Voxel-based comparison of state-of-the-art reconstruction algorithms for ^{18}F-FDG PET brain imaging using simulated and clinical data

K. Vunckx[a,b,*], P. Dupont[b,c], K. Goffin[a,b,d],
W. Van Paesschen[b,e,f], K. Van Laere[a,b,d], J. Nuyts[a,b]

[a] KU Leuven - University of Leuven, Department of Imaging & Pathology, Nuclear Medicine & Molecular Imaging, Herestraat 49, B-3000 Leuven, Belgium.

[b] KU Leuven - University of Leuven, University Hospitals Leuven, Medical Imaging Research Center (MIRC), Herestraat 49, B-3000 Leuven, Belgium.

[c] KU Leuven - University of Leuven, Department of Neurosciences, Lab. for Cognitive Neurology, Herestraat 49, B-3000 Leuven, Belgium.

[d] University Hospitals Leuven, Department of Nuclear Medicine, Herestraat 49, B-3000 Leuven, Belgium

[e] KU Leuven - University of Leuven, Department of Neurosciences, Lab. for Epilepsy Research, Herestraat 49, B-3000 Leuven, Belgium.

[f] University Hospitals Leuven, Department of Neurology, Herestraat 49, B-3000 Leuven, Belgium.

*** Corresponding author:**
Kathleen Vunckx, PhD
Kathleen.Vunckx@uz.kuleuven.be

Abstract
The resolution of a PET scanner (2.5-4.5 mm for brain imaging) is similar to the thickness of the cortex in the (human) brain (2.5 mm on average), hampering accurate activity distribution reconstruction. Many techniques to compensate for the limited resolution during or post-reconstruction have been proposed in the past and have been shown to improve the quantitative accuracy. In this study, state-of-the-art reconstruction techniques are compared on a voxel-basis for quantification accuracy and group analysis using both simulated and measured data of healthy volunteers and patients with epilepsy. **Methods:** Maximum a posteriori (MAP) reconstructions using either a segmentation-based or a segmentation-less anatomical prior were compared to maximum likelihood expectation maximization (MLEM) reconstruction with resolution recovery. As anatomical information, a spatially aligned 3-D T1-weighted magnetic resonance image was used. Firstly, the algorithms were compared using normal brain images to detect systematic bias with respect to the true activity distribution, as well as systematic differences between two methods. Secondly, it was verified whether the algorithms yielded similar results in a group comparison study. **Results:** Significant differences were observed between the reconstructed and the true activity, with the largest errors when using (post-smoothed) MLEM. Only 5-10 % underestimation in cortical gray matter voxel activity was found for both MAP reconstructions. Higher errors were observed at GM edges. MAP with the segmentation-based prior also resulted in a significant bias in the subcortical regions due to segmentation inaccuracies, while MAP with the anatomical prior which does not need segmentation did not. Significant differences in reconstructed activity were also found between the algorithms at similar locations (mainly in gray matter edge voxels and in cerebrospinal fluid voxels) in the simulated as well as in the clinical data sets. Nevertheless, when comparing two groups, very similar regions of significant hypometabolism were detected by all algorithms. **Conclusion:** Including anatomical a priori information during reconstruction in combination with resolution modeling yielded accurate gray matter activity estimates, and a significant improvement in quantification accuracy was found when compared to post-smoothed MLEM reconstruction with resolution modeling. AsymBowsher provided more accurate subcortical GM activity estimates. It is also reassuring that the differences found between the algorithms did not hamper the detection of hypometabolic regions in the gray matter when performing a voxel-based group comparison. Nevertheless, the size of the detected clusters differed. More elaborated and application-specific studies are required to decide which algorithm is best for a group analysis.

Keywords
Iterative reconstruction; Anatomical prior; Voxel-based comparison; ^{18}F-FDG PET; Brain imaging

1. Introduction

Fluorine-18 fluoro-2-deoxyglucose (^{18}F-FDG) positron emission tomography (PET) imaging of the brain is a well-established way to detect changes in neuronal activity induced by a variety of diseases through the alteration in glucose consumption *(1)*. Both filtered backprojection and iterative reconstruction algorithms are currently being used in clinical practice. However, the limited spatial resolution of PET (2.5-4.5 mm for brain imaging *(2,3)*) in combination with the small cortical thickness (average thickness of 2.5 mm *(4)*) compromises accurate quantification of the tracer uptake in the brain *(5)*.

Two approaches have been recommended to improve the quantitative accuracy. The first one is to correct for the partial volume effect after reconstruction, by estimating the point spread function and correcting for its effect on the reconstructed activity, given spatially aligned tissue segmentation information. Different partial volume correction methods for brain ^{18}F-FDG PET imaging have been compared in *(6)*. The second approach is to model the spatial resolution during image reconstruction to sharpen the image ('resolution recovery' or 'resolution modeling'). Even in high-resolution PET systems this can reduce the partial volume effect and improve the quantitative accuracy *(7)*, but it is worth noting that resolution modeling may induce some unwanted secondary effects, such as an increase in inter-voxel correlations and Gibbs artifacts. Hence, an improvement in one figure of merit does not imply an improvement of all possible figures of merit *(8)*. To suppress noise and Gibbs artifacts, resolution modeling is best combined with the use of anatomical prior information during, e.g., a maximum a posteriori (MAP) image reconstruction *(9)*. In addition, detection accuracy of hypometabolic regions can be significantly improved in this way *(10,11)*.

Unfortunately, segmentation of, e.g., structural magnetic resonance (MR) images does not yield perfect brain tissue classification. Especially subcortical structures are very hard to segment, although increasingly accurate segmentation methods are being developed *(12)*. Partial volume correction (PVC) techniques and anatomical priors that heavily rely on segmentations will automatically translate the segmentation errors into quantification errors and reconstruction artifacts. Therefore, segmentation-free anatomical priors have been proposed as well in the past *(13-15)*.

Because noise suppression through the use of anatomical information was found to be superior with MAP reconstruction compared to with post-processed maximum likelihood reconstruction *(16)*, we concentrated on the evaluation of reconstruction-based partial volume correction algorithms in this work. Furthermore, diagnosis based on ^{18}F-FDG PET brain images is often performed by looking for locally deviating patterns or small hypometabolic regions. Therefore, we focused on voxel-based rather than region of interest-based image quality evaluation. Also in group analyses, voxel-based statistical parametric mapping techniques are typically used to highlight differences in tracer uptake.

The aim of the present study was to investigate where systematic differences are to be expected when using MLEM with resolution recovery, MAP reconstruction with a segmentation-based prior or MAP with a segmentation-free anatomical prior (three state-of-the-art reconstruction algorithms), and whether the choice of the algorithm influences the outcome of a group analysis. Through the use of a simulation study, results can also be compared to the true activity distribution. The evaluation on clinical data sets enabled us to verify whether similar results are obtained in a realistic setting.

2. Material and methods

2.1. Reconstruction algorithms

2.1.1. MLEM with resolution modeling

Some of the resolution lost due to blur during the detection process can be recovered by modeling the resolution of the detector during iterative reconstruction. In these iterative algorithms, the currently estimated activity distribution $\Lambda^{old} = [\lambda_1 \dots \lambda_J]$ (with J the number of reconstruction image voxels) is gradually improved based on a comparison between the measurement $Q = [q_1 \dots q_I]$ (with I the number of detector pixels) and the forward projected Λ^{old}, which can be written as $\sum_k c_{ik}\lambda_k^{old}$ (with c_{ik} the probability of detecting a photon pair at detector i if it was emitted at image voxel j).

The most commonly used iterative reconstruction algorithm, MLEM, maximizes the log-likelihood L *(17)*. The additive formulation of the algorithm is given by

$$\lambda_j = \lambda_j^{old} + \lambda_j^{old} \frac{\sum_i c_{ij} \left\langle \frac{q_i - \sum_k c_{ik}\lambda_k^{old}}{\sum_k c_{ik}\lambda_k^{old}} \right\rangle}{\sum_i c_{ij}} \left[= \lambda_j^{old} + \frac{\frac{\partial L}{\partial \lambda_j}}{\frac{\sum_i c_{ij}}{\lambda_j^{old}}} \right]$$

The expression between square brackets is provided for comparison to the MAP algorithm, described in the next paragraphs. The resolution model can be included in c_{ik}, which also contains information about the sensitivity of the camera and the patient attenuation.

