Digital Image Computing: Techniques and Applications (DICTA), Hobart, Australia, pp. 1–8.

Chen Y, Juttukonda M, Su Y, Benzinger T, Rubin BG, et al. , 2015. Probabilistic air segmentation and sparse regression estimated pseudo CT for PET/MR attenuation correction. Radiology 275, 562–569. [PubMed: 25521778]

Collins DL, Zijdenbos A, Baaré WC, Evans A, 1999. ANIMAL+INSECT: improved cortical structure segmentation. In: Kuba A, Šáamal M, Todd-Pokropek A (Eds.), Information Processing in Medical Imaging Springer, Berlin, Heidelberg, 210–223.

Delso G, Fürst S, Jakoby B, Ladebeck R, Ganter C, et al. , 2011. Performance measurements of the Siemens mMR integrated whole-body PET/MR scanner. J. Nucl. Med 52, 1914–1922. [PubMed: 22080447]

Delso G, Carl M, Wiesinger F, Sacolick L, Porto M, et al. , 2014a. Anatomic evaluation of 3-dimensional ultrashort-echo-time bone maps for PET/MR attenuation correction. J. Nucl. Med 55, 780–785. [PubMed: 24639457]

Delso G, Zeimpekis K, Carl M, Wiesinger F, Hullner M, et al. , 2014b. Cluster-based segmentation of dual-echo ultra-short echo time images for PET/MR bone localization. EJNMMI Phys 1, 7. [PubMed: 26501449]

Dickson JC, O'Meara C, Barnes A, 2014. A comparison of CT- and MR-based attenuation correction in neurological PET. Eur. J. Nucl. Med. Mol. Imaging 41, 1176–1189. [PubMed: 24425423]

Drzezga A, Barthel H, Minoshima S, Sabri O, 2014. Potential clinical applications of PET/MR imaging in neurodegenerative diseases. J. Nucl. Med 55, 47S–55S. [PubMed: 24819417]

Dukart J, Mueller K, Horstmann A, Barthel H, Moller HE, et al. , 2011. Combined evaluation of FDG-PET and MRI improves detection and differentiation of dementia. PLoS One 6, e18111. [PubMed: 21448435]

Fei B, Yang X, Nye JA, Aarsvold JN, Raghunath N, et al. , 2012. MRPET quantification tools: registration, segmentation, classification, and MR-based attenuation correction. Med. Phys. 39, 6443–6454. [PubMed: 23039679]

Fonov VS, Evans AC, McKinstry RC, Almli C, Collins D, 2009. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. Neuroimage 47, S102.

Grodzki DM, Jakob PM, Heismann B, 2012. Ultrashort echo time imaging using pointwise encoding time reduction with radial acquisition (PETRA). Magn. Reson. Med 67, 510–518. [PubMed: 21721039]

Heiss WD, 2009. The potential of PET/MR for brain imaging. Eur. J. Nucl. Med. Mol. Imaging 36, 105–112.

Hillner BE, Siegel BA, Liu D, Shields AF, Gareen IF, et al. , 2008. Impact of positron emission tomography/computed tomography and positron emission tomography (PET) alone on expected management of patients with cancer: initial results from the National Oncologic PET Registry. J. Clin. Oncol 26, 2155–2161. [PubMed: 18362365]

Hitz S, Habekost C, Furst S, Delso G, Forster S, et al. , 2014. Systematic comparison of the performance of integrated whole-body PET/MR imaging to conventional PET/ CT for 18F-FDG b.rain imaging in patients examined for suspected dementia. J. Nucl. Med 55, 923–931. [PubMed: 24833495]

Hofmann M, Steinke F, Scheel V, Charpiat G, Farquhar J, et al. , 2008. MRI-based attenuation correction for PET/MRI: a novel approach combining pattern recognition and atlas registration. J. Nucl. Med 49, 1875–1883. [PubMed: 18927326]

Izquierdo-Garcia D, Hansen AE, Forster S, Benoit D, Schachoff S, et al. , 2014. An SPM8-based approach for attenuation correction combining segmentation and nonrigid template formation: application to simultaneous PET/MR brain imaging. J. Nucl. Med 55, 1825–1830. [PubMed: 25278515]

Jakoby BW, Bercier Y, Watson CC, Bendriem B, Townsend DW, 2009. Performance characteristics of a new LSO PET/CT scanner with extended axial field- of-view and PSF reconstruction. IEEE Trans. Nucl. Sci 56, 633–639.

Johansson A, Karlsson M, Nyholm T, 2011. CT substitute derived from MRI sequences with ultrashort echo time. Med. Phys 38, 2708–2714. [PubMed: 21776807]

Juttukonda MR, Mersereau BG, Chen Y, Su Y, Rubin BG, et al. , 2015. MR-based attenuation correction for PET/MRI neurological studies with continuous-valued attenuation coefficients for bone through a conversion from R2* to CT-Hounsfield units. Neuroimage 112, 160–168. [PubMed: 25776213]

Keereman V, Fierens Y, Broux T, De Deene Y, Lonneux M, et al. , 2010. MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences. J. Nucl. Med 51, 812–818. [PubMed: 20439508]

Khateri P, Saligheh Rad H, Jafari AH, Fathi Kazerooni A, Akbarzadeh A, et al. , 2015. Generation of a four-class attenuation map for MRI-based attenuation correction of PET data in the head area using a novel combination of STE/Dixon-MRI and FCM clustering. Mol. Imaging Biol 17, 884–892. [PubMed: 25917750]

Koesters T, Friedman KP, Fenchel M, Zhan Y, Hermosillo G, et al. , 2016. Dixon sequence with superimposed model-based bone compartment provides highly accurate PET/MR attenuation correction of the brain. J. Nucl. Med 57, 918–924. [PubMed: 26837338]

Ladefoged CN, Hansen AE, Keller SH, Holm S, Law I, et al. , 2014. Impact of incorrect tissue classification in Dixon-based MR-AC: fat-water tissue inversion. EJNMMI Phys 1, 101. [PubMed: 26501459]

