Evaluation of Anatomy Based Reconstruction for Partial Volume Correction in Brain FDG-PET

Kristof Baete [a,*], Johan Nuyts [a], Koen Van Laere [a], Wim Van Paesschen [b], Sarah Ceyssens [a], Liesbet De Ceuninck [a], Olivier Gheysens [a], Annemarie Kelles [a], Jimmy Van den Eynden [a], Paul Suetens [c], and Patrick Dupont [a]

[a] *Department of Nuclear Medicine, UZ Gasthuisberg, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium*

[b] *Department of Neurology, UZ Gasthuisberg, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium*

[c] *Laboratory for Medical Image Computing, Radiology-ESAT/PSI, UZ Gasthuisberg, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium*

Abstract

FDG-PET contributes to the diagnosis and management of neurological diseases. In some of these diseases, pathological gray matter (GM) areas may have a reduced FDG uptake. Detection of these regions can be difficult and some remain undiscovered using visual assessment. The main reason for this detection problem is the relatively small thickness of GM compared to the spatial resolution of PET, known as the partial volume effect. We have developed an anatomy based maximum-a-posteriori reconstruction algorithm (A-MAP) which corrects for this effect during the reconstruction using segmented magnetic resonance (MR) data. Monte-Carlo based 3-D brain software phantom simulations were used to investigate the influence of the strength of anatomy based smoothing in GM, the influence of misaligned MR data, and the effect of local segmentation errors. A human observer study was designed to assess the detection performance of A-MAP versus post-smoothed maximum-likelihood (ML) reconstruction. We demonstrated the applicability of A-MAP using real patient data. The results for A-MAP showed improved recovery values and robustness for local segmentation errors. Misaligned MR data reduced the recovery values towards those obtained by post-smoothed ML, for small registration errors. In the human observer study, detection accuracy of hypometabolic regions was significantly improved using A-MAP, compared to post-smoothed ML ($p < 0.004$). The patient study confirmed the applicability of A-MAP in clinical practice. Conclusion: A-MAP is a promising technique for voxel-based partial volume correction of FDG-PET of the human brain.

Paper accepted for publication in NeuroImage *27th April 2004*

Key words: iterative reconstruction, partial volume, positron emission tomography, spill-over correction

1 Introduction

Positron emission tomography (PET) of the cerebral glucose metabolism, reflected by the uptake of radioactive labelled 2-[^{18}F]fluoro-2-deoxy-D-glucose (FDG), has shown to be very useful for the diagnosis and management of a number of neurological diseases. In some neurological diseases, gray matter (GM) areas associated with the pathological region show a reduced FDG uptake. Some of these hypometabolic regions may be subtle so that they are overlooked by nuclear medicine physicians when examining the FDG-PET images visually. The main reasons for this detection problem are the limited spatial resolution and the noise affecting PET images.

The average GM thickness, which is about 2.5 mm (Fischl and Dale, 2000), is relatively small compared to the spatial resolution of PET, which is about 5 mm (Brix et al., 1997), and the voxel size used for brain PET images, which is typically about 2 mm in all directions. Since the human brain has a natural variation in GM thickness and geometry, measuring small brain structures will lead to an underestimation of the tracer activity. This well-known partial volume effect (PVE) can lead to spurious hypometabolic regions (Fazio and Perani, 2000). In the diagnosis of neurological diseases, PVE can yield a misinterpretation of the true glucose consumption (false positive or false negative).

Many groups have developed techniques to correct for PVE. Some of these groups have explored techniques for the use of anatomical information. In general, literature show two methodological approaches: (a) use of anatomical information after the reconstruction process, or post-processing methods, and (b) use of anatomical information as a priori knowledge in a Bayesian reconstruction framework. One of the post-processing techniques was originally designed to correct for PVE due to the cerebrospinal fluid (CSF) (Meltzer et al., 1990). Later, this technique was modified to allow voxel based partial volume correction (PVC) in GM (Müller-Gärtner et al., 1992; Labbé et al., 1996). The Bayesian reconstruction framework has been shown to be useful for modelling different kinds of physical effects, including the incorporation of anatomical information. For example, anatomy based weighting of Gibbs smoothing priors (Fessler et al., 1992; Gindi et al., 1993; Comtat et al., 2002), intensity distribution priors (Bowsher et al., 1996; Lipinski et al., 1997; Sastry and Carson, 1997), minimum cross-entropy reconstruction with anatomy based smoothing (Ardekani et al., 1996), or using a joint mixture framework (Rangarajan et al., 2000).

Recently, we have developed an anatomy based maximum-a-posteriori iterative re-

* Corresponding author is Kristof Baete
Email address: Kristof.Baete@uz.kuleuven.ac.be (Kristof Baete).
[1] This work was supported by K.U.Leuven grants IDO-99/005, OT-00/32 and F.W.O. grant G.0174.03.

construction algorithm, called A-MAP (Baete et al., 2004). A-MAP includes anatomical information as a priori knowledge in a Bayesian reconstruction framework. However, the approximation of a tissue composition model, as it is used in some post-processing methods (Müller-Gärtner et al., 1992), was included in the A-MAP technique as well. A-MAP inherently applies PVC for activity in the GM region. For that purpose, it makes use of anatomical information derived from high-resolution magnetic resonance imaging (MRI) data that has been aligned with the PET data and processed through state-of-the-art segmentation algorithms.

In the following sections, we give a concise description of the A-MAP reconstruction algorithm. A-MAP was originally developed for application in epilepsy. The assumptions made in the derivation of the method are valid in many other PET brain imaging applications. This paper is mainly focussed on the in-depth evaluation of A-MAP. We designed a number of simulation experiments in which we examined the assigned weight to the anatomical information during the image reconstruction, the effect of misaligned MRI data, and the effect of segmentation errors. A human observer study was conducted in which small lesions with reduced activity had to be detected to evaluate the performance of A-MAP. Finally, A-MAP was applied to FDG-PET projection data of a well-documented patient to illustrate the algorithm in a true three-dimensional (3-D) clinical setting. We compared A-MAP systematically with post-smoothed maximum-likelihood (ML) reconstructions of the projection data.

2 Materials and methods

2.1 Anatomical based MAP reconstruction algorithm

We assumed that a spatial distribution of tracer activity $\Lambda = \{\lambda_j\}$ inside the field-of-view of a PET scanner results in the detection of y_i photon pairs along line-of-response i. If λ_j represents the activity in voxel j, and c_{ij} the probability that photons emitted in j are detected in detector pair i, then the expected amount of photon pairs is $E(y_i) = \sum_j c_{ij}\lambda_j$. The system matrix $\{c_{ij}\}$ describes the projection process with the associated physical effects that take place between emission and detection of photons. We included photon attenuation and position independent detector resolution, which is implemented as a convolution in sinogram space. The Bayesian maximum-a-posteriori (MAP) estimation of reconstruction image Λ is computed by maximizing $L(\Lambda, Y) + M(\Lambda)$, in which L is the logarithm of the likelihood of measuring $Y = \{y_i\}$ given Λ, and M is the logarithm of the prior that is used to impose additional constraints on the solution.