2.1.2. MAP with segmentation-based anatomical prior

If a priori information about the activity distribution is available, it can be included during reconstruction using a maximum a posteriori (MAP) algorithm. In this study, a gradient ascent optimization method is used *(18)*, which can be formulated as follows:

$$\lambda_j = \lambda_j^{old} + \left(\frac{\partial L}{\partial \lambda_j} + \beta \frac{\partial M}{\partial \lambda_j} \right) \Big/ \left(\frac{\sum_i c_{ij}}{\lambda_j^{old}} - 2\beta \frac{\partial^2 M}{\partial \lambda_j^2} \right)$$

M is the log-prior function, which encourages that the reconstructed activity distribution is in agreement with the a priori information, which can be based on the reconstructed image itself or, e.g., on a spatially aligned structural MR (anatomical prior). The choice of β determines the strength of the regularization provided by the prior.

The MAP with segmentation-based anatomical prior used in this work is the A-MAP algorithm proposed in *(19)*. It intrinsically smooths the reconstructed activity within specific tissue classes, but not across the tissue boundaries. Furthermore, it copes with the fact that some voxels consist of multiple tissues.

The algorithm has been slightly modified to enhance the reconstruction accuracy and the convergence speed. Fuzzy gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) segmentation images are assumed to be available. From these images, a label image is generated, where all voxels with at least 10 % GM are labeled as GM. All other voxels with at least 10 % WM get label WM. All other voxels with at least 50 % CSF are denoted as CSF and the rest get label OTHER. At every iteration of the MAP algorithm, the current estimate Λ^{old} is converted into a full-class activity concentration image, such that each voxel – even if it consists of multiple tissues – represents only the activity concentration inside the tissue corresponding to the label that was assigned to that voxel. This is done by subtracting the estimated contributions of the other tissue classes from the activity in the voxel and converting the remaining activity to a 100 % contribution. Next, the same kind of prior is applied to all voxels with the same label (the relative difference prior *(18)* to GM, WM and OTHER and an absolute intensity prior with mean zero to CSF). The effect of the prior is computed in every

voxel and is weighted by the fraction of the labeled tissue in that voxel, such that more confidence is given to single-tissue voxels than to multiple-tissue voxels. This method will further be referred to as *A-MAP*.

2.1.3. **MAP with segmentation-less anatomical prior**

A promising anatomical prior, which does not rely on segmentations, is the Bowsher prior *(13)*. It is based on the assumption that neighboring voxels that are similar in the anatomical image will most probably also have a similar activity value in the PET image. Practically, a predefined neighborhood of voxels is investigated around each voxel. For each voxel *j*, the *N* voxels in this neighborhood with most similar intensity in the anatomical image compared to that of voxel *j* are determined and similarity between the reconstructed PET activity in each of these voxels and the activity in voxel *j* is encouraged. In a MAP reconstruction algorithm, this means that a (log-)prior function, e.g. the relative difference prior, is only applied to these sets of voxels indicated by the anatomical image.

Recently, we proposed a heuristic modification to this algorithm *(20)*, which avoids that – due to the symmetrical behavior of the prior – too much weight is given to the prior at edge voxels. This prior, referred to as the asymmetrical Bowsher (*AsymBowsher*) prior, yields clearly sharper and quantitatively more accurate images compared to the original Bowsher prior *(20)*. Therefore, the AsymBowsher algorithm was selected for the current study.

2.1.4. **Implementation details**

All reconstruction algorithms were implemented in-house using IDL (ITT Visual Information Solutions) and C programming software. The scanner modeled was the ECAT EXACT HR+ PET scanner (Siemens, Knoxville, TN, USA). Attenuation, scatter, detector resolution and decay were corrected for during reconstruction. Randoms were not modeled during measurement simulation, and the real measurements were precorrected for randoms. A spatially invariant Gaussian was used to model the detector resolution in image space during reconstruction. By estimating the full-width at half-maximum (FWHM) of 2D (non-isotropic) Gaussians to simulated and acquired point source measurements (at various spatial locations), we found that PET-SORTEO *(21)*, which was used to generate our simulated PET scans (see section 2.2.2.), models a spatially variant resolution that is similar but slightly different from the one we measured in our HR+ scanner (see Fig. 3 in *(9)*). Therefore, different spatially invariant Gaussians were used to model the resolution of the simulated and the real scanner. The transaxial and axial full-width at half-maximum (FWHM) of the Gaussian were set to 4.0 mm and 5.45 mm, respectively for simulations and 4.5 mm and 4.5 mm, respectively for real measurements (in good agreement with values found in *(3)*). These values were obtained by averaging the spatially invariant resolution in the central part of the field of view, which is used for brain imaging.

The images were reconstructed with 2.25 mm x 2.25 mm x 2.425 mm voxels. To accelerate convergence, ordered subsets *(22)* were used for all algorithms, with a total number of iterations corresponding to 250 MLEM iterations. The reconstruction parameters were optimized by minimizing the figure of merit (FOM) suggested in *(9)*. This FOM encourages a high quantitative accuracy in GM voxels, both in a normal brain ($\mathrm{FOM_{GM,N}}$) and in a brain with lesions $\mathrm{FOM_{GM,L}}$, as well as an accurate signal in the lesions $l = \{1, \ldots, 20\}$ ($\mathrm{FOM_S}$):

$$\mathrm{FOM} = \sqrt{\frac{\mathrm{FOM_{GM,N}^2} + \mathrm{FOM_{GM,L}^2}}{4} + \frac{\mathrm{FOM_S^2}}{2}}$$

with $\mathrm{FOM}_{\mathrm{GM,X}} = \sqrt{\frac{1}{n_{\mathrm{GM80}}}\sum_{j\in \mathrm{GM80}}\left(\frac{\lambda_j^{\mathrm{X}}-\lambda_j^{\mathrm{X,true}}}{\text{GM activity}}\right)^2}$ and $\mathrm{FOM}_{\mathrm{S}} = \sqrt{\frac{1}{20}\sum_{l=1}^{20}\left(\frac{(\bar{\lambda}_l^{\mathrm{L}}-\bar{\lambda}_l^{\mathrm{N}})-(\bar{\lambda}_l^{\mathrm{L,true}}-\bar{\lambda}_l^{\mathrm{N,true}})}{\text{GM activity}}\right)^2}$, and X = {N,L}. GM80 denotes all GM voxels with at least 80 % GM tissue fraction, n_{GM80} is the number of voxels in GM80, "GM activity" is the true GM activity concentration, $\bar{\lambda}_l^{\mathrm{L}}$ and $\bar{\lambda}_l^{\mathrm{N}}$ indicate the mean reconstructed value in the lesion volume of interest (VOI) l in the brain with hypointense lesions L and in the normal brain N, respectively, and the superscript $^{\mathrm{true}}$ refers to the ground truth activity distribution image (resampled to PET voxel size). Lesion VOIs were defined in PET space and restricted to those voxels from GM80 that fall at least for 80 % inside the high resolution lesion boundary. For MLEM, the isotropic Gaussian post-smooth filter had an optimal full-width at half-maximum of 4 mm. For A-MAP, the prior weights of the GM, WM, CSF and OTHER prior were set to gradually increase to 200, 750, 125 and 30, respectively. For the AsymBowsher MAP reconstruction, a relative difference prior was applied to the 13 most similar neighbors out of each spherical 5x5x5 neighborhood (i.e. 80 neighbors) with a prior weight of 10.

For the measured data, motion artifacts were reduced by applying a multiple acquisition frame correction method *(23)*, where the frame duration was fixed to 5 minutes in advance, because no motion tracking had been performed during acquisition. To correct for emission-transmission misalignment, a 2D filtered backprojection reconstructed attenuation map was registered to a fast non-attenuation corrected MLEM reconstruction of each frame. Next, the sinograms were reconstructed using MLEM with resolution modeling (and all corrections). Scatter was estimated using software provided by Siemens *(24)*. The high-resolution structural MR image was registered to the final MLEM reconstruction of each frame, such that it could be used as an anatomical prior for the MAP reconstructions. In the end, since all frames had already been co-registered with the MR, all frames that did not show obvious intra-frame motion were transformed to the original MR space (while keeping the PET voxel size) and averaged. In that way, an average motion-corrected reconstruction image was obtained. The same inter-frame transformations were used for all reconstruction methods. All spatial alignments were performed using in-house developed rigid registration software using normalized mutual information as the cost function.