Ladefoged CN, Benoit D, Law I, Holm S, Kjaer A, et al. , 2015. Region specific optimization of continuous linear attenuation coefficients based on UTE (RESOLUTE): application to PET/MR brain imaging. Phys. Med. Biol 60, 8047–8065. [PubMed: 26422177]

Martinez-Möller A, Souvatzoglou M, Delso G, Bundschuh RA, Chefd'hotel C, et al. , 2009. Tissue classification as a potential approach for attenuation correction in whole-body PET/MRI: evaluation with PET/CT data. J. Nucl. Med 50, 520–526. [PubMed: 19289430]

Mehranian A, Zaidi H, 2015. Joint estimation of activity and attenuation in whole-body TOF PET/MRI using constrained Gaussian mixture models. IEEE Trans. Med. Imaging 34, 1808–1821. [PubMed: 25769148]

Mehranian A, Arabi H, Zaidi H, 2016. Quantitative analysis of MRI-guided attenuation correction techniques in time-of-flight brain PET/MRI. Neuroimage 130, 123–133. [PubMed: 26853602]

Merida I, Costes N, Heckemann RA, Drzezga A, Forster S, et al. , 2015. Evaluation of several multi-atlas methods for PSEUDO-CT generation in brain MRI-PET attenuation correction. IEEE. In: Proceedings of the 12th International Symposium on Biomedical Imaging (ISBI), New York, NY, USA. pp. 1431–1434.

Navalpakkam BK, Braun H, Kuwert T, Quick HH, 2013. Magnetic resonance-based attenuation correction for PET/MR hybrid imaging using continuous valued attenuation maps. Invest. Radiol 48, 323–332. [PubMed: 23442772]

Nuyts J, Dupont P, Stroobants S, Benninck R, Mortelmans L, et al. , 1999. Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms. IEEE Trans. Med. Imaging 18, 393–403. [PubMed: 10416801]

Paulus DH, Quick HH, Geppert C, Fenchel M, Zhan Y, et al. , 2015. Whole-body PET/MR imaging: quantitative evaluation of a novel model-based MR attenuation correction method including bone. J. Nucl. Med 56, 1061–1066. [PubMed: 26025957]

Poynton CB, Chen KT, Chonde DB, Izquierdo-Garcia D, Gollub RL, et al. , 2014. Probabilistic atlas-based segmentation of combined T1-weighted and DUTE MRI for calculation of head attenuation maps in integrated PET/MRI scanners. Am. J. Nucl. Med. Mol. Imaging 4, 160–171. [PubMed: 24753982]

Rezaei A, Defrise M, Nuyts J, 2014. ML-reconstruction for TOF-PET with simultaneous estimation of the attenuation factors. IEEE Trans. Med. Imaging 33, 1563–1572. [PubMed: 24760903]

Rezaei A, Defrise M, Bal G, Michel C, Conti M, et al. , 2012. Simultaneous reconstruction of activity and attenuation in time-of-flight PET. IEEE Trans. Med. Imaging 31, 2224–2233. [PubMed: 22899574]

Roy S, Wang WT, Carass A, Prince JL, Butman JA, et al. , 2014. PET attenuation correction using synthetic CT from ultrashort echo-time MR imaging. J. Nucl. Med 55, 2071–2077. [PubMed: 25413135]

Salomon A, Goedicke A, Schweizer B, Aach T, Schulz V, 2011. Simultaneous reconstruction of activity and attenuation for PET/MR. IEEE Trans. Med. Imaging 30, 804–813. [PubMed: 21118768]

Schreibmann E, Nye JA, Schuster DM, Martin DR, Votaw J, et al. , 2010. MR-based attenuation correction for hybrid PET-MR brain imaging systems using deformable image registration. Med. Phys 37, 2101–2109. [PubMed: 20527543]

Sekine T, Buck A, Delso G, Ter Voert EE, Huellner M, et al. , 2016a. Evaluation of atlas-based attenuation correction for integrated PET/MR in human brain: application of a head atlas and

comparison to true CT-based attenuation correction. J. Nucl. Med 57, 215–220. [PubMed: 26493207]

Sekine T, Burgos N, Warnock G, Huellner M, Buck A, et al. , 2016b. Multi-atlas-based attenuation correction for brain 18F-FDG PET imaging using a time-of-flight PET/MR scanner: comparison with clinical single-atlas- and CT-based attenuation correction. J. Nucl. Med 57, 1258–1264. [PubMed: 27013697]

Su KH, Hu L, Stehning C, Helle M, Qian P, et al. , 2015. Generation of brain pseudo-CTs using an undersampled, single-acquisition UTE-mDixon pulse sequence and unsupervised clustering. Med. Phys 42, 4974–4986. [PubMed: 26233223]

Torrado-Carvajal A, Herraiz JL, Alcain E, Montemayor AS, Garcia-Canamaque L, et al. , 2016. Fast patch-based pseudo-CT synthesis from T1-weighted MR images for PET/MR attenuation correction in brain studies. J. Nucl. Med 57, 136–143. [PubMed: 26493204]

Wiesinger F, Sacolick LI, Menini A, Kaushik SS, Ahn S, et al. , 2016. Zero TE MR bone imaging in the head. Magn. Reson. Med 75, 107–114. [PubMed: 25639956]

Wollenweber SD, Ambwani S, Delso G, Lonn AHR, Mullick R, et al. , 2013. Evaluation of an atlas-based PET head attenuation correction using PET/CT & MR patient data. IEEE Trans. Nucl. Sci 60, 3383–3390.

Yang Z, Choupan J, Sepehrband F, Reutens D, Crozier S, 2013. Tissue classification for PET/MRI attenuation correction using conditional random field and image fusion. Int. J. Mach. Learn. Comput 3, 87–92.