The probability that a voxel belongs to GM, white matter (WM), or CSF is identified by segmented high-resolution MRI data. In practice, the high-resolution anatomical information is aligned and down-sampled to the PET sampling grid. Then, we can reasonably assume that the result of the aligned tissue probabilities approximates the underlying tissue composition (Sastry and Carson, 1997). We denote the set of segmentations $\mathcal{F} = \{f^G, f^W, f^C, f^O\}$, with G for GM, W for WM, C for CSF, and O for "other". Each voxel in these segmentation images contains a value between 0 and 1, which represents the tissue fraction for that particular voxel. A voxel in

the reconstruction image Λ can be thought of as a radioactive source containing an amount of tracer activity equal to the total amount of the fractional activities of the contributing tissues, or

$$\lambda_j = f_j^G \lambda_j^G + f_j^W \lambda_j^W + f_j^C \lambda_j^C + f_j^O \lambda_j^O \tag{1}$$

The measured data provides only sufficient information to estimate the sum of all fractional activities in each voxel. However, that approach causes problems when an operator is used within a tissue class (e.g. a smoothing operator to suppress noise). Because some voxels contain more than one tissue, a method is needed to restrict the operator to a single tissue class within one voxel. This requires knowledge of the fractional activities. In this study, the tissue of interest is GM. We define $\mathbb{G} = \{j \,|\, f_j^G > \varepsilon\}$ to be the set of voxels in which the GM tissue fraction exceeds the value of a segmentation threshold parameter ε, with $0 < \varepsilon \ll 1$. To allow estimation of the GM activity, assumptions about the other tissue classes must be introduced. In many applications, it is reasonable to assume that the uptake of tracer activity in WM and CSF can be well modelled as a narrow Gaussian distribution with a position independent mean value. Moreover, CSF activity is expected to approximate zero. For voxels containing GM, we will assume that the activities of WM and CSF are exactly equal to the mean of their distribution. In addition, we will assume that GM is surrounded by WM and CSF only, such that $f_j^O \approx 0$ for all voxels $j \in \mathbb{G}$. With these assumptions, the activity of GM can be directly computed from the total activity of the voxel. Then, for all $j \in \mathbb{G}$, equation (1) is approximated by

$$\forall j \in \mathbb{G} : \lambda_j \approx f_j^G \lambda_j^G + f_j^W \bar{\lambda}^W + f_j^C \bar{\lambda}^C \tag{2}$$

with $\bar{\lambda}^W$ the mean activity in "white matter" region $\mathbb{W} = \{j \,|\, f_j^W > (1-\varepsilon)\}$, and $\bar{\lambda}^C$ the mean activity in "CSF" region $\mathbb{C} = \{j \,|\, f_j^C > (1-\varepsilon)\}$. Then, instead of estimating $\Lambda = \{\lambda_j\}$, we compute the MAP estimate of a new set of variables $\Lambda^* = \{\lambda_j^*\}$, defined by

$$\Lambda^* = \{\lambda_j^G \,|\, j \in \mathbb{G}\} \cup \{\lambda_j \,|\, j \notin \mathbb{G}\} \tag{3}$$

such that the expected amount of measured photon pairs equals $E(y_i) = \sum_i c_{ij}^* \lambda_j^*$ and for which we derived a new system matrix c_{ij}^* (Baete et al., 2004),

$$c_{ij}^* = \delta_j^{\mathbb{G}} f_j^G c_{ij} + (1 - \delta_j^{\mathbb{G}}) c_{ij} + \delta_j^{\mathbb{W}} f_j^W \frac{\sum\limits_{k \in \mathbb{G}} c_{ik} f_k^W}{\sum\limits_{l \in \mathbb{W}} f_l^W} + \delta_j^{\mathbb{C}} f_j^C \frac{\sum\limits_{k \in \mathbb{G}} c_{ik} f_k^C}{\sum\limits_{l \in \mathbb{C}} f_l^C} \tag{4}$$

where

$$\delta_j^{\mathbb{X}} \equiv \begin{cases} 1 & \iff \quad j \in \mathbb{X} \\ 0 & \iff \quad j \notin \mathbb{X} \end{cases} \tag{5}$$

In $\mathbb{W}$, we impose physiological a priori knowledge using a Gaussian distribution prior $M^{\mathbb{W}}$, because $\mathbb{W}$ contains voxels composed of mainly WM, and because we assumed that the WM tracer activity is approximately uniform, compared to the activity in GM. We use a similar prior for voxels in $\mathbb{C}$, i.e. $M^{\mathbb{C}}$. The set of voxels containing only a mixture of WM and CSF is defined as $\mathbb{R} = \{j \notin (\mathbb{G} \cup \mathbb{W} \cup \mathbb{C}) \,|\, (f_j^W + f_j^C) > (1-\varepsilon)\}$. We use a similar prior for voxels in $\mathbb{R}$, i.e. $M^{\mathbb{R}}$. The propagation of noise in $\mathbb{G}$ is suppressed by a Gibbs smoothing prior $M^{\mathbb{G}}$, penalizing small differences between neighboring λ_j^G, but tolerating larger differences (e.g. using the prior function defined

by Nuyts et al., 2002). Then, the logarithm of the prior is $M = M^{\mathbb{W}}+M^{\mathbb{C}}+M^{\mathbb{R}}+M^{\mathbb{G}}$, with

$$M^{\mathbb{W}} = -\beta_{\mathbb{W}} \sum_{j\in\mathbb{W}} (\lambda_j - \bar{\lambda}^W)^2 \tag{6}$$

$$M^{\mathbb{C}} = -\beta_{\mathbb{C}} \sum_{j\in\mathbb{C}} \lambda_j^2 \tag{7}$$

$$M^{\mathbb{R}} = -\beta_{\mathbb{R}} \sum_{j\in\mathbb{R}} (\lambda_j - f_j^W \bar{\lambda}^W)^2 \tag{8}$$

$$M^{\mathbb{G}} = -\beta_{\mathbb{G}} \sum_{j\in\mathbb{G}} \sum_{k\in\mathcal{N}_j^{\mathbb{G}}} \Phi(\lambda_j^G, \lambda_k^G) \tag{9}$$

with $\mathcal{N}_j^{\mathbb{G}} = \mathcal{N}_j \cap \mathbb{G}$, $\mathcal{N}_j$ the neighborhood of voxel j with $j \notin \mathcal{N}_j$, and Φ the local prior weight function based on a measure of the activity in voxel j and voxel k. A heuristic gradient ascent form of the expectation maximization approach was used for the MAP optimization procedure (Baete et al., 2004). After the MAP estimation of Λ^*, we compute the A-MAP reconstruction Λ by

$$\lambda_j = (1-\delta_j^{\mathbb{G}})\lambda_j^* + \delta_j^{\mathbb{G}} \left(f_j^G \lambda_j^* + f_j^W \frac{\sum_{k\in\mathbb{W}} f_k^W \lambda_k^*}{\sum_{l\in\mathbb{W}} f_l^W} + f_j^C \frac{\sum_{k\in\mathbb{C}} f_k^C \lambda_k^*}{\sum_{l\in\mathbb{C}} f_l^C} \right) \tag{10}$$

2.2 Simulation experiments

2.2.1 A-MAP and partial volume correction

The A-MAP reconstruction algorithm corrects for the PVE. We illustrate this property using a simple two-dimensional (2-D) software phantom (see Fig. 1a). The image dimensions are 200×200 voxels. In the center of the image, we inserted a uniform disc with a radius of 90 voxels and an intensity of 0.5. Along the equator of the disc, we added five smaller discs with a radius of 10, 8, 6, 4, and 2 voxels, in which the intensity was set to 1. We also inserted a disc with a radius of 4 voxels in which the intensity was set to 0.75. A PET acquisition was simulated by projecting the phantom over 120 angles. The detector resolution effect was simulated by smoothing the projections along the detector using a one-dimensional (1-D) Gaussian kernel with a full-width at half-maximum (FWHM) of 5 voxels. We ignored attenuation and the contribution of noise.

The projection data were reconstructed using the A-MAP reconstruction algorithm and the maximum-likelihood (ML) reconstruction algorithm (Shepp and Vardi, 1982). We considered the smaller discs as "GM" and the large disc as "WM." The contribution of "CSF" was ignored. We simulated fuzzy "GM" and "WM" segmentation images by smoothing the exact anatomical information, derived from the phantom data, using a 2-D Gaussian kernel with an isotropic FWHM of 1.5 voxels. The segmentation threshold parameter was set to $\varepsilon = 0.01$, the weights of the Gaussian priors were $\beta_{\mathbb{W}} = 10$, $\beta_{\mathbb{R}} = \beta_{\mathbb{C}} = 0$, and the weight of the Gibbs smoothing prior was $\beta_{\mathbb{G}} = 10$. For the local prior weight function Φ, we used the relative difference prior (Nuyts et al., 2002). For both reconstruction methods, the finite

resolution was included into the projector and back-projector. An iteration scheme with a decreasing number of ordered subsets (Hudson and Larkin, 1994), consisting of (subsets × iterations): 30 × 5, 24 × 5, 20 × 5, 15 × 5, 12 × 5, 10 × 5, 8 × 5, 6 × 5, 4 × 5, 3 × 5, 2 × 5, 1 × 5. This scheme is approximately equivalent to 675 regular iterations. The ML reconstructions were post-smoothed using a 2-D Gaussian kernel with an isotropic FWHM of 5 voxels. Along the equator, we computed line profiles.