2.2. **Simulation studies**

2.2.1. **Phantom creation**

For a recent study, where we compared three MR-based anatomical priors for quantitative PET brain imaging *(9)*, two software PET brain phantoms were generated. The phantoms were created based on an anatomical model of a normal brain, which consists of ten fuzzy tissue class segmentation images, available at the BrainWeb website *(25,26)*. The same activity concentrations as in *(9,10)* were assigned to the different tissue classes to generate a ground truth normal activity distribution (12500 Bq/ml in GM, 3125 Bq/ml in WM, 0 Bq/ml in air, CSF and bone, and 1000 Bq/ml in all other tissues; values based on typical clinical ^{18}F-FDG PET images and 4:1 GM:WM ratio (see *(27)*)). In addition, an attenuation map was created. In the second phantom, 20 small to medium-sized hypometabolic focal lesions (1.51 to 3.11 cm^3, 25 % reduction in tracer uptake) were inserted in the GM, to mimic a patient with, e.g., multifocal vascular disease or cortical dysplastic lesions *(28,29)*. On the BrainWeb website, also T1-weighted MR images simulated based on these tissue class images are available, such that we have a perfectly matching MR with perfect segmentation information for our study. In addition, GM, WM and CSF segmentation images were also computed using the Statistical Parametric Mapping (SPM8) software package (Wellcome Trust Centre for Neuroimaging (London, UK),

http://www.fil.ion.ucl.ac.uk/spm) in combination with Matlab R2011a (The MathWorks, Inc., Natick, MA). The noiseless MR was used for our simulation experiment and the focal lesions were not inserted, and thus invisible in the MR image.

2.2.2. **PET measurement simulations**

To be predictive for how well the algorithms will reconstruct clinical brain data, a PET Monte Carlo simulator, PET-SORTEO *(21)*, was used to generate 60 sinograms which realistically represent five-minute ^{18}F-FDG PET scans on a Siemens ECAT EXACT HR+ PET scanner. Among other effects, the spatially variant resolution of the scanner was modeled during simulation. A realistic, noiseless scatter model was also generated from simulated SORTEO data and used to add Poisson corrupted scatter to the sinograms (see *(9)* for more details).

2.3. **Human studies**

To verify whether the outcome of the simulation studies is predictive for what can be expected in clinical practice, the studies were repeated on two data sets of ^{18}F-FDG PET brain scans, i.e. on data from a group of 20 healthy volunteers and on data from a set of 16 patients diagnosed with mesial temporal lobe epilepsy with hippocampal sclerosis (MTLE-HS). In the patient brains, the metabolism is known to be locally reduced in specific brain regions, e.g., hippocampus and locations in the lateral temporal and frontal lobe at the side of the epileptogenic focus *(30)*.

2.3.1. **Healthy volunteers**

A brain ^{18}F-FDG PET and a high-resolution 3D T1-weighted Magnetization Prepared Rapid Acquisition Gradient Echo (3D-MPRAGE) MR scan of twenty neurologically normal volunteers (9 male/11 female, median age: 29 (range 21-49)) were available from a previous study approved by the Local Ethics Committee *(30)*. Informed consent had been obtained from all participants before the investigations. The raw PET data were reconstructed with the different algorithms and used for a retrospective analysis.

2.3.2. **Patients with MTLE-HS**

A group of 18 patients with well-documented MTLE-HS (12 left- and 6 right-sided) was also investigated retrospectively. Only patients were selected that underwent both a brain ^{18}F-FDG PET and a high-resolution 3D-MPRAGE scan between October 2001 and November 2008 as part of their diagnostic program, prior to resection of the epileptogenic zone.
Two patients with left-sided MTLE-HS were excluded from the current study. The first one was omitted, because she had a strongly enlarged ventricle, making accurate segmentation and spatial normalization impossible. The second one was excluded, because he had a frontal lesion, which would influence the outcome of the voxel-based analysis. The final cohort consisted of 7 men and 9 women (median age: 40 (range 15-58)).

2.3.3. **PET acquisition**

Both groups were scanned on the same Siemens ECAT EXACT HR+ PET scanner under standard conditions with low ambient noise and in a dimmed room. For attenuation correction purposes, a transmission scan using ^{68}Ge rod sources was acquired in 2-D mode.
The healthy controls were scanned dynamically for one hour starting immediately after the injection of approximately 150 MBq of ^{18}F-FDG to study the kinetics of ^{18}F-FDG in the brain. In this work, only the six 5-minute frames between 30 and 60 min post injection were reconstructed and analyzed to enable comparison to the reconstructed images of the patient scans.

The patients with MTLE-HS were scanned thirty minutes after intravenous injection of approximately 150 MBq of ^{18}F-FDG. Six 5-minute frames were acquired in 3-D mode, such that in case of head motion, a subset of frames could be used. For the patients, interictal state was confirmed through continuous EEG-monitoring during the first 30 minutes after injection.

2.4. **Comparison algorithms**

2.4.1. **Bias with respect to true activity and differences between algorithms**

In a first step, the reconstruction algorithms were compared by verifying in which voxels statistically different activity values were obtained with respect to the true activity. This comparison was performed on the reconstructions of the 60 simulated normal brain scans. Note that significant differences can also show up in case of subtle, but very systematic differences, which are not necessarily relevant. Therefore, the mean percentage difference between the reconstructed activity and the true activity is also calculated for each algorithm.

In a second step, the different reconstruction algorithms were compared to each other, first using the 60 simulated normal brains and later in a clinical and more challenging setting using the 20 measured brains of the healthy controls. Because a voxel-based analysis on the clinical data requires spatial normalization and smoothing to cope with inter-patient functional variability, the effect of these steps was first analyzed on the simulated data.

2.4.2. **Detection of group differences**

A specific reconstruction algorithm is expected to yield similar reconstruction accuracy when processing patient data or data of healthy controls, such that only differences should be found where the two groups really differ. Some algorithms might, however, yield higher or lower contrast between normal metabolism and hypometabolism compared to other, or smooth away low contrast lesions. Therefore, we evaluated how the reconstruction algorithms performed in a voxel-based group analysis. Firstly, a group analysis on the reconstructed simulated data was performed, one for each of the reconstruction algorithms. In the group analysis, the group of 60 reconstructed normal brains was compared voxel-wise to the group of 60 reconstructed brains with hypometabolic lesions. Because of the high noise in the MLEM images, the analysis was also run on post-smoothed MLEM images (Gaussian filter with 4 mm FWHM). Secondly, the analyses were repeated on the spatially normalized and smoothed images to predict the outcome of a group analysis on measured human data. Finally, the images of the 20 healthy controls were compared to those of the 16 patients with MTLE-HS for each of the reconstruction methods.

2.5. **Image processing and statistical analysis**

2.5.1. **Image segmentation**

The MR images were segmented using SPM8. This yielded fuzzy GM, WM and CSF images in native, i.e. PET reconstruction space, which were used during A-MAP reconstruction.

2.5.2. **SPM analysis**

All voxel-based analyses were performed using flexible factorial statistical designs in SPM8, except for the voxel-based comparison with respect to the ground truth, for which in-house developed software was used. The reconstructed images of the patient data were first flipped if needed to make sure the epileptogenic focus was always at the left side. All reconstructions of the measured human data were then first normalized to cancel out differences in average WM activity, as proposed in *(30)*. Next, these images were spatially normalized to the stereotactic Montreal Neurological Institute (MNI) space using the non-linear warping obtained during MR segmentation. Finally, these images were smoothed with an isotropic Gaussian with a FWHM

of 12 mm to account for gyral variations between subjects. The reconstruction images of the simulated data could be analyzed without activity and spatial normalization and without smoothing, highlighting the differences purely induced by the reconstruction algorithm used. Nevertheless, to make the outcome of the simulation study more predictive for the differences to be expected in a clinical study, the reconstructed images of the simulated data were also analyzed after application of the typical post-processing steps, i.e. spatial normalization and smoothing.
Differences were studied on a voxel-wise basis in a treatment-comparison design (3 algorithms, paired t-tests) for the algorithm comparison study and in a population-comparison design (patients versus controls, unpaired t-tests) for the group analysis. For the simulated data, 60 scans of one subject were used. For the clinical data, 1 scan per subject was available. The images were brain-masked prior to voxel-based analysis. The significance threshold at the voxel level was set to a family wise error (FWE) corrected $p \leq 0.05$. Only clusters of at least 20 suprathreshold voxels (which corresponds to 0.16 cm^3 or about 1/10 of the smallest lesion) were considered in the group analysis.