Zaidi H, Montandon ML, Slosman DO, 2003. Magnetic resonance imaging-guided attenuation and scatter corrections in three-dimensional brain positron emission tomography. Med. Phys 30, 937–948. [PubMed: 12773003]

Zaidi H, Ojha N, Morich M, Griesmer J, Hu Z, et al. , 2011. Design and performance evaluation of a whole-body Ingenuity TF PET-MRI system. Phys. Med. Biol 56, 3091–3106. [PubMed: 21508443]

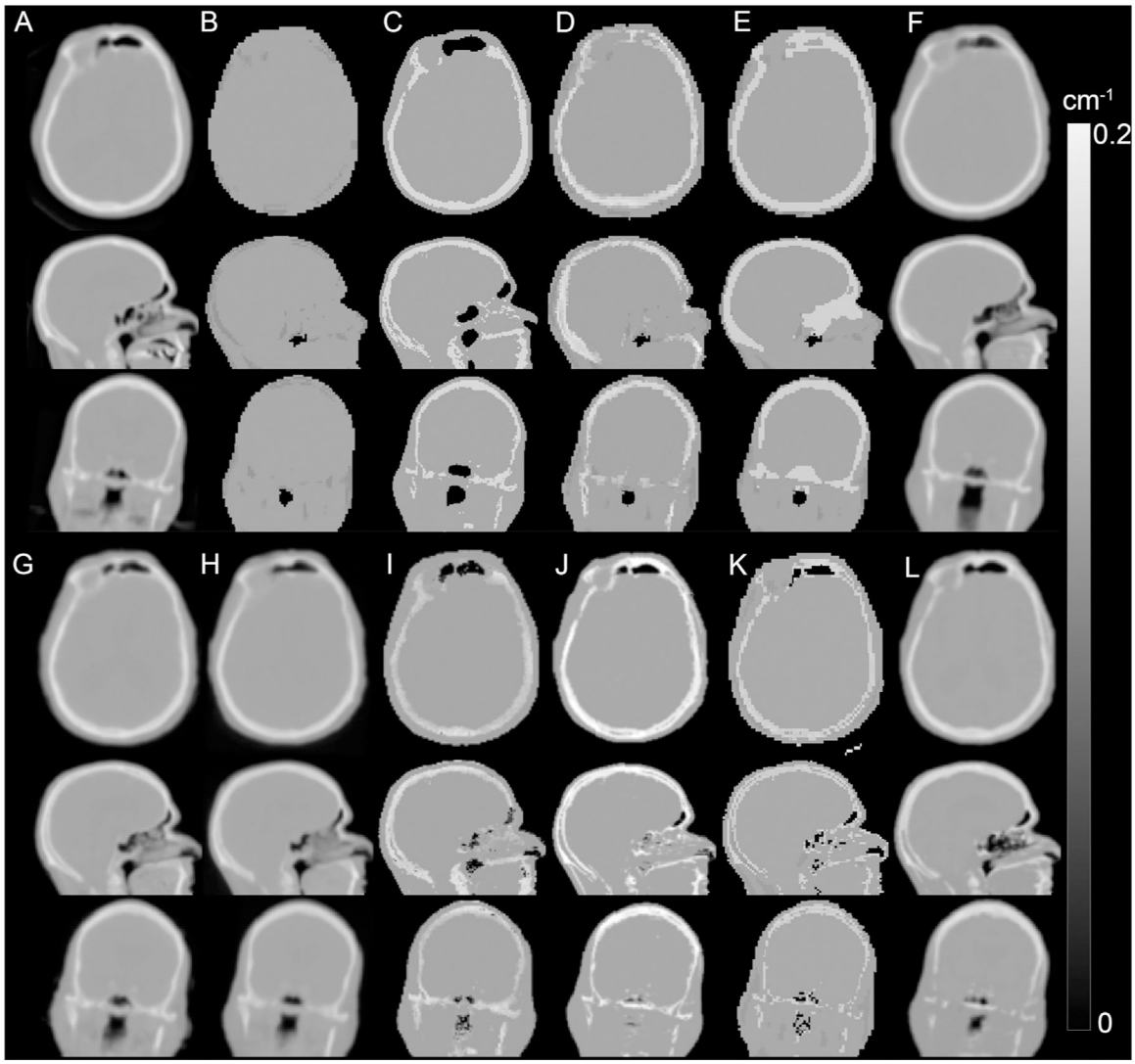


Fig. 1.

Attenuation correction images for a sample patient that minimizes the difference of the overall brain error to the median error across all methods. (A) CT, (B) Dixon, (C) UTE, (D) Segbone, (E) Ontario, (F) Boston, (G) UCL, (H) MaxProb, (I) MLAA, (J) Munich, (K) CAR-RiDR, (L) RESOLUTE.