2.2.2 Phantom data, acquisition and ML reconstruction

We constructed a clinically realistic 3-D software phantom for subsequent Monte-Carlo simulation experiments. The construction of the phantom was based on 30 consecutive axial slices of the digital software phantom provided by the BrainWeb database at McGill University (Montreal, Canada), `http://www.bic.mni.mcgill.ca/brainweb/`. The software phantom dimensions are 218 × 218 × 30 voxels with an isotropic voxel size of 1 mm. GM, WM, and CSF were identified by the discrete anatomical model as described by Collins et al. (1998). In that model, each voxel was classified as either GM, WM, CSF, or other. A baseline FDG-PET tracer distribution, representing the glucose metabolism of a normal human brain, was constructed. Tracer activity uptake was set to 12.5 counts/mm^3 in GM, 3.125 counts/mm^3 in WM, and no uptake in CSF. In a duplicate of the baseline phantom, called the hypometabolic phantom, tracer activity uptake in four 3-D regions within GM was decreased by 25 %.

The PET data acquisition process of a clinical PET scanner was simulated by projecting the baseline and the hypometabolic phantoms over 144 angles using a uniform attenuation image derived from the contour of the digital BrainWeb phantom. The linear attenuation coefficient was set to 0.095 cm^{-1}. The projections were smoothed along the detector grid using a 2-D Gaussian kernel with 5 mm FWHM. Thereby, we accounted for a realistic detector resolution effect of a clinical PET scanner. The projections were resampled to a realistic sampling frequency of 2 × 2 mm^2 detector elements using linear interpolation. Poisson noise realizations were computed for each sinogram.

The projection data were reconstructed using the ML reconstruction algorithm for further comparison with the A-MAP reconstructions. During the reconstruction process, we included correction for the attenuation and the finite resolution effect. An iteration scheme with a decreasing number of ordered subsets was used, which was consisting of (subsets × iterations): 36 × 6, 24 × 6, 18 × 6, 16 × 6, 12 × 6, 9 × 6, 8 × 6, 6 × 6, 4 × 6, 3 × 6, 2 × 6, 1 × 6. This scheme is approximately equivalent to 834 regular iterations. The ML reconstructions were post-smoothed using a 3-D Gaussian kernel with an isotropic FWHM varying from 4 to 8 mm.

In the following three simulation experiments, we study the influence of the strength of anatomy based smoothing in GM, the influence of misaligned MR data, and the effect of local segmentation errors. The construction of the anatomical information is given in the description of each experiment. Unless stated otherwise, we used the following parameter settings for the A-MAP reconstruction algorithm in the simulation experiments. The segmentation threshold parameter was set to $\varepsilon = 0.01$, the weights of the Gaussian priors were $\beta_{\mathbb{W}} = \beta_{\mathbb{C}} = \beta_{\mathbb{R}} = 0.4$, and the weight of the Gibbs smoothing prior was $\beta_{\mathbb{G}} = 10$. For the local prior weight function Φ, we used

the relative difference prior (Nuyts et al., 2002). Correction for the attenuation and the finite resolution effect, as well as the iteration scheme, was identical to that of the ML reconstruction algorithm.

2.2.3 Experiment 1: influence of anatomical prior weight

A-MAP has been based on the assumption that changes of the cerebral glucose metabolism, due to lesions, should only occur in GM. No anatomical information about the presence, location, or extent of possible lesions is used during the reconstruction process. Consequently, a strong smoothing prior within GM could reduce the visibility of lesions. For that reason, we investigated the influence of smoothing within GM for A-MAP, i.e. the strength of the anatomical prior weight, and compared that to the influence of post-smoothing in ML using 3-D simulated FDG-PET data.

The projection data were reconstructed using A-MAP. For that purpose, the ML reconstructions post-smoothed with 4 mm FWHM were used as the initial image for the iterative process of A-MAP. The anatomical information was based on fuzzy tissue classification data of the GM, WM, and CSF, provided by the BrainWeb database (Collins et al., 1998). The fuzzy tissue classification data was resampled to the PET grid using linear interpolation. Dimensions are $109 \times 109 \times 15$ voxels and an isotropic voxel size of 2 mm. The emission phantom data and the corresponding anatomical information are by design perfectly aligned. A-MAP reconstructions were performed using $\beta_{\mathbb{G}} = 5$, $\beta_{\mathbb{G}} = 10$, and $\beta_{\mathbb{G}} = 15$, denoted A-MAP-1, A-MAP-2, and A-MAP-3, respectively.

2.2.4 Experiment 2: effect of small misalignments

Differences in geometrical alignment are typical for images of the same patient using two different imaging modalities. Using state-of-the-art image registration algorithms, it is possible to compute the affine transformation that is needed to align the MR with the PET data. The influence of errors in the alignment of the anatomical information was assessed.

Usually, registration errors are small (Meyer et al., 1997). We limited this experiment to: (a) a transaxial shift of the "fuzzy" tissue classification data over 1 mm to the right, and (b) an axial rotation over 1°. We considered the size of these errors realistic for current state-of-the-art registration algorithms. The anatomical information was resampled to the PET grid using linear interpolation after applying the transformation.

2.2.5 Experiment 3: effect of segmentation errors

It is clear that the performance of A-MAP will depend on the accuracy of the anatomical information. For that purpose, we assessed the influence of small segmentation errors. To speed up the Monte-Carlo simulations, we used a smaller 3-D brain software phantom, i.e. a subset of the digital BrainWeb phantom in the surroundings of the middle frontal gyrus. The phantom dimensions were $68 \times 68 \times 26$

voxels with an isotropic voxel size of 1 mm. GM and WM tissue and CSF were identified in the corresponding discrete anatomical model. For all slices, the phantom tissue information outside an inscribed circle was put to zero in order to avoid truncated projection data. The binary tissue segmentation images were smoothed using a 3-D isotropic Gaussian kernel with 1.5 mm FWHM to simulate the output of fuzzy segmentation algorithms. A baseline FDG-PET tracer distribution, representing the glucose metabolism of a normal brain, was constructed.

The PET data acquisition procedure used in the previous experiments was applied. Poisson noise realizations of the projections were computed, and the projection data was reconstructed using the ML algorithm. The reconstructions were corrected for attenuation and resolution effect, and the ML iteration scheme defined in section 2.2.2 was used. The ML reconstructions were post-smoothed using a 3-D isotropic Gaussian kernel with a FWHM varying from 4 to 8 mm.

Three types of segmentation errors, often observed in segmentation images of clinical MR data, were semi-automatically constructed using standard morphological dilation and erosion operations. In the discrete tissue segmentations of GM, WM, and CSF, we simulated (a) the "closing" of a sulcus, (b) the "thinning" of a gyrus, and (c) the "thickening" of a gyrus. The modified binary tissue maps were smoothed using a 3-D isotropic Gaussian kernel with 1.5 mm FWHM, to simulate fuzzy segmentations, and resampled to the PET grid using linear interpolation, consisting of $34 \times 34 \times 13$ voxels and an isotropic voxel size of 2 mm. The A-MAP reconstructions were performed using the non-modified and the modified GM, WM and CSF fuzzy tissue classifications.

2.2.6 Evaluation measures

The performance of each reconstruction algorithm was measured using a number of metrics. For notational convenience, we use the following symbols: B^{nl} and H^{nl} for the reconstructions of the noiseless PET projection data of the baseline and hypo-metabolic phantom, respectively; $B^{(r)}$ and $H^{(r)}$ for the reconstructions of noise realization r of the PET projection data of the baseline and hypo-metabolic phantom, respectively; and B^{true} and H^{true} for the "true" baseline and hypo-metabolic phantom, respectively.