2.5.3. **Overlap quantification**

In order to evaluate how well each algorithm can localize voxels that are systematically different in one group compared to in the other group, a set a figures of merit was calculated from the images obtained with each of the reconstruction algorithms. The percentages overlap, under- and overestimation by algorithm A compared to B were computed as follows:

$$\%\ \text{overlap} = 100\ \frac{2\ \#(\mathrm{SV_A}\ \cap\ \mathrm{SV_B})}{\#\mathrm{SV_A} + \#\mathrm{SV_B}},$$

$$\%\ \text{underestimation} = 100\frac{\#\mathrm{SV}_B\ - \#(\mathrm{SV_A}\ \cap\ \mathrm{SV_B})}{\#\mathrm{SV_A} + \#\mathrm{SV_B}},$$

$$\%\ \text{overestimation} = 100\ \frac{\#\mathrm{SV_A}\ - \#(\mathrm{SV_A}\ \cap\ \mathrm{SV_B})}{\#\mathrm{SV_A} + \#\mathrm{SV_B}},$$

with $\cap$ denoting the intersection between two sets of voxels, with # the number of elements in a set and with SV_A and SV_B the sets of significant voxels found with algorithm A and B, respectively. In case of simulated data, these figures of merit were also computed with respect to the true lesion location by replacing SV_B by the set of true lesion voxels SV_T, which we defined in the resampled ground truth images as all voxels that showed at least 1 % reduction in activity in the brain with lesions compared to in the normal brain. For the simulation study, we also computed the % lesion coverage and % volume estimation error as

$$\%\ \text{lesion coverage} = 100\ \frac{\#(\mathrm{SV_A}\ \cap\ \mathrm{SV_T})}{\#\mathrm{SV_T}},\ \%\ \text{volume estimation error} =\ 100\frac{\#\mathrm{SV_A} - \#\mathrm{SV_T}}{\#\mathrm{SV_T}}$$

2.5.4. **Statistical analysis**

Statistical significance is reported as being at or above the 95 % limit (corrected $p \leq 0.05$, family-wise error).

3. **Results**

3.1. **Bias with respect to true activity and differences between algorithms**

The results of the comparison studies are shown in Fig. 1-4. The colored voxels indicate the percentage difference in reconstructed activity, but only in those voxels in which these differences were considered to be significant. For the mutual comparison of the reconstruction algorithms (Fig. 2-4), three voxel-based analyses were performed, i.e. on reconstructions of

simulated normal brain data in native space, of simulated data after spatial normalization and smoothing for comparison to real data, and of measured healthy control data.

3.1.1. **Post-smoothed MLEM vs. ground truth**

MLEM post-smoothed with a Gaussian with 4 mm FWHM (PS4-MLEM) systematically underestimates the true activity values in GM voxels (-20 to -40 %) with larger errors in GM voxels close to CSF than in those surrounded by WM (see Fig. 1, top row). A huge positive bias was observed in CSF voxels, as well as in WM voxels close to the gyri (> +70 %). Inline Supplementary Figure 1 allows the reader to study the significant differences found in post-smoothed MLEM images in more detail.

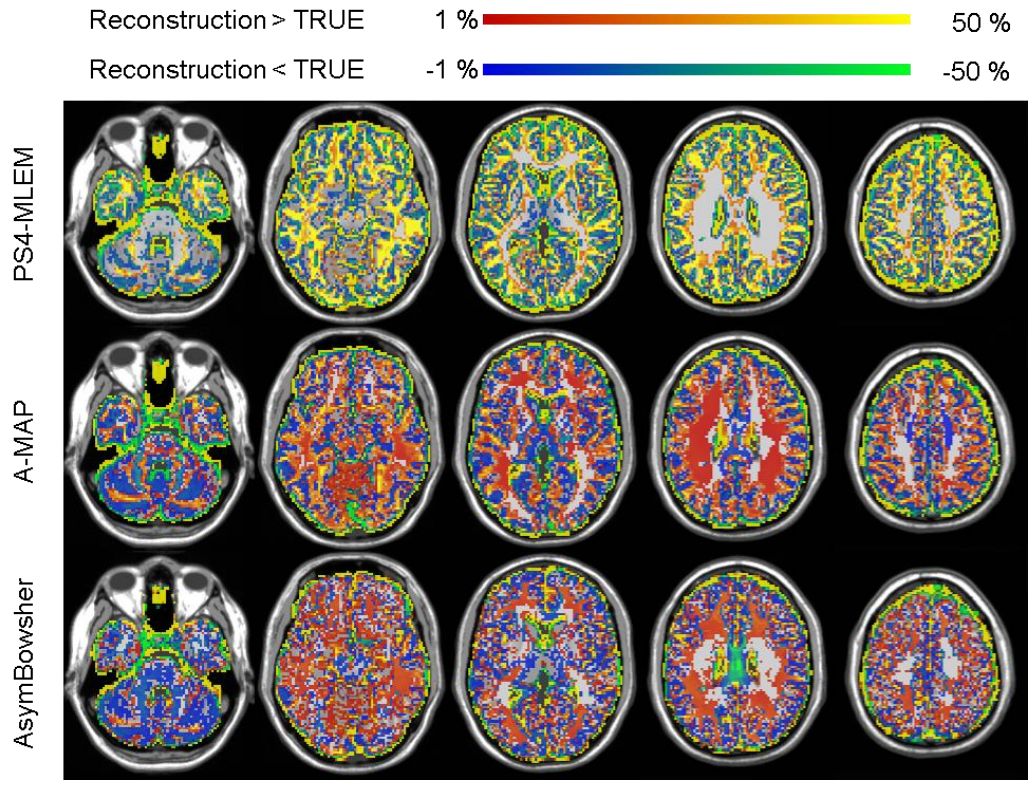

Figure 1: *Comparison of the three reconstruction algorithms with respect to the ground truth using a simulated brain phantom with a normal ^{18}F-FDG activity distribution: (top) post-smoothed MLEM, (center) A-MAP, (bottom) AsymBowsher. The colored voxels represent the percentage difference between the reconstructed value and the true activity concentration. Only significant differences are shown ($p_{cor} \leq 0.05$).*

3.1.2. **A-MAP vs. ground truth**

In the A-MAP images (Fig. 1, central row), edge GM-WM voxels received too high a value (+10 to +25 %), as well as WM voxels close to the gyri (+3 to +5 %), except for those at the top-center of the brain (-3 to -5 %). GM voxels close to CSF, on the contrary, were reconstructed with too low a value (-5 to -10 %). Edge GM-CSF voxels were even more underestimated. Clear quantification errors were seen in the subcortical regions, i.e. an

underestimation (-3 to -8 %) in dense regions and overestimation (+10 and +20) in less dense regions. This was not observed for the other two reconstruction algorithms, except for a slight overestimation (+3 to + 5 %) in the mid-caudate in AsymBowsher images. Inline Supplementary Figure 2 allows the reader to study the significant differences found in A-MAP images in more detail.

3.1.3. **AsymBowsher vs. ground truth**

In the AsymBowsher images, clear significant overestimation was seen in the WM areas (+7 to +15 %, see Fig. 1, bottom row). Furthermore, CSF was reconstructed with too high a value like with post-smoothed MLEM, while the activity in GM-WM edge voxels was underestimated (-5 to -10 %). The GM voxels next to the outer CSF were underestimated, but less than in the A-MAP images. Inline Supplementary Figure 3 allows the reader to study the significant differences found in AsymBowsher images in more detail.