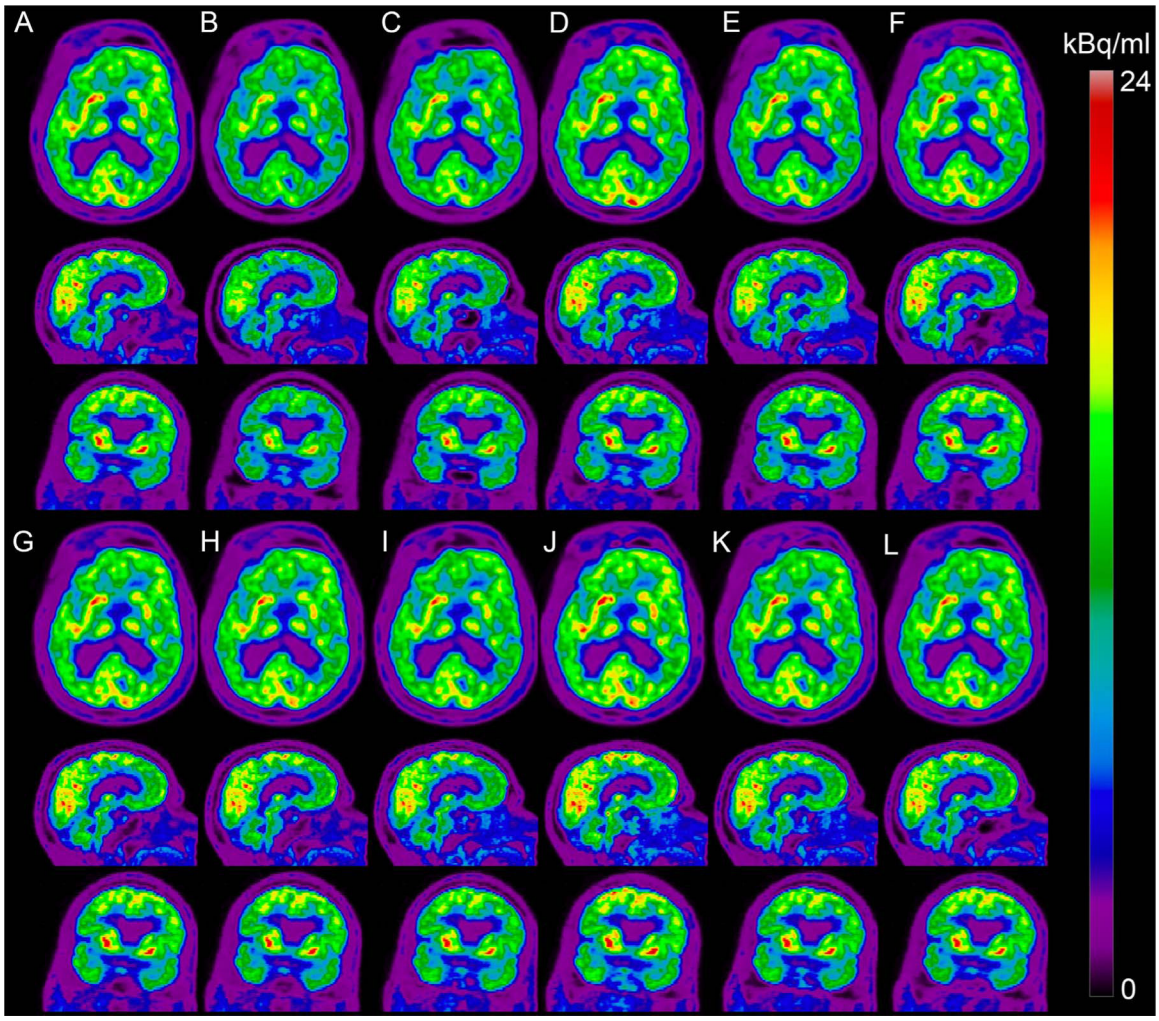
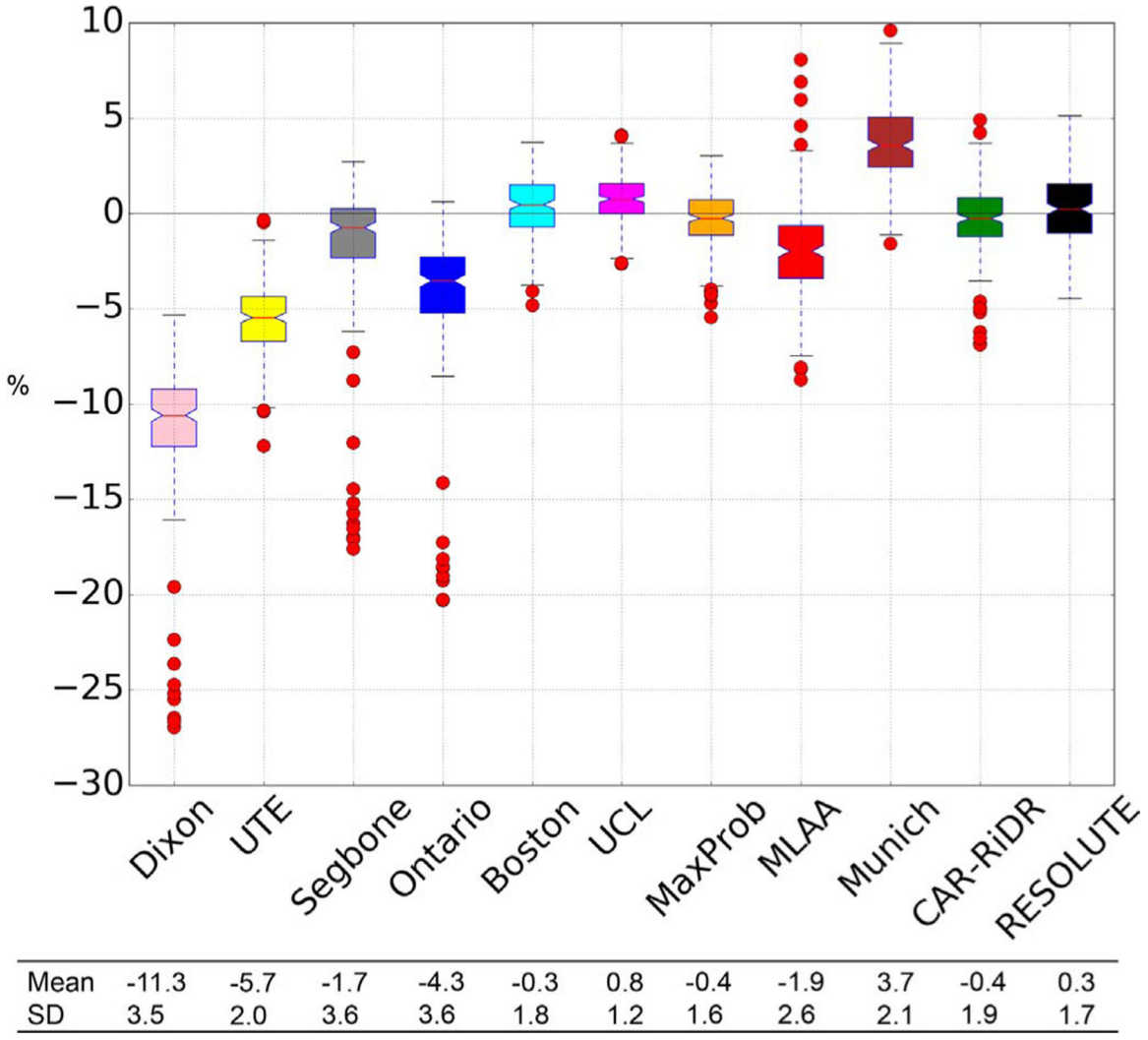


Fig. 2.

PET images for a sample patient that minimizes the difference of the overall brain error to the median error across all methods. (A) CT, (B) Dixon, (C) UTE, (D) Segbone, (E) Ontario, (F) Boston, (G) UCL, (H) MaxProb, (I) MLAA, (J) Munich, (K) CAR-RiDR, (L) RESOLUTE.