A non-prewhitening numerical computer observer was used for the quantification of the detection of the induced hypometabolic regions (Wagner and Brown, 1985). For each reconstruction algorithm, and for each reconstruction of a noise realization r of the projection data, we computed

$$s(I^{(r)}) = \sum_{j \in R} \left(B_j^{\mathrm{nl}} - H_j^{\mathrm{nl}} \right) I_j^{(r)} \tag{11}$$

for $I = B$ and for $I = H$, and where R is a region of interest (ROI). The signal-to-noise ratio (SNR) in R was computed using

$$\mathrm{SNR} = \sqrt{2 \frac{\left[\bar{s}(B) - \bar{s}(H) \right]^2}{\sigma_s^2(B) + \sigma_s^2(H)}} \tag{12}$$

where $\bar{s}(I)$ is the mean and $\sigma_s^2(I)$ the variance of $s(I^{(r)})$ over the reconstructions of

all noise realizations of the projection data.

We measured the partial volume correction performance of each reconstruction algorithm by computing the recovery of activity in every reconstructed voxel. The recovery value ρ was calculated as the ratio of the reconstructed activity divided by the true activity. We computed the mean μ_ρ and the standard deviation $\tilde{\sigma}_\rho$ of the recovery values within ROIs for the reconstructions based on the baseline and the hypometabolic phantom data, using

$$\mu_\rho(I) = \frac{1}{n_R} \sum_{j \in R} \left(\frac{1}{P} \sum_{r=1}^{P} \frac{I_j^{(r)}}{I_j^{\text{true}}} \right) \tag{13}$$

$$\tilde{\sigma}_\rho(I) = \sqrt{\frac{1}{n_R} \sum_{j \in R} \left[\frac{1}{P} \sum_{r=1}^{P} \left(\frac{I_j^{(r)} - \bar{I}_j}{I_j^{\text{true}}} \right)^2 \right]} \tag{14}$$

for $I = B$ and for $I = H$, where $\bar{I}_j$ is the mean of $I_j^{(r)}$ over the reconstructions of all noise realizations r of the projection data, P is the number of noise realizations, and n_R is the number of voxels in ROI R. Additionally, we measured the distribution of the recovery values within GM.

2.3 Human observer study

A human observer study was designed to assess the performance for detecting hypometabolic regions using the A-MAP versus the ML reconstruction algorithm. Six nuclear medicine physicians and residents-in-training were asked to localize regions with a reduced metabolism in 2-D simulated FDG-PET images of the brain. In order to make the observations as realistic as possible, but on the other hand not too time consuming, we constructed a 2-D realistic software phantom based on a single transaxial slice of the digital BrainWeb phantom. Dimensions are 218×218 voxels with an isotropic voxel size of 1 mm. GM, WM and CSF tissues were identified using the discrete anatomical model. A baseline FDG-PET tracer distribution representing the glucose metabolism of a normal brain was constructed. Throughout the 2-D phantom, we delineated 25 regions which were restricted to the GM tissue. The mean size of the regions was 3.7 cm^2 ($\sigma = 0.4$ cm^2). For each of these regions, we constructed 20 hypometabolic phantoms in which the region had a proportional metabolic reduction value ranging from 5 % to 100 % in steps of 5 %, defined as the contrast level. All phantoms were resampled to a realistic PET sampling grid of 109×109 voxels and an isotropic voxel size of 2 mm using linear interpolation. The baseline tracer activity was set to 12.5 counts/mm^3 in the GM tissue, 3.125 counts/mm^3 in the WM and 0 counts/mm^3 in the CSF. The PET data acquisition procedure used in the previous experiments was applied. It was left to the observing nuclear medicine physician to control the amount of post-smoothing required for further observation of the ML reconstruction images.

A clinically realistic high-resolution noisy MR image, corresponding to the digital BrainWeb phantom, was generated using the BrainWeb MR simulator at McGill University (Montreal, Canada), `http://www.bic.mni.mcgill.ca/brainweb/`. The

noise level was 7 %. The noisy MR image was segmented into GM, WM and CSF tissues using the Statistical Parametric Mapping (SPM99) software package, Wellcome department of cognitive neurology (London, UK), `http://www.fil.ion.ucl.ac.uk/spm`. The fuzzy tissue segmentation images were resampled to the PET grid using linear interpolation and used in an A-MAP reconstruction of all simulated PET projection data of the baseline and hypometabolic phantoms.

An interactive computer program was written as a user interface for the human observers in order to evaluate the performance of detecting and localizing the induced hypometabolic regions. The user interface was based on in-house developed medical image viewing software that is also used in the daily clinical environment. Physicians were invited to carry out two separate observing sessions, one for the A-MAP reconstructions and one for the ML reconstructions. In a learning session for the observers, the program started with presenting the reconstructed images of noise realizations of the baseline phantom projection data. Observers were asked to get acquainted with the image quality, the influence of noise, and they were asked to tune the display settings. After the learning session, the program presented the reconstructed images of the hypometabolic phantom data in an order of decreasing contrast of the hypometabolic regions. The program selected for each of the 20 contrast levels randomly 10 of the 25 hypometabolic reconstructed images belonging to that specific contrast level. Throughout the session, physicians were able to adjust the width of a 2-D Gaussian smoothing kernel applied during the image displaying, as well as using a variety of color lookup tables, scale parameters, cursor settings, and zoom factors.

The task of the observer was then to specify the most likely position of the hypometabolic region using a mouse click in the image field. The location and distance between true lesion barycenter and observer localization was measured and stored for further analysis. We defined a hit-range h. The hit-range is the radius of an imaginary sphere surrounding the true lesion barycenter. If the indicated position by the observer was within the imaginary sphere, the localization was considered to be correct. After every 10 images, the observer received feedback about the number of correct localizations for a hit-range of $h = 5$ mm. Display of the images was accompanied by the MR image from which the anatomical data was derived. This image served as additional anatomical information during the observation of both ML and A-MAP reconstructed images. The cursor was linked between the images. Both the reconstructed image Λ as well as the intermediate reconstructed image Λ^* were shown when the user interface was used for observing A-MAP reconstructions. It was explained in advance to the observers what the A-MAP images represented and how they were created. During the observation session, physicians were able to consult a printout of the phantom image with small dots indicating the 25 possible location sites of the hypometabolic regions. One half of the physicians was asked to complete first the observation of the ML reconstructions, and after that start with the observation of the A-MAP reconstructions. The other half was asked to follow the opposite order. For each reconstruction algorithm, we measured the score $S(h, \mathcal{O}, \mathcal{C})$, defined as the percentage of correct localizations, as function of the hit-range h, the observer $\mathcal{O}$, and the contrast level $\mathcal{C}$ of the induced hypometabolic region. We calculated (a) the overall score, defined as the mean score over all contrast

levels and for $h = 5$ mm, or

$$\text{overall score}\,(\mathcal{O}) = \frac{1}{n_{\mathcal{C}}} \sum_{\mathcal{C}} S(h = 5, \mathcal{O}, \mathcal{C}) \tag{15}$$

where $n_{\mathcal{C}}$ is the number of contrast levels, (b) the average score, defined as the mean score over all observers for $h = 5$ mm, or

$$\text{average score}\,(\mathcal{C}) = \frac{1}{n_{\mathcal{O}}} \sum_{\mathcal{O}} S(h = 5, \mathcal{O}, \mathcal{C}) \tag{16}$$

where $n_{\mathcal{O}}$ is the number of observers, and (c) the average overall score, defined as the mean score over all contrast levels and observers, or

$$\text{average overall score}\,(h) = \frac{1}{n_{\mathcal{C}} \cdot n_{\mathcal{O}}} \sum_{\mathcal{C}} \sum_{\mathcal{O}} S(h, \mathcal{O}, \mathcal{C}) \tag{17}$$

2.4 Patient study

We performed a patient study to illustrate the applicability of A-MAP in clinical practice. We selected FDG-PET and T1-weighted MR data of a patient with well-documented epilepsy of age 41. She developed refractory temporal lobe epilepsy at age 18. Before she entered the presurgical evaluation procedure in our hospital, she suffered each month from about 20 complex partial seizures, for which no etiology was established. Two separate video combined electro-encephalography (video-EEG) sessions during and between seizures showed left temporal epileptogenic EEG patterns. The patient underwent left temporal lobe resection, including the amygdala but not the hippocampus. The patient has been free of seizures since surgery during a follow-up period of 2 years.