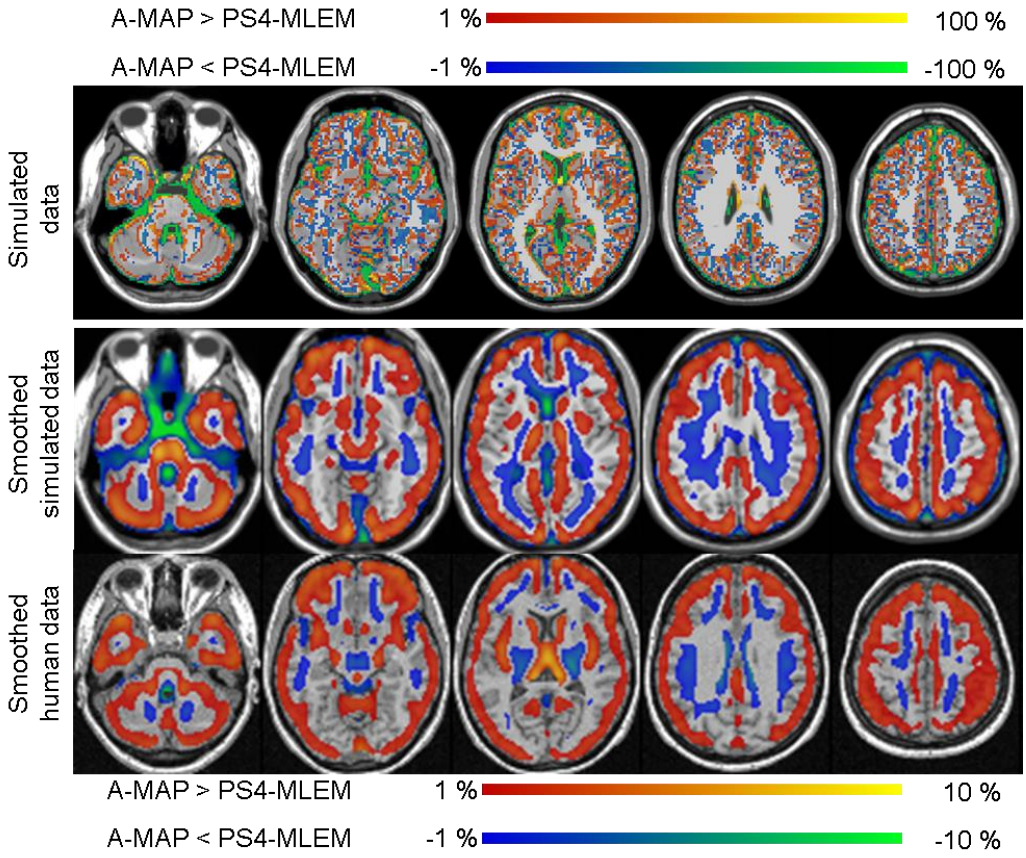

Figure 2: *Comparison between A-MAP and post-smoothed MLEM reconstructions of normal PET brain scans. The colored voxels represent the percentage difference between the activity concentration reconstructed with A-MAP and post-smoothed MLEM. Only significant differences are shown ($p_{cor} \leq 0.05$).*

3.1.4. **A-MAP vs. post-smoothed MLEM**

From the top row images in Fig. 2, A-MAP was found to yield significantly higher values (+15 tot +50 %) inside the outer cortex and in GM-WM edge voxels when compared to MLEM with resolution modeling and post-smoothed with a Gaussian with 4 mm FWHM (red-yellow voxels). A-MAP also showed reduced activity between the gyri both in CSF voxels (-80 to -90 %) and in WM voxels (-30 to -40 %) (blue-green voxels),compared to PS4-MLEM. After spatial normalization and smoothing (see Fig. 2, middle row), these trends became even more clear and the differences became more significant, but also much smaller (in the order of a few percent). From the analysis of the clinical data (Fig. 2, bottom row), we found clusters at similar locations, i.e. red-yellow clusters on GM voxels and blue-green clusters next to the gyri. Due to the inter-patient variability, the significance of the differences was smaller and less clusters showed up, but similar intensity differences were present.

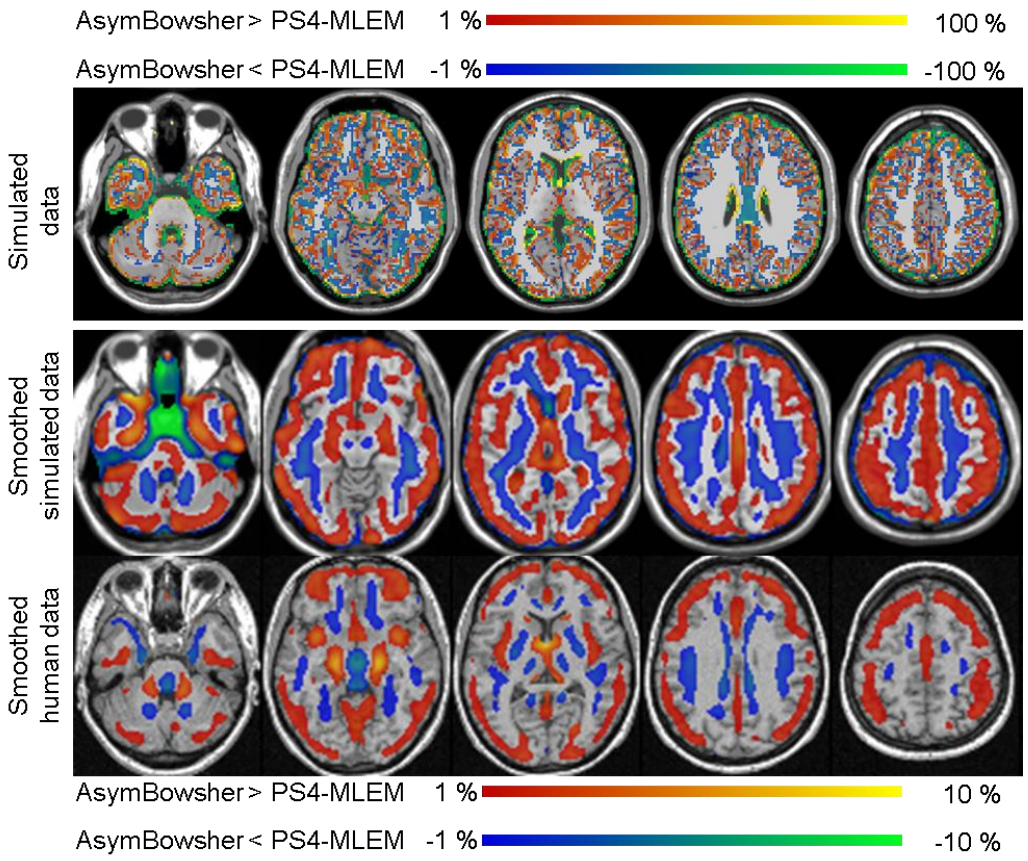

Figure 3: *Comparison between AsymBowsher and post-smoothed MLEM reconstructions of normal PET brain scans. The colored voxels represent the percentage difference between the activity concentration reconstructed with AsymBowsher and post-smoothed MLEM. Only significant differences are shown ($p_{cor} \leq 0.05$).*

3.1.5. **AsymBowsher vs. post-smoothed MLEM**

Similar to A-MAP, AsymBowsher resulted in much higher reconstruction values (+20 to +50 %) in the outer cortex and in GM-WM edge voxels than PS4-MLEM did (see red-yellow voxels in the top row images of Fig. 3). Furthermore, it yielded a 30 % reduction in activity in WM voxels that are next to GM voxels and a 70 % reduction in uptake estimation in CSF voxels (blue-green voxels). From the spatially normalized and smoothed images (central and bottom row images of Fig. 3), similar conclusions could be drawn, but with smaller absolute differences. Again, many differences lost their significance due to inter-patient variabilities, when comparing the human images (e.g., see outer CSF voxels, bottom row in Fig. 3).