	Dixon	UTE	Segbone	Ontario	Boston	UCL	MaxProb	MLAA	Munich	CAR-RiDR	RESOLUTE
Mean	-11.3	-5.7	-1.7	-4.3	-0.3	0.8	-0.4	-1.9	3.7	-0.4	0.3
SD	3.5	2.0	3.6	3.6	1.8	1.2	1.6	2.6	2.1	1.9	1.7

Fig. 3.

Global performance of all methods for the FDG patients (n=201, including patients with fat/water tissue inversion). The median (red line), 25th and 75th percentiles (box), 1.5*IQR (whiskers), outliers (red dots), mean and SD are shown for each method. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

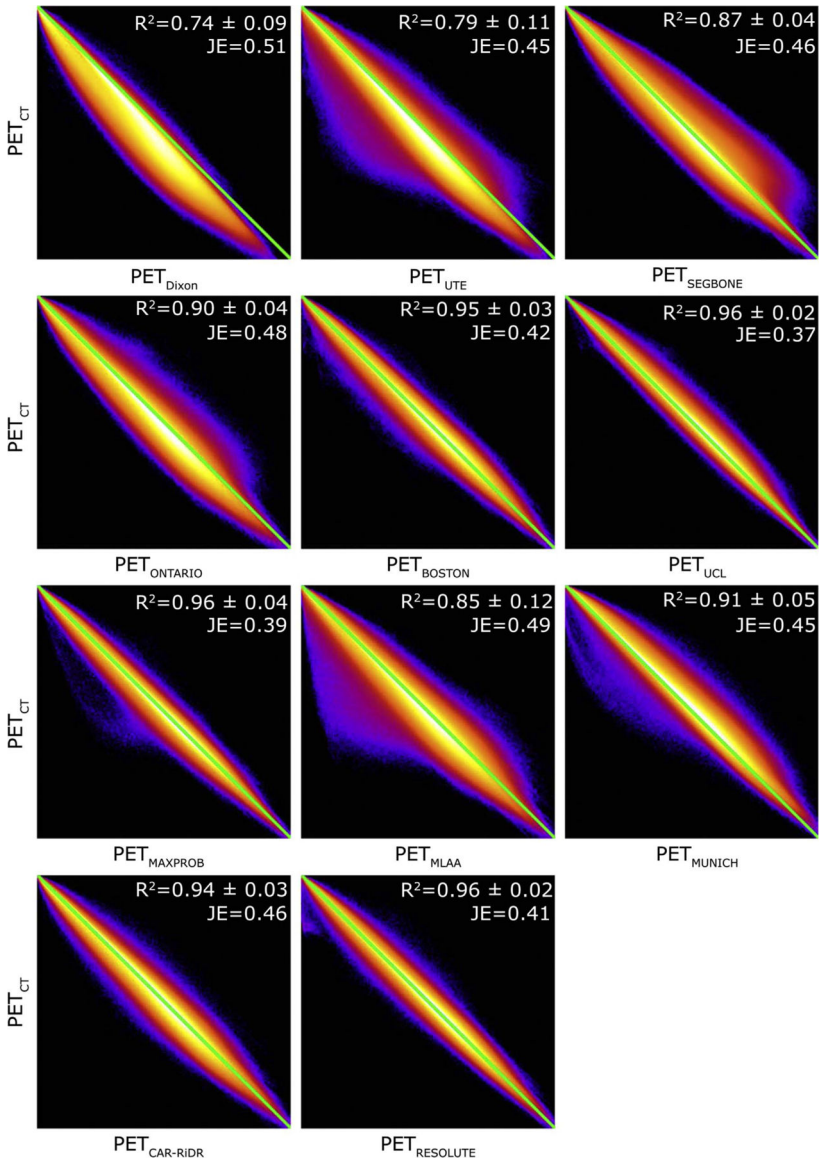


Fig. 4.

Summed joint histograms of PET activity within the brain mask for FDG (n=192, excluding patients with fat/water tissue inversion) for PET_{CT} versus each of the methods. The R^2 scores are average ± SD of the individual patients. The joint entropy (JE) is calculated for all patients.

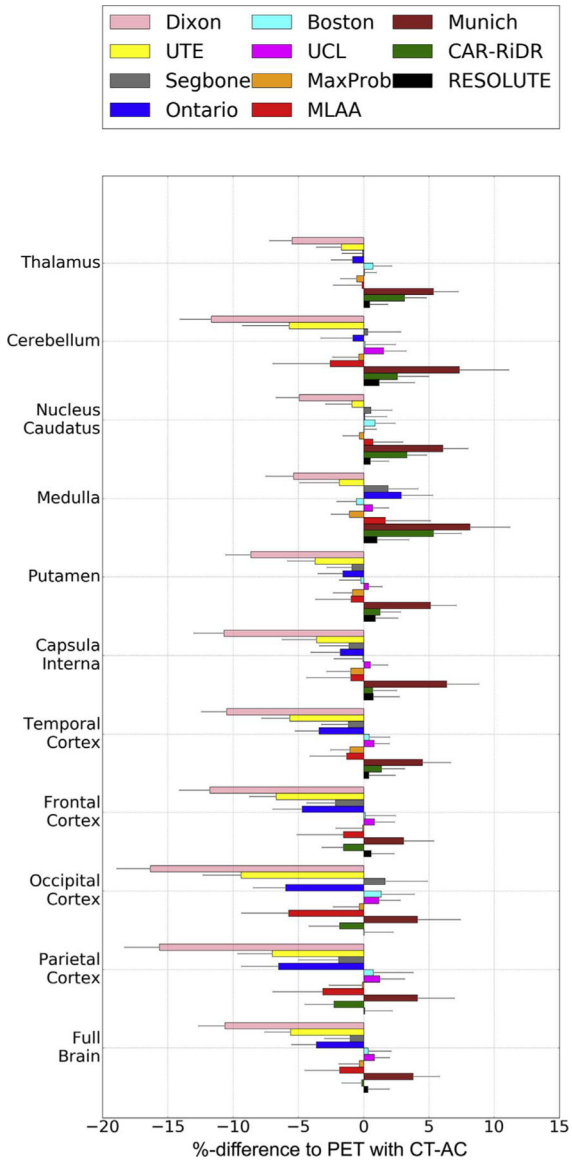


Fig. 5.