Presurgical FDG-PET acquisitions were performed on a CTI-Siemens ECAT EXACT HR+ scanner (Brix et al., 1997). The patient received a catheter inserted into the left brachial vein for tracer administration, and was positioned on the scanner bed. The head was immobilized and centered in the field-of-view. EEG monitoring was installed and recorded throughout the acquisition process to confirm the interictal condition of the patient. A transmission scan of 10 min, using ^{68}Ge rod sources, was acquired in 2-D mode to correct for photon attenuation. Thereafter, 150 MBq FDG was injected intravenously, and a 3-D mode emission scan of 30 min was started at 30 min post-injection.

The transmission scan was reconstructed using a maximum-a-posteriori reconstruction algorithm (Nuyts et al., 1999). Randoms correction of the 3-D emission scan was performed on-line using a delayed time window. The reconstructed attenuation image was used to apply a scatter correction method to the 3-D emission data (Watson, 2000), followed by the Fourier Rebinning (FORE) algorithm (Defrise, 1995). The obtained 2-D projection data were reconstructed using ML (Shepp and Vardi, 1982). The dimensions of the reconstructed image are $110 \times 110 \times 63$ voxels with an in-plane pixel size of 2.25×2.25 mm^2 and a plane separation of 2.425 mm. The ordered subsets technique was used (Hudson and Larkin, 1994). An iteration scheme with a decreasing number of subsets and an increasing number of iterations was used,

consisting of (subsets × iterations): 36×5, 24×5, 18×5, 16×6, 12×6, 9×7, 8×7, 6×8, 4×8, 3×9, 2×9, 1×10. This scheme is approximately equivalent to 812 regular iterations. We included correction for attenuation and a finite resolution of 5 mm FWHM (Brix et al., 1997) during the reconstruction process.

Four T1-weighted MR acquisitions of the patient were obtained consecutively on a Siemens 1.5 Tesla MAGNETOM Vision scanner using the Magnetization Prepared Rapid Gradient Echo (3-D MPRAGE) sequence (Mugler and Brookman, 1990). The following parameter settings were used: TR = 10 msec, TE = 4 msec, flip angle = 12°, inversion time = 300 msec, 256×256 matrix size, 160 sagittal slices, and 1 mm isotropic voxel size. All 4 MR images were spatially aligned using a mutual information based image registration algorithm (Maes et al., 1997), and the mean MR image was computed. The MRI data were visually examined by an expert radiologist. The MRI data were reported to show no abnormalities.

Segmentation of the GM, WM and CSF tissues was performed using the Statistical Parametric Mapping (SPM2) software package, Wellcome department of cognitive neurology (London, UK), `http://www.fil.ion.ucl.ac.uk/spm`. The affine transformation was computed to align the MR image with the post-smoothed ML reconstruction of the PET data. Then, the same transformation was used to align the segmentation images of the MRI data with the ML reconstruction. The anatomical information was resampled to the PET grid using linear interpolation. The ML reconstruction, post-smoothed using a Gaussian kernel with 5 mm FWHM, was used as the initial image for the A-MAP iterative reconstruction. We used the same iteration scheme, attenuation correction and resolution recovery, as for the ML reconstruction.

3 Results

3.1 Simulation experiments

3.1.1 A-MAP and partial volume correction

The phantom used for illustrating the PVE is shown in Fig. 1(a). A line profile of the phantom image through the higher intensity regions is shown in Fig. 1(b). In Fig. 1(c), the corresponding line profile through the fuzzy "GM" segmentation image is shown. A line profile through the ML reconstruction without post-smoothing and the post-smoothed ML reconstruction using 5 voxels FWHM are shown in Fig. 1(d). A line profile through the A-MAP reconstruction is shown in Fig. 1(e). The mean and the standard deviation of the recovery values within the "GM" regions are: 0.97 ± 0.05 for A-MAP, 0.95 ± 0.11 for ML without post-smoothing, and 0.89 ± 0.10 for post-smoothed ML using 5 voxels FWHM.

3.1.2 Data construction and acquisition

The construction of the phantom data, the implementation of the A-MAP and ML reconstruction algorithms, as well as the simulation experiments were developed and

performed on a Sun450 Model-4300 workstation (SUN Microsystems, Inc.) using the software package IDL 5.1 (Research Systems, Inc.). The location and the extent of the induced hypometabolic regions used in the phantom data are shown in Fig. 2. The simulated noiseless PET emission data of the 3-D baseline phantom contained 117 Mcounts. We computed 100 Poisson noise realizations of the PET projection data of the baseline and 100 of the hypometabolic phantom.

3.1.3 Experiment 1: influence of anatomical prior weight

Fig. 3 shows the mean recovery value (13) results plotted against the standard deviation of the recovery values (14) for the ML reconstructions and for the A-MAP reconstructions using $\beta_{\mathbb{G}} = 5.0$, 10.0, and 15.0. The upper panel of Fig. 3 shows the results for the simulations based on the baseline phantom data, and computed within an overall 3-D ROI, defined by the GM and WM tissues in the phantom data for the 5 most central planes. The lower panel of Fig. 3 shows the results for the simulations based on the hypometabolic phantom data, and computed within four 3-D ROI around the corresponding hypometabolic regions.

A-MAP shows an overall improved mean recovery compared to post-smoothed ML for the reconstructions of the baseline phantom data. A similar improvement is seen in the mean recovery values in the ROIs around the hypometabolic regions of the reconstructions of the hypometabolic phantom data. Although the results depend on the ROI that was used, as can be seen in the lower panel of Fig. 3, the above pattern is the same for each ROI. ML without post-smoothing (including resolution recovery) shows comparable mean recovery values as A-MAP, but with a high standard deviation of these values.

The SNR results for post-smoothed ML and A-MAP for different ROI are shown in Table 1. The SNR of post-smoothed ML increases with increasing FWHM in the range from 0 to 8 mm. On the other hand, in all ROI, the SNR is almost unchanged for A-MAP when the prior weight is varied between $\beta_{\mathbb{G}} = 5.0$ and 15.0.

Fig. 4 shows the distribution of the recovery values within GM for ML without post-smoothing, a post-smoothed ML reconstruction using a Gaussian kernel with 5 mm FWHM, and an A-MAP reconstruction using $\beta_{\mathbb{G}} = 10.0$, of a single noise realization of the baseline projection data. The mean and the standard deviation of the recovery values are: 0.89 ± 0.15 for A-MAP, 0.90 ± 0.24 for ML without post-smooth, and 0.77 ± 0.13 for post-smoothed ML using 5 mm FWHM.

3.1.4 Experiment 2: effect of small misalignments

The mean and the standard deviation of the recovery values for A-MAP using misaligned anatomical information are shown in Fig. 5. The results for post-smoothed ML and A-MAP using perfectly aligned anatomical information, i.e. those of experiment 1, were added for comparison.

The use of shifted anatomical information in the A-MAP reconstructions shows a reduction of the mean recovery value in the overall ROI and in the ROI around the hypometabolic regions, but has little effect on the standard deviation of the recovery.

The use of rotated anatomical information shows a similar pattern. The SNR results for A-MAP using misaligned anatomical information are shown in Table 1. The results of the SNR in all ROI are comparable with the SNR for A-MAP using perfectly aligned anatomical information.

Fig. 6 shows the distribution of the recovery values within GM, as defined by the baseline phantom, for ML without post-smoothing, a post-smoothed ML reconstruction using a Gaussian kernel with 5 mm FWHM, and an A-MAP reconstruction using shifted anatomical information, of a single noise realization of the baseline projection data. The mean and the standard deviation of the recovery values are 0.84 ± 0.23 for A-MAP using shifted anatomical information. The results for ML are identical to those obtained in experiment 1 (see section 3.1.3).

3.1.5 Experiment 3: effect of segmentation errors

The simulated GM fuzzy tissue segmentation images, including the segmentation errors, are shown in Fig. 7. The noiseless projection data of the smaller 3-D phantom contained 12 Mcounts. Hundred Poisson noise realizations of the projection data were computed. The mean and standard deviation of the recovery values for assessing the influence of local segmentation errors within the anatomical information are shown in Fig. 8.