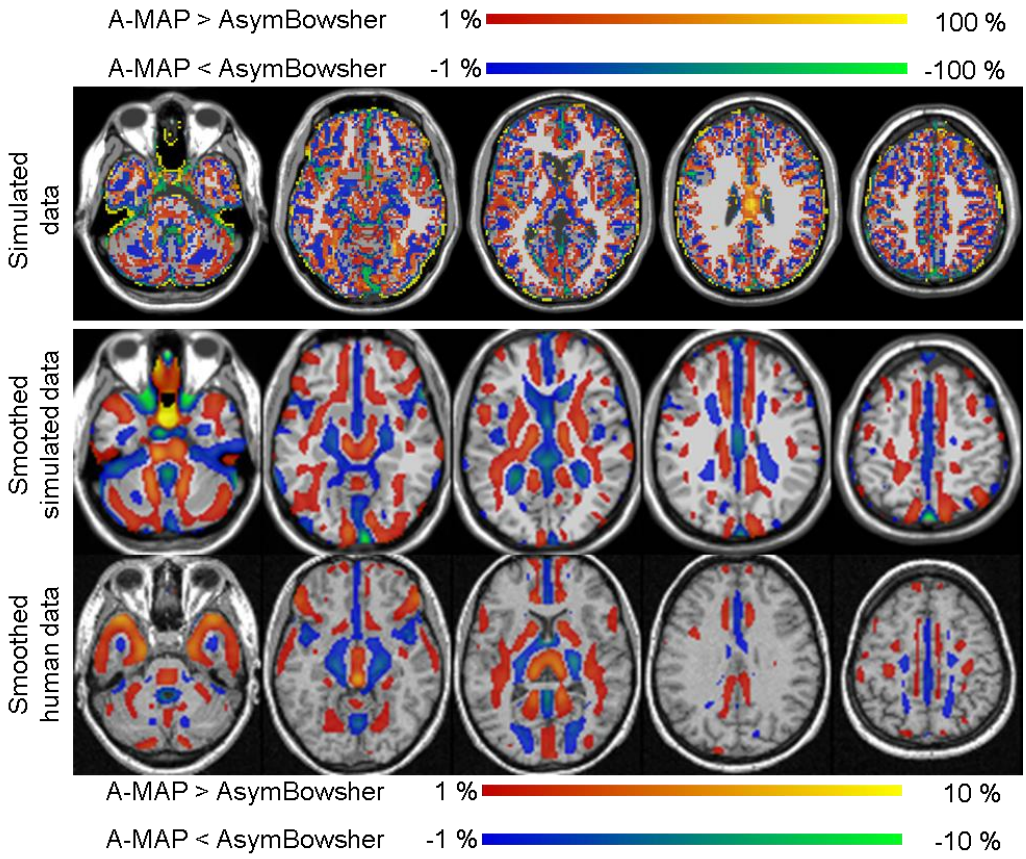

Figure 4: *Comparison between A-MAP and AsymBowsher reconstructions of normal PET brain scans. The colored voxels represent the percentage difference between the activity concentration reconstructed with A-MAP and AsymBowsher. Only significant differences are shown ($p_{cor} \leq 0.05$).*

3.1.6. **A-MAP vs. AsymBowsher**

Comparing the MAP reconstructions with two different anatomical priors in Fig. 4 (top row), it is seen that A-MAP generates significantly higher values (+20 to +40 %) in GM-WM edge voxels than AsymBowsher (red-yellow voxels). The outer CSF voxels have clearly higher

values in the A-MAP images, but these differences are less significant than those found in GM-WM edge voxels. GM-CSF edge voxels typically have a lower reconstructed value with A-MAP than with AsymBowsher (about 10 % reduction between the gyri and much more when a lot of CSF is present, see blue-green voxels in the top row of Fig. 4). It should also be noted that the subcortical structures, like caudate, putamen and thalamus have in full-GM regions 6-12 % lower and in GM-WM transition regions 10-25 % higher values in A-MAP images than in AsymBowsher images. In the smoothed images, both in the simulated data and in the measured human data (central and bottom row of Fig. 4), we see the same main trends, i.e. GM-WM edges have higher values in A-MAP images, whereas GM-CSF edges have higher estimates in AsymBowsher images.

Table 1: *Percentage overlap, lesion coverage and volume estimation error of significantly different voxels in simulated normal brains versus brains with hypometabolic lesions ($p_{cor} \leq 0.05$; cluster extent > 20 voxels), detected with different reconstruction algorithms and compared to the true lesion locations. The reconstructed images were analyzed in native space without additional smoothing (except for PS4-MLEM).*

	MLEM	PS4-MLEM	A-MAP	AsymBowsher
% underestimation	76.0 %	30.9 %	30.2 %	8.1 %
% overlap	**24.0 %**	**69.0 %**	**69.6 %**	**71.5 %**
% overestimation	0.0 %	0.1 %	0.2 %	20.3 %
% lesion coverage	13.6 %	52.7 %	53.6 %	81.4 %
% volume estimation error	-86.4 %	-47.1 %	-46.1 %	27.7 %

Table 2: *Percentage overlap of significantly different voxels in simulated normal brains versus brains with hypometabolic lesions ($p_{cor} \leq 0.05$; cluster extent > 20 voxels), detected with MLEM, A-MAP and AsymBowsher, and compared to those detected with PS4-MLEM. The reconstructed images were spatially normalized and smoothed with a Gaussian with FWHM of 12 mm.*

	MLEM	A-MAP	AsymBowsher
% underestimation	4.9 %	0.0 %	0.0 %
% overlap	**95.1 %**	**91.7 %**	**88.4 %**
% overestimation	0.0 %	8.3 %	11.6 %

3.2. **Detection of Group Differences**

3.2.1. **Simulated Data**

The outcome of the group analysis, performed on the simulated data, is listed in Table 1. From the group analysis on the MLEM images, 17 out of the 20 hypointense lesions were observed. However, they were clearly smaller than the true lesions, i.e. the total volume was only 13.6 % of the true total volume and the overlap with the true lesions was only 24.0 %. The overlap improved when using post-smoothed MLEM (PS4-MLEM). Very similar results were obtained with A-MAP. The smallest absolute volume estimation error and highest overlap (71.5 %) was achieved with AsymBowsher. In this case, the lesion size was over- instead of underestimated, so some voxels were falsely indicated as hypointense. All simulated lesions were detected with PS4-MLEM, A-MAP and AsymBowsher. In none of the cases were additional lesions found.

To predict how well the different algorithms perform in a realistic group analysis, the reconstructions of the simulated data were processed similarly as typically done in a group analysis of measured human data, i.e. they were spatially normalized and smoothed with a Gaussian with 12 mm FWHM to account for gyral variations between subjects. In the smoothed images, some lesions merged to one, but all were detected and no additional lesions were observed. Due to the heavy smoothing, the size of the lesions was highly overestimated (total volume estimation error ranging from 517 % to 758 % for MLEM and AsymBowsher, respectively). Again, the outcome of the different reconstruction methods only differed in the size of the clusters with significantly different voxels, with MLEM yielding the smallest and AsymBowsher the biggest ones. This can also be concluded from Table 2, which summarizes the quantified overlap between the lesions detected by using PS4-MLEM and those detected by MLEM (95.1 %), A-MAP (91.7 %) and AsymBowsher (88.4 %).