Global and regional ROI analysis across all FDG patients (n=192, excluding patients with fat/water tissue inversion). The gray lines indicate 1 SD.

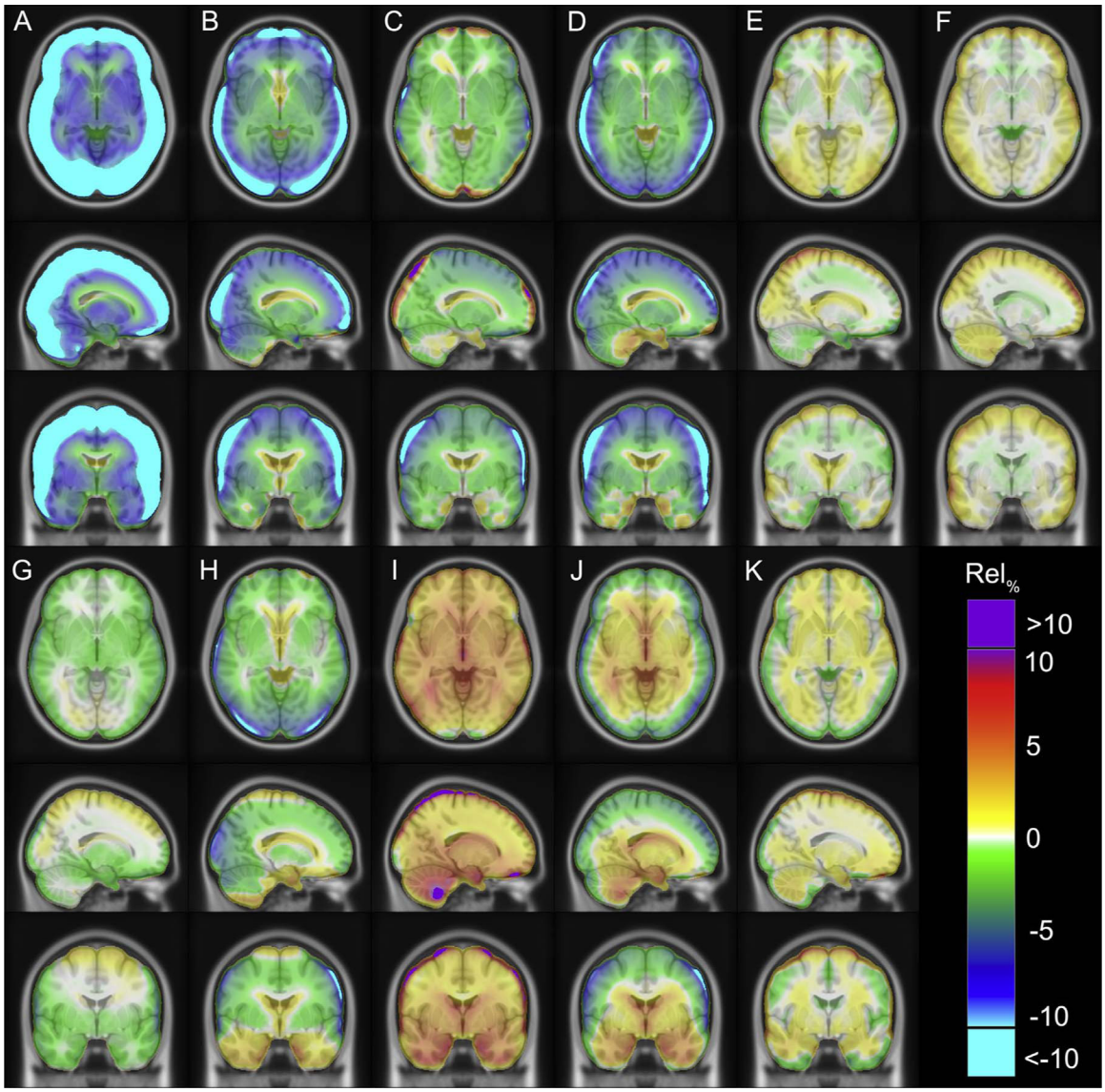


Fig. 6.

Averaged $Rel_{\%}$ images across all FDG patients (n=201, including patients with fat/water tissue inversion) for each method: (A) Dixon, (B) UTE, (C) Segbone, (D) Ontario, (E) Boston, (F) UCL, (G) MaxProb, (H) MLAA, (I) Munich, (J) CAR-RiDR, (K) RESOLUTE.