The influence of segmentation errors in A-MAP shows a small reduction of the mean recovery value in the overall ROI. A similar reduction is observed in the ROI around the modification sites. However, the mean recovery values are still better than those obtained by post-smoothed ML.

Fig. 9 shows the recovery values within GM for A-MAP using exact and modified anatomical information, and for post-smoothed ML, using a single Gaussian kernel with 5 mm FWHM, of a noise realization of the projection data.

3.2 Human observer study

The results of the human observer study are shown in Fig. 10. The A-MAP and ML average overall score, i.e. the proportional score over all contrast levels and observers, as function of the hit-range h, are shown in Fig. 10(a). Both curves level off at a hit-range of about 5 mm. The A-MAP and ML average score, i.e. the proportional score over all observers for $h = 5$ mm, and as function of the contrast level of the hypometabolic regions, are shown in Fig. 10(b). Since there were 25 possible locations for the hypometabolic region, the probability of obtaining 2 or more correct answers by chance alone is 0.058. Fig. 10(c) shows the overall score, i.e. the proportional score over all contrast levels, for $h = 5$ mm, for each observer, and for using ML and A-MAP. For all observers, there is a systematic improvement of detection performance using the A-MAP algorithm. The ML algorithm shows a much larger variance than A-MAP. A Paired t-Test of the overall score data of ML (average overall score for $h = 5$ mm is 68.4 %) versus A-MAP (average overall score for $h = 5$ mm is 81.8 %) showed a significantly higher average overall score value for A-MAP ($p < 0.004$).

3.3 Patient study

The results of the patient study are shown in Fig. 11. A coronal slice through the temporal lobe was selected for display (panel A in Fig. 11). The A-MAP reconstruction image is shown for voxels j where $f_j^G + f_j^W + f_j^C > \varepsilon$, i.e. voxels belonging to the brain. The PET images show clearly a left temporal hypometabolism, which is concordant with the findings of the presurgical video-EEG registrations. We computed a line profile through the right and left temporal lobes of the post-smoothed ML and A-MAP reconstructed images. The line profile is shown in panel B of Fig. 11 and indicated in the coronal views.

4 Discussion

4.1 Tissue composition model and partial volume correction

This study aimed at the evaluation of A-MAP, an iterative reconstruction algorithm that was developed for PVC during the reconstruction of FDG-PET brain projection data.

A-MAP was originally designed for FDG-PET imaging of epilepsy patients (Baete et al., 2004). The aim was to improve the detection of hypometabolic regions in the presurgical evaluation of FDG-PET images. For that purpose, we have looked at the use of anatomical information derived from the MR data to correct for the partial volume effect. We used an approximation of a tissue composition model in order to compute the GM activity in every voxel, while using prior assumptions about the WM and CSF activity distribution, see equation (2). This concept forms the basis for the PVC technique described by Müller-Gärtner et al. (1992). As a post-processing method, this method has a major drawback, because it lacks the use of a noise model. Aston et al. (2002) developed a more general mathematical framework for PVC after the reconstruction process, in which explicit assumptions about the noise model are used. Recently, we compared the noise characteristics of A-MAP and post-processed ML for the incorporation of anatomical information (Nuyts et al., 2003). Post-processed ML showed inferior performance. If a prewhitening step was introduced, results became comparable. Unfortunately, prewhitening involves the application of a shift variant filter which makes the implementation more complicated than that of A-MAP.

For A-MAP, we have used the approximation of a tissue composition model to carry out a parameter transformation of the image variables, see equation (3). The new set of variables were included in a Bayesian reconstruction framework. A similar approach was described by Sastry and Carson (1997). The parameter transformation was used to construct an adjusted projection and back-projection operator. For that purpose, we derived a new system matrix, see equation (4). The correction for the partial volume effect of WM and CSF in the GM region can be seen in the third and fourth term of (4). Our approach has the advantage that noise suppression within GM has been restricted to GM activity only, even if the voxels contain a mixture of GM and WM. In that way, smoothing over tissue mixtures is avoided, as

opposed to other techniques, where noise suppression is applied to the total voxel activity (Fessler et al., 1992; Ardekani et al., 1996; Bowsher et al., 1996; Lipinski et al., 1997; Comtat et al., 2002). Moreover, because the GM glucose metabolism is locally varying throughout the brain, and because lesions can change the local FDG uptake, we applied a Gibbs smoothing prior that tends to preserve strong edges (Nuyts et al., 2002). We used additional regularization in true WM and true CSF regions, by imposing a Gaussian activity distribution in each region.

The PVE as function of structure size and intensity was illustrated using a simple 2-D software phantom simulation (Fig. 1). Inclusion of the finite resolution into the projector and back-projector of an iterative reconstruction algorithm leads to the compensation by the algorithm for the blurring (i.e. the resolution recovery). However, perfect deblurring is impossible, and the ML reconstruction algorithm may produce over- and undershoot artifacts (Snyder et al., 1987), as can be seen in Fig. 1(d). These artifacts disappear with post-smoothing, which is required to suppress the noise. Unfortunately, post-smoothing leads to an underestimation of the true intensity value for the smaller structures inside the phantom. In A-MAP, the deblurring problem is stabilized using anatomical based smoothing. The over- and undershoot artifacts are strongly reduced, which results in a superior resolution recovery.

Using the 3-D BrainWeb software phantom, we measured and compared the recovery values of A-MAP and post-smoothed ML. We measured how far the reconstruction was deviating from the true solution using mean recovery values, and how stable that solution was under the influence of noise using the standard deviation of the recovery values. The results showed an overall improved recovery of the activity using A-MAP, as compared to ML. A-MAP showed to be robust for variation of the weight of smoothing prior in the GM region. Within the hypometabolic regions, we observed similar recovery characteristics, indicating that the recovery improvement for A-MAP was not critically depending on the amount of smoothing within GM, at least for the range tested in the simulation experiments.

We have demonstrated the partial volume correction properties of A-MAP and post-smoothed ML using GM recovery values. The distribution of the recovery values for the reconstruction of a noise realization showed a clear underestimation of the GM activity for post-smoothed ML, whereas the use of A-MAP improved the GM recovery values. The ML reconstruction without post-smoothing showed less underestimation of the GM activity, but an increased spread of the recovery values (Fig. 4).

4.2 Registration and segmentation errors

Two important steps in the A-MAP procedure are the registration between an initial PET image and the MR image, and the segmentation of the MRI data. For small registration errors, we observed a reduction of the mean recovery values at a comparable level of standard deviation. The influence of segmentation errors was more limited. The use of shifted anatomical information introduced local underestimations of the GM activity. This can be seen in the left tail of the distribution of the recovery values, which is more extended towards lower values (Fig. 6).

We measured lesion detection performance using the SNR measure, reflecting the performance of a computer observer. With the simulation experiments, we observed that the SNR results for A-MAP were lower than the optimal SNR results for post-smoothed ML. For ML, the optimal results corresponded to large FWHM of the post-smoothing kernel. In the human observer study, we observed that the detection was improved using A-MAP, as opposed to post-smoothed ML. These results reflect the differences between computer observers and human observers for lesion detection.

4.3 Detection of hypometabolic regions

We hypothesized that physicians would demonstrate an improved detection of real hypometabolic regions, as opposed to regions with a decreased intensity due to PVE, using the A-MAP reconstruction algorithm. To test this hypothesis, we investigated the performance of visually detecting hypometabolic regions in A-MAP reconstructions compared to post-smoothed ML reconstructions by means of a human observer study.

Human observers were asked to specify the most likely position of the hypometabolic region, knowing that there was only one lesion in the image. This approach is also known as a multiple alternative forced choice experiment (Burgess, 1995). As opposed to receiver operating characteristic analysis (ROC), or localizing ROC analysis (Gifford et al., 2000), our method avoids the use of a confidence rating or an explicit threshold which makes the procedure easier for the observers (Nuyts et al., 1999).