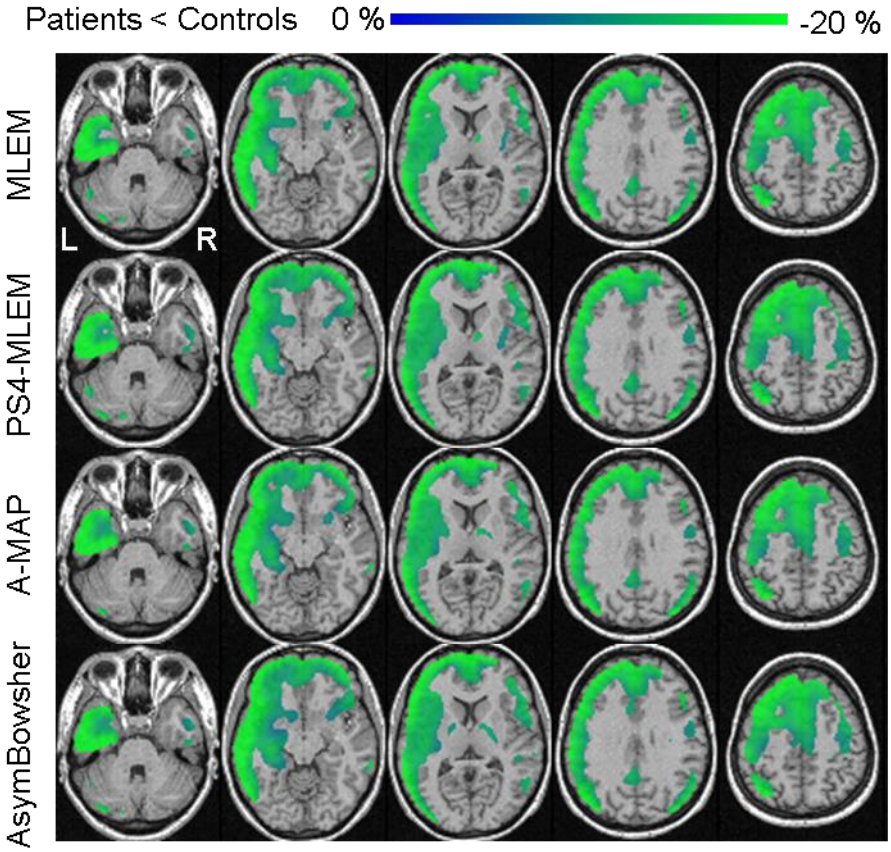

Figure 5: *Outcome of the group analyses on the measured human data. The clusters with significantly decreased activity (MTLE-HS patients < healthy controls; $p_{cor} \leq 0.05$; cluster extent > 20 voxels) were overlaid on transaxial slices of an individual MR (in MNI space) of a healthy volunteer. The blue-green color scale indicates increasing percentage difference in activity. Top to bottom: MLEM, post-smoothed MLEM, A-MAP and AsymBowsher.*

Table 3: *Percentage overlap of significantly different voxels in patients versus healthy controls ($p_{cor} \leq 0.05$; cluster extent > 20 voxels), detected with MLEM, A-MAP and AsymBowsher, compared to those detected with PS4-MLEM. The reconstructed images were analyzed in MNI space and smoothed with a Gaussian with FWHM of 12 mm.*

	MLEM	A-MAP	AsymBowsher
% underestimation	5.4 %	4.5 %	2.9 %
% overlap	**94.6 %**	**94.5 %**	**94.2 %**
% overestimation	0.0 %	1.0 %	2.9 %

3.2.2. **Measured Human Data**

Figure 5 shows which voxels have a significantly lower activity value in the MTLE-HS patients compared to healthy controls. The locations are in agreement with the findings in *(30)*, i.e. most pronounced in the frontal and temporal lobe at the side of the epileptogenic focus, which is at the left (remember that right-sided cases have been flipped prior to image processing). The color scale from blue to green indicates the percentage difference between the mean reconstruction images of the two groups, with green a stronger decrease than blue. As can be seen, the amount of decrease is also very similar. Finally, the sets of significantly different voxels obtained with the different reconstruction methods were quantitatively compared in Table 3. Again, the clusters were smallest for MLEM, and those obtained with MLEM, A-MAP and AsymBowsher overlapped for 94.6 %, 94.5 % and 94.2 %, respectively, with those from PS4-MLEM.

4. **Discussion**

All reconstruction methods under investigation include resolution recovery, which makes that even MLEM yields a strong partial volume correction. Because resolution recovery requires many iterations to reach convergence, the MLEM images are too noisy to be useful in clinical practice. Therefore, its quantification accuracy was only investigated after post-smoothing, which obviously resulted in a great loss of the recovered resolution, apparent as a significant underestimation of the GM activity and overestimation of the values in WM and CSF voxels close to GM voxels.

A-MAP and AsymBowsher both yielded very good resolution recovery in the cortex, with less than 10 % underestimation in GM voxels close to CSF and WM, respectively. Larger errors were seen in edge voxels and in WM and CSF voxels (especially in AsymBowsher images). A-MAP was found to provide less accurate estimates for the activity in subcortical regions than AsymBowsher.

When comparing the two anatomy-enhanced MAP algorithms, it was seen that the activity at the edges of the gyri did not coincide perfectly and that the edges were not equally steep in both images. This was , e.g., highlighted by the significantly higher activity in GM-WM edge voxels in A-MAP images compared to in AsymBowsher images. The reason for this is twofold. On the one hand, the GM-WM edge voxels were overestimated in the A-MAP images because of a slight oversegmentation. On the other hand, AsymBowsher always smooths over an equal amount of neighbors (in our case 13 out of 80), which might be suboptimal at the edges. The figure of merit used to tune this parameter did only consider the reconstruction accuracy of voxels with a high fraction of GM and of voxels inside lesions, not that of WM voxels (nor of edge voxels). The selected reconstruction parameters are thus not ideal for quantifying the uptake in the WM. This was also pointed out by the systematic overestimation in the WM areas in the AsymBowsher vs. true activity analysis. All of these findings indicate a ‘leakage’ of the GM activity into the WM regions during reconstruction. Nevertheless, the bias is small in the

GM voxels, which are the most important ones in ^{18}F-FDG PET brain imaging of patients with epilepsy. This was confirmed by the fact that no significant differences were found between the values in the inner GM voxels estimated by AsymBowsher or A-MAP.

A-MAP was found to be the only of the investigated algorithms that can nicely drive the CSF values to very small values, thanks to the absolute intensity prior applied to the CSF voxels during A-MAP reconstruction and due to the full exploitation of the fuzzy segmentation information, which effectively prevents smoothing over tissue boundaries. MLEM and AsymBowsher have more problems to recover the very steep GM-CSF edges and to yield small values in CSF voxels. Nevertheless, the large percentage difference with respect to the ground truth in the voxels with large fractions of CSF overamplify the absolute errors, because their true activity is near zero, such that even small absolute errors translate into big relative errors. Finally, the (positive and negative) bias which was apparent in the subcortical regions in the A-MAP images can be explained by the lower segmentation accuracy of SPM8 in these regions compared to in cortical regions.

From the group analysis of the reconstructions of the simulated normal and abnormal brain data, it could be concluded that all lesions could be detected, without detecting false positives, using post-smoothed MLEM, A-MAP and AsymBowsher. Using unprocessed MLEM, a few lesions were missed. Furthermore, it was found that the locations of the detected lesions were very similar, but that the size differed a bit, with in order of increasing size: MLEM, post-smoothed MLEM, A-MAP, and AsymBowsher. However, one should not rely too much on the exact numbers of the overlap, because the definition of the exact lesion location is not perfect due to the need to resample the true activity image or lesion location mask. Indeed, we found that changing the definition introduced clear changes in the overlap measure. Nevertheless, the trends remained the same. Unfortunately, this type of analysis can only be applied in case the same patient is scanned repeatedly over time, which will rarely be more than twice in practice.

More relevant was the group analysis on the processed simulated brain data, which taught us that all investigated methods (also MLEM without initial post-smoothing) were able to correctly detect all lesions. Furthermore, the same trend of increase in detected cluster size from MLEM over post-smoothed MLEM and A-MAP to AsymBowsher was found, as well as a high amount of overlap between the clusters of each of the methods. It should, however, be noted that the detected lesion size is a very large overestimation of the true lesion size (5 to 7.5 times too big for these small to medium-sized lesions). Although no inter-patient gyral variability was modeled, the results of the second simulation study showed to be quite predictive for the outcome of the group analysis applied to the measured human data, from which the same conclusions could be drawn. The obtained clusters did not only overlap very well, they also yielded very similar decrease in activity, although this should not be taken as the true decrease due to the large amount of smoothing that has been applied. It is without doubt that a simulation study which resembles more closely the problem at hand, e.g., using multiple brain phantoms and modeling both functional and anatomical variability as in *(31)*, would be needed to enable us to determine which algorithm is more suited for this task. Ideally, one would even optimize the parameters of each algorithm for each task at hand. Unfortunately, it is unlikely that this will happen in practice. Therefore, it is probably more relevant to perform a robustness analysis, by checking how variable the outcome of each of the reconstruction methods is when changing the parameters a bit. This study was considered out of the scope of this work.