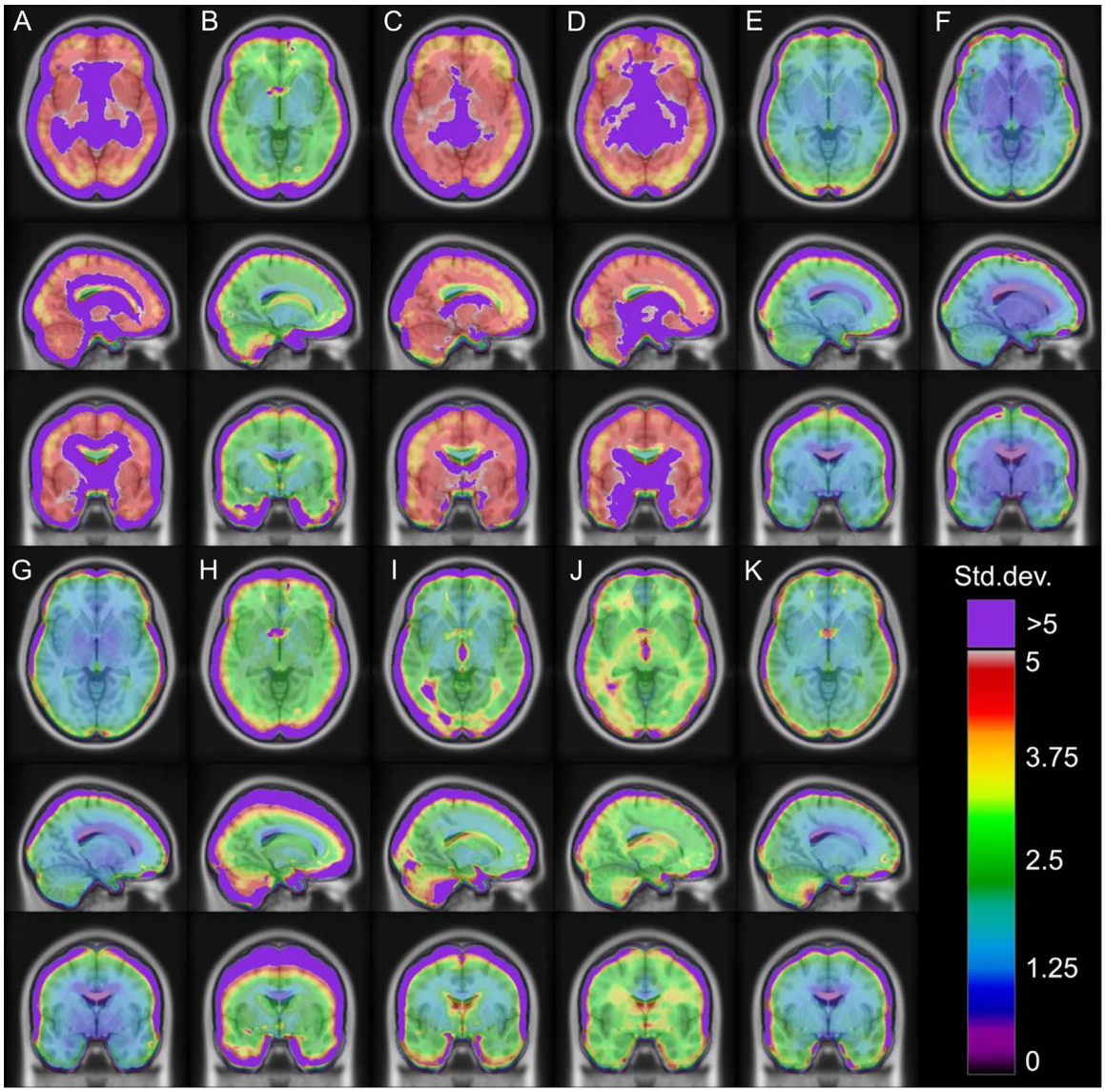


Fig. 7.

Standard deviation images across all FDG $Rel_{\%}$ images (n=201, including patients with fat/water tissue inversion) for each method: (A) Dixon, (B) UTE, (C) Segbone, (D) Ontario, (E) Boston, (F) UCL, (G) MaxProb, (H) MLAA, (I) Munich, (J) CAR-RiDR, (K) RESOLUTE.

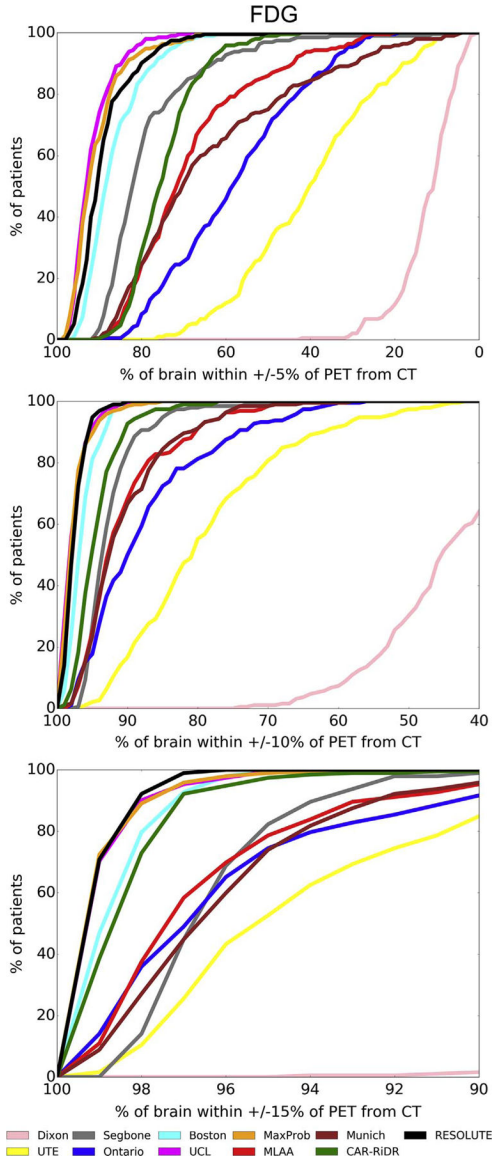


Fig. 8.

Outlier analysis for the FDG patients (n=192, excluding patients with fat/water tissue inversion). Note different scale of x-axis.

Table 1

Patient information.

Tracer	Originating center	N (fat/water inverted)	Male/ Female	Age (min-max) in years	Injected tracer (± SD) in MBq	Scan start p.i. (min-max) in min
[^{18}F]FDG	Rigshospitalet, Copenhagen Denmark	201 (9)	108/93	68 (23–96)	203 (± 20)	51 (24–134)
[^{11}C]PiB	Rigshospitalet, Copenhagen Denmark	48 (1)	2424	68 (39–85)	426 (± 75)	43 (37–100)
[^{18}F]florbetapir	Washington University, North Carolina, USA	88 (3)	21/67	71 (41–91)	353 (± 30)	0 n=75, or 52 (49–60) n=13

p.i.: post injection.

Table 2

Jaccard distance for overlap with bone ($\mu > 0.101$ cm^{-1}) segmented from MR-AC$_{CT}$ shown for each method and tracer. Results are shown for full head and further subdivided into top and bottom representing the area above and below the eyes. The best results are highlighted in bold. Note the Dixon method is left out due to there not being any bone present.