If available, MRI data are often used in clinical practice for the examination of FDG-PET. We included the MR image in the user interface during the observations of both A-MAP and post-smoothed ML to make sure that differences found between the two methods could not be attributed to the fact whether the MRI was present or not.

The average overall score as function of the hit-range levelled off at about 5 mm for both A-MAP and ML. This means that from that distance on, the score is not critically depending on the distance between true lesion barycenter and the observer indicated position. The average score of the observers for A-MAP as function of contrast, going from 100 % to about 80 %, showed a small initial increase. This is probably a residual learning effect. This effect was not observed for post-smoothed ML, because physicians are more used to observe that kind of images.

The overall score for A-MAP was for each observer higher than that for ML, indicating a benefit of using the A-MAP reconstruction method for the detection of hypometabolic regions in FDG-PET images. Furthermore, the ML algorithm showed a larger variance for the overall score than A-MAP. This could suggest that the experience of the observer is more important for ML than for A-MAP.

4.4 Clinical application

The patient study was used to illustrate the feasibility of the A-MAP reconstruction method applied to measured FDG-PET projection data. Interestingly, A-MAP

provides an intermediate image showing the GM activity distribution λ^G inside GM region $\mathbb{G}$. Further evaluation of A-MAP should reveal whether this information improves the diagnostic use of FDG-PET data for a range of different pathologies (e.g. epilepsy, dementia, movement disorders), and if the approach is applicable to other radioligands. The A-MAP reconstruction, shown in Fig. 11, was restricted to the brain voxels only. Outside the brain, the A-MAP reconstruction algorithm is equivalent to unconstrained ML algorithm, in which noise was not suppressed.

5 Conclusion

We can conclude that partial volume correction during the reconstruction process leads to a more accurate detection of hypometabolic regions in FDG-PET of the brain. The A-MAP reconstruction algorithm shows to be robust for small registration and segmentation errors. We demonstrated the applicability of A-MAP on real patient data.

Acknowledgment

The authors wish to thank Dirk Vandermeulen, Siddharth Srivastava, Dirk Bequé, and André Palmini for useful discussions regarding this study.

References

Ardekani, B., Braun, M., Hutton, B., Kanno, I., Iida, H., 1996. Minimum cross-entropy reconstruction of PET images using prior anatomical information. Phys. Med. Bio. 41, 2497–2517.

Aston, J., Cunningham, V., Asselin, M.-C., Hammers, A., Evans, A., Gunn, R., 2002. Positron emission tomography partial volume correction: estimation and algorithms. J. Cereb. Blood Flow Metab. 22 (8), 1019–1034.

Baete, K., Nuyts, J., Van Paesschen, W., Suetens, P., Dupont, P., 2004. Anatomical-based FDG-PET reconstruction for the detection of hypo-metabolic regions in epilepsy. IEEE Trans. Med. Imag. 23 (4), 510–519.

Bowsher, J., Johnson, V., Turkington, T., Jaszczak, R., Floyd, C., Coleman, R., 1996. Bayesian reconstruction and use of anatomical a priori information for emission tomography. IEEE Trans. Med. Imag. 15 (5), 673–686.

Brix, G., Zaers, J., Adam, L., Bellemann, M., Ostertag, H., Trojan, H., Haberkorn, U., Doll, J., Oberdorfer, F., Lorenz, W., 1997. Performance evaluation of a whole-body PET scanner using the NEMA protocol. J. Nucl. Med. 38 (10), 1614–1623.

Burgess, A., 1995. Comparison of receiver operating characteristic and forced choice observer performance measurement methods. Med. Phys. 22 (5), 643–655.

Collins, D., Zijdenbos, A., Kollokian, V., Sled, J., Kabani, N., Holmes, C., Evans, A., 1998. Design and construction of a realistic digital brain phantom. IEEE Trans. Med. Imag. 17 (3), 463–468.

Comtat, C., Kinahan, P., Fessler, J., Beyer, T., Townsend, D., Defrise, M., Michel, C., 2002. Clinically feasible reconstruction of 3d whole-body PET/CT data using blurred anatomical labels. Phys. Med. Bio. 47, 1–20.

Defrise, M., 1995. A factorization method for the 3D X-ray transform. Inverse Problems 11 (5), 983–994.

Fazio, F., Perani, D., 2000. Importance of partial-volume correction in brain PET studies. J. Nucl. Med. 41 (11), 1849–1850.

Fessler, J., Clinthorne, N., Rogers, W., 1992. Regularized emission image reconstruction using imperfect side information. IEEE Trans. Nucl. Sc. 39 (5), 1464–1471.

Fischl, B., Dale, A., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. USA 97 (20), 11050–11055.

Gifford, H., King, M., Wells, R., Hawkins, W., Narayanan, M., Pretorius, P., 2000. LROC analysis of detector-response compensation in SPECT. IEEE Trans. Med. Imag. 19 (5), 463–473.

Gindi, G., Lee, M., Rangarajan, A., Zubal, I., 1993. Bayesian reconstruction of functional images using anatomical information as priors. IEEE Trans. Med. Imag. 12 (4), 670–680.

Hudson, H., Larkin, R., 1994. Accelerated image reconstruction using ordered subsets of projection data. IEEE Trans. Med. Imag. 13 (4), 601–609.

Labbé, C., Froment, J., Kennedy, A., Ashburner, J., Cinotti, L., 1996. Positron emission tomography metabolic data corrected for cortical atrophy using magnetic resonance imaging. Alzheimer Dis. Assoc. Disord. 10 (3), 141–170.

Lipinski, B., Herzog, H., Rota Kops, E., Oberschelp, W., Müller-Gärtner, H., 1997. Expectation maximization reconstruction of positron emission tomography images using anatomical magnetic resonance information. IEEE Trans. Med. Imag. 16 (2), 129–136.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P., 1997. Multimodality image registration by maximization of mutual information. IEEE Trans. Med. Imag. 16 (2), 187–198.

Meltzer, C., Leal, J., Mayberg, H., Wagner, Jr., H., Frost, J., 1990. Correction of PET data for partial volume effects in human cerebral cortex by MR imaging. J. Comput. Assist. Tomogr. 14 (4), 561–570.

Meyer, C., Boes, J., Kim, B., Bland, P., Zasadny, K., Kison, P., Koral, K., Frey, K., Wahl, R., 1997. Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations. Med. Image Anal. 1 (3), 195–206.

Mugler, J., Brookman, J., 1990. Three-dimensional magnetization-prepared rapid gradient-echo imaging (3-D MPRAGE). Magn. Reson. Med. 15 (1), 152–157.

Müller-Gärtner, H., Links, J., Prince, J., Bryan, R., McVeigh, E., Leal, J., Davatzikos, C., Frost, J., 1992. Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects. J. Cereb. Blood Flow Metab. 12 (4), 571–583.

Nuyts, J., Baete, K., Bequé, D., Dupont, P., 2003. Comparison between MAP and post-processed ML for incorporating anatomical knowledge in emission tomography. In: Proceedings of the IEEE Medical Imaging Conference. Portland, USA.

Nuyts, J., Bequé, D., Dupont, P., Mortelmans, L., 2002. A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography. IEEE Trans. Nucl. Sc. 49 (1), 56–60.

Nuyts, J., Dupont, P., Stroobants, S., Maes, A., Mortelmans, L., Suetens, P., 1999. Evaluation of maximum-likelihood based attenuation correction in positron emis-

sion tomography. IEEE Trans. Nucl. Sci. 46 (4), 1136–1141.
Rangarajan, A., Hsiao, I.-T., Gindi, G., 2000. A bayesian joint mixture framework for the integration of anatomical information in functional image reconstruction. J. Math. Imag. Vision 12 (3), 119–217.
Sastry, S., Carson, R., 1997. Multimodality bayesian algorithm for image reconstruction in positron emission tomography: a tissue composition model. IEEE Trans. Med. Imag. 16 (6), 750–761.
Shepp, L., Vardi, Y., 1982. Maximum likelihood reconstruction for emission tomography. IEEE Trans. Med. Imag. 1 (2), 113–122.
Snyder, D., Miller, M., Thomas, L., Politte, D., 1987. Noise and edge artifacts in maximum-likelihood reconstructions for emission tomography. IEEE Trans. Med. Imag. MI-6 (3), 228–238.
Wagner, R., Brown, D., 1985. Unified SNR analysis of medical imaging systems. Phys. Med. Bio. 30 (6), 489–518.
Watson, C., 2000. New, faster, image-based scatter correction for 3D PET. IEEE Trans. Nucl. Sc. 47 (4), 1587–1594.