5. **Conclusion**

Three state-of-the-art iterative algorithms for PET image reconstruction, i.e. (post-smoothed) MLEM with resolution recovery, and two MAP algorithms using either a segmentation-based (A-MAP) or a segmentation-less (AsymBowsher) anatomical prior, have been compared for

their use in human ^{18}F-FDG PET brain imaging using voxel-based statistical analysis techniques. In a first study, the algorithms were compared to the true activity and to each other. Significant differences were mainly found at GM edges and in non-GM voxels. Systematic bias of A-MAP and AsymBowsher in estimating the activity in cortical voxels consisting of a high fraction of GM was limited to -5 to -10 %. A-MAP also showed significant errors in the subcortical regions due to segmentation inaccuracies of SPM8 at the gradual GM-WM transitions. When comparing AsymBowsher to A-MAP, significant differences in GM activity only showed up at edges and in subcortical regions. Furthermore, both A-MAP and AsymBowsher provided much higher quantitative accuracy than post-smoothed MLEM, with A-MAP yielding the sharpest edges, but being sensitive to segmentation errors. Especially for studying subcortical regions, the use of AsymBowsher is recommended. Finally, similar trends were seen when the methods were applied to a set of ^{18}F-FDG PET brain scans of 20 healthy volunteers.

In a second study, the algorithms were compared for their performance to detect group differences. An idealized group comparison of the reconstructions of normal brain scans versus those of scans of a brain with hypometabolic lesions revealed that post-smoothed MLEM, A-MAP and AsymBowsher all can yield perfect detection of the simulated lesions in this situation, whereas unprocessed MLEM missed 3 out of 20 lesions. Differences in the size of the detected lesions were observed (smallest for MLEM, largest for AsymBowsher). However, most group analyses require spatial normalization and severe smoothing to compensate for inter-patient gyral variations. Therefore, a more realistic simulation-based group analysis and a real data analysis comparing a group of 20 healthy volunteers to a group of 16 patients with mesial temporal lobe epilepsy with hippocampal sclerosis were also performed. Very similar clusters were found with all investigated reconstruction methods.

6. Acknowledgements

K. Vunckx is a postdoctoral researcher of the Research Foundation - Flanders (FWO Vlaanderen). K. Van Laere is a Senior Clinical Investigator of the Research Foundation - Flanders.

References

1. Varrone A, Asenbaum S, Vander Borght T, et al. EANM guidelines for PET brain imaging using [18F]FDG, version 2. *Eur J Nucl Med Mol Imaging.* 2009;36:2103-2110.
2. Wienhard K, Schmand M, Casey ME, et al. The ECAT HRRT: Performance and first clinical application of the new high resolution research tomograph. *IEEE Trans Nucl Sci.* 2002;49:104-110.
3. Adam L-E, Zaers J, Ostertag H, et al. Performance evaluation of the whole-body PET scanner ECAT EXACT HR+ following the IEC standard. *IEEE Trans Nucl Sci.* 1997;44:1172-1179.
4. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A.* 2000;97:11050-11055.
5. Hoffman EJ, Huang SC, Phelps ME. Quantitation in positron emission computed tomography: 1. Effect of object size. *J Comput Assist Tomogr.* 1979;3:299–308.
6. Yang J, Huang SC, Mega M, et al. Investigation of partial volume correction methods for brain FDG PET studies. *IEEE Trans Nucl Sci.* 1996;43:3322-3327.
7. Sureau FC, Reader AJ, Comtat C, et al. Impact of image-space resolution modeling for studies with the high-resolution research tomograph. *J Nucl Med.* 2008;49:1000-1008.
8. Alessio AM, Rahmim A, Orton CG. Point/counterpoint. Resolution modeling enhances PET imaging. Med Phys. 2013;40:120601.

9. Vunckx K, Atre A, Baete K, et al. Evaluation of three MRI-based anatomical priors for quantitative PET brain imaging. *IEEE Trans Med Imaging*. 2012;31:599-612.
10. Baete K, Nuyts J, Van Laere K, et al. Evaluation of anatomy based reconstruction for partial volume correction in brain FDG-PET. *NeuroImage.* 2004;23:305-317.
11. Goffin K, Van Paesschen W, Dupont P, et al. Anatomy-based reconstruction of FDG-PET images with implicit partial volume correction improves detection of hypometabolic regions in patients with epilepsy due to focal cortical dysplasia diagnosed on MRI. *Eur. J. Nucl. Med. Mol. Imaging.* 2010;37(6):1148-55.
12. Babalola KO, Patenaude B, Aljabar P, et al. An evaluation of four automatic methods of segmenting the subcortical structures in the brain. *NeuroImage.* 2009;47:1435-1447.
13. Bowsher JE, YuanH, Hedlund LW, et al. Using MRI information to estimate F18-FDG distributions in rat flank tumors. *IEEE Nucl Sci Symp Conf Record*. 2004;2488-2492.
14. Somayajula S, Asma E, Leahy RM. PET image reconstruction using anatomical information through mutual information based priors. *IEEE Nucl Sci Symp Conf Record.* 2005;2722-2726.
15. Nuyts J. The use of mutual information and joint entropy for anatomical priors in emission tomography. *IEEE Nucl Sci Symp Conf Record.* 2007;4149-4154.
16. Nuyts J, Baete K, Bequé D, Dupont P. Comparison between MAP and post-processed ML for image reconstruction in emission tomography when anatomical knowledge is available. *IEEE Trans Med Imaging.* 2005;24:667-675.
17. Shepp LA, Vardi Y. Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging.* 1982;1:113-122.
18. Nuyts J, Bequé D, Dupont P, Mortelmans L. A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography. *IEEE Trans Nucl Sci.* 2002;49:56-60.
19. Baete K, Nuyts J, Van Paesschen W, Suetens P, Dupont P. Anatomical based FDG-PET reconstruction for the detection of hypometabolic regions in epilepsy. *IEEE Trans. Med. Imaging.* 2004;23:510-519.
20. Vunckx K, Nuyts J. Heuristic modification of an anatomical Markov prior improves its performance. *IEEE Nucl Sci Symp Conf Record.* 2010;3262-3266.
21. Reilhac A, Lartizien C, Costes N, et al. PET-SORTEO: A Monte Carlo-based simulator with high count rate capabilities. *IEEE Trans Nucl Sci.* 2004;51:46-52.
22. Hudson HM, Larkin RS. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging.* 1994;13:601-609.
23. Picard Y, Thompson CJ. Motion correction of PET images using multiple acquisition frames. *IEEE Trans Med Imaging.* 1997;16:137-144.
24. Watson CC. New, faster, image-based scatter correction for 3D PET. *IEEE Trans Nucl Sci.* 2000;47:1587-1594.
25. website: "http://www.bic.mni.mcgill.ca/brainweb/".
26. Collins DL, Zijdenbos AP, Kollokian V, et al. Design and construction of a realistic digital brain phantom. *IEEE Trans Med Imaging.* 1998;17:463-468.
27. Vander Borght Th, Laloux P, Maes A, et al. Guidelines for brain radionuclide imaging. *Acta Neurol Belg.* 2001;101:196-209.
28. De Reuck J, Decoo D, Hasenbroekx MC, et al. Acetazolamide vasoreactivity in vascular dementia: a positron emission tomographic study. *Eur Neurol.* 1999;41:31-36.
29. Besson P, Andermann F, Dubeau F, Bernasconi A. Small focal cortical dysplasia lesions are located at the bottom of a deep sulcus. *Brain*. 2008;131: 3246-3255.

30. Nelissen N, Van Paesschen W, Baete K, et al. Correlations of interictal FDG-PET metabolism and ictal SPECT perfusion changes in human temporal lobe epilepsy with hippocampal sclerosis. *NeuroImage.* 2006;32:684-695.
31. Aguiar P, Pareto D, Gispert JD, et al. Effect of anatomical variability, reconstruction algorithms and scattered photons on the SPM output of brain PET studies. *NeuroImage.* 2008;39:1121-1128.