	FDG			**PiB**			**Florbetapir**		
	Full	**Top**	**Bottom**	**Full**	**Top**	**Bottom**	**Full**	**Top**	**Bottom**
UTE	0.56 ± 0.06	0.62 ± 0.10	0.51 ± 0.05	0.57 ± 0.06	0.65 ± 0.09	0.51 ± 0.04	0.32 ± 0.07	0.38 ± 0.12	0.27 ± 0.04
Segbone	0.54 ± 0.04	0.63 ± 0.07	0.48 ± 0.04	0.56 ± 0.03	0.65 ± 0.04	0.49 ± 0.03	N/A	N/A	N/A
Ontario	0.60 ± 0.04	0.74 ± 0.05	0.52 ± 0.05	0.60 ± 0.05	0.73 ± 0.07	0.52 ± 0.04	0.63 ± 0.05	0.75 ± 0.07	0.55 ± 0.04
Boston	0.74 ± 0.06	0.79 ± 0.08	0.69 ± 0.07	0.78 ± 0.04	0.83 ± 0.05	0.73 ± 0.04	0.78 ± 0.03	0.81 ± 0.06	0.75 ± 0.03
UCL	0.80 ± 0.04	0.83 ± 0.05	**0.77** ± 0.05	0.80 ± 0.04	0.83 ± 0.05	0.78 ± 0.04	**0.82** ± 0.03	**0.85** ± 0.04	**0.80** ± 0.03
MaxProb	**0.81** ± 0.04	**0.85** ± 0.05	**0.77** ± 0.04	**0.82** ± 0.02	0.84 ± 0.04	**0.79** ± 0.03	0.80 ± 0.03	0.83 ± 0.04	0.78 ± 0.03
MLAA	0.59 ± 0.07	0.67 ± 0.09	0.53 ± 0.06	0.51 ± 0.07	0.60 ± 0.09	0.44 ± 0.06	0.36 ± 0.09	0.37 ± 0.15	0.35 ± 0.05
Munich	0.72 ± 0.04	**0.85** ± 0.05	0.63 ± 0.05	0.73 ± 0.04	**0.86** ± 0.05	0.64 ± 0.04	0.72 ± 0.03	**0.85** ± 0.04	0.63 ± 0.04
CAR-RiDR	0.50 ± 0.07	0.62 ± 0.09	0.40 ± 0.06	0.50 ± 0.06	0.62 ± 0.08	0.41 ± 0.06	0.46 ± 0.06	0.60 ± 0.09	0.36 ± 0.05
RESOLUTE	0.68 ± 0.04	**0.85** ± 0.04	0.56 ± 0.04	0.69 ± 0.03	0.84 ± 0.04	0.58 ± 0.04	0.68 ± 0.03	**0.85** ± 0.03	0.54 ± 0.04

Neuroimage. Author manuscript; available in PMC 2019 October 29.

Table 3

Analysis of worst three outlier patients based on amount of brain with errors greater than ± 10% (excluding patients with fat/water inversion). The worst outlier is also shown in Supplemental Fig. 15.

		Subject with highest % of brain > ± 10%	2nd highest % of brain > ± 10%	3rd highest % of brain > ± 10%
Dixon	*%brain (category)* *Rel %* *(region)*	83% (A) −30% (cortical)	82% (A) −20% (cortical)	81% (A) −25% (cortical)
UTE	*%brain (category)* *Rel %* *(region)*	58% (A) −15% (cortical)	55% (A) −15% (cortical)	54% (A) −15% (cortical)
Segbone	*%brain (category)* *Rel %* *(region)*	35% (A) −15% (cortical)	28% (B) +20% (cortical)	26% (A) −15% (cortical)
Ontario	*% of brain > ± 10%* *Rel %* *(region)*	44% (A) −20% (cortical)	39% (A) −20% (cortical)	39% (A) −20% (cortical)
Boston	*% of brain > ± 10%* *Rel %* *(region)*	10% (B) 20% (cortical)	10% (B) 10% (cortical)	9% (A) −15% (cortical)
UCL	*% of brain > ± 10%* *Rel %* *(region)*	10% (B) 30% (cortical)	9% (B) 15% (cortical)	9% (B) 30% (cortical)
MaxProb	*% of brain > ± 10%* *Rel %* *(region)*	14% (D) −25% (r brain)	12% (A) −20% (cortical)	10% (A) −20% (cortical)
MLAA	*% of brain > ± 10%* *Rel %* *(region)*	34% (A) −15% (cortical)	34% (A) −15% (cortical)	30% (A) −15% (cortical)

		Subject with highest % of brain > ± 10%	2nd highest % of brain > ± 10%	3rd highest % of brain > ± 10%
Munich		41% (C)	34% (C)	31% (C)
	% of brain > ± 10%	25% (cerebellum)	20% (cerebellum)	15% (cerebellum)
	Rel %			
	(region)			
CAR-RiDR		23% (A)	23% (C)	18% (A)
	% of brain > ± 10%	−20% (cortical)	20% (cerebellum)	−20% (cortical)
	Rel %			
	(region)			
RESOLUTE		10% (B/C)	9% (B/C)	7% (A)
	% of brain > ± 10%	10% (cerebellum)	10% (cerebellum)	−10% (cortical)
	Rel %			
	(region)			

Category explanation: (A) Bone density or bone width underestimated, (B) bone density or bone width overestimated, (C) incorrect air/bone interface segmentation, and (D) template/atlas registration errors.

Table 4

Requirements of the proposed methods.

	T1 MPRAGE (acq. time: 300 s)	Dixon (acq time: 19 s)	UTE (acq. time: 100 s)	Prior data	Requirements (license)	Runtime (approx.)
Segbone		✓		Needs single atlas	Proprietary	1–2 m
Ontario	✓	✓			SPM (GNU)	30 m
Boston	✓			Needs single atlas	FreeSurfer, SPM (GNU)	30 m
UCL	✓			Needs multiple atlases	NiftyReg (BSD)	1 h
MaxProb	✓			Needs multiple atlases	NiftyReg (BSD)	2 h
MLAA	✓		✓	Optimization needed per tracer type		1 h
Munich			✓	Calibration to specific MR sequence required		1 m
CAR-RiDR		✓	✓	Calibration to specific MR sequence required		1 m
RESOLUTE			✓	Calibration to specific MR sequence required	MINC (BSD), ANTs (BSD)	15 m