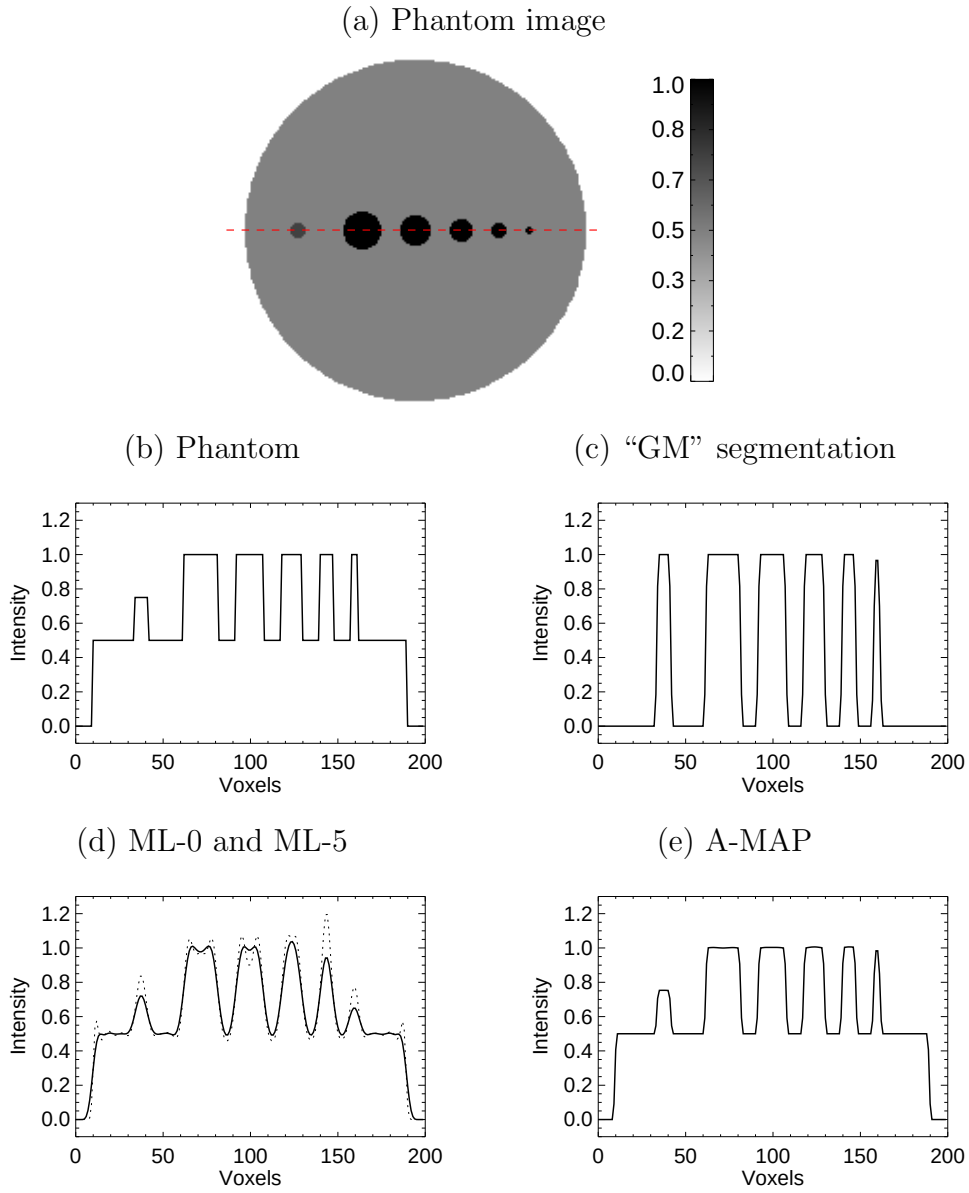
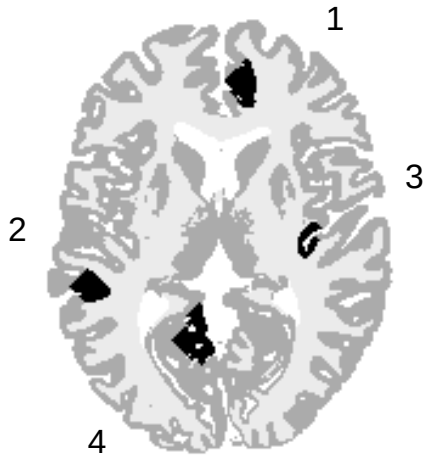


Figure 1. The 2-D phantom image used for illustrating PVE is shown in (a). The plot in (b) shows the line profile through the phantom image along the equator. Plot (c) shows the line profile through the fuzzy segmentation image of "GM". Plot (d) shows the line profile through the ML without post-smoothing, i.e. ML-0 (dotted line), and post-smoothed ML using 5 voxels FWHM, i.e. ML-5 (solid line). Plot (e) shows the line profile through the A-MAP reconstruction.



Region	Size (cm^3)
1	1.5
2	1.2
3	0.8
4	1.7

Figure 2. The central plane of the 3-D hypometabolic phantom that was used in the simulation experiments. The locations of the hypometabolic regions are visualized in black. The size of the regions is given.

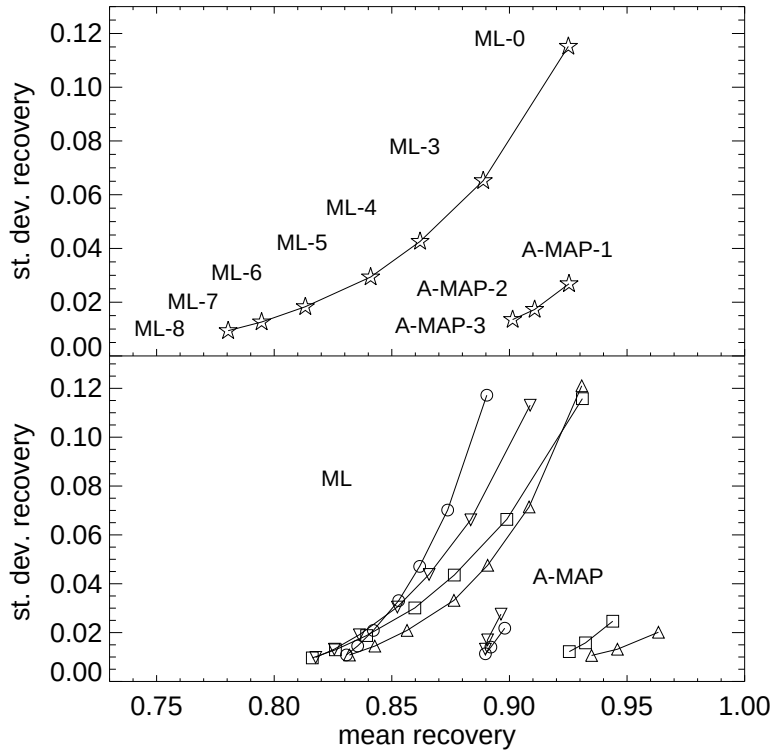


Figure 3. The plots show the results of the mean and the standard deviation of the recovery values of simulation experiment 1 for post-smoothed ML, and for A-MAP using different values for the GM prior weight (A-MAP-1, 2, and 3). ML-X means post-smoothed ML using a 3-D isotropic Gaussian kernel with X mm FWHM. Star shaped symbols in the upper panel show the results for the baseline phantom data, computed in an overall 3-D region-of-interest (ROI). The lower panel shows the results for the hypometabolic phantom data, computed within 3-D ROI around hypometabolic region 1 ($\triangledown$), 2 ($\square$), 3 ($\circ$), and 4 ($\triangle$). For ML and A-MAP, symbols belonging to the same ROI are connected by a solid line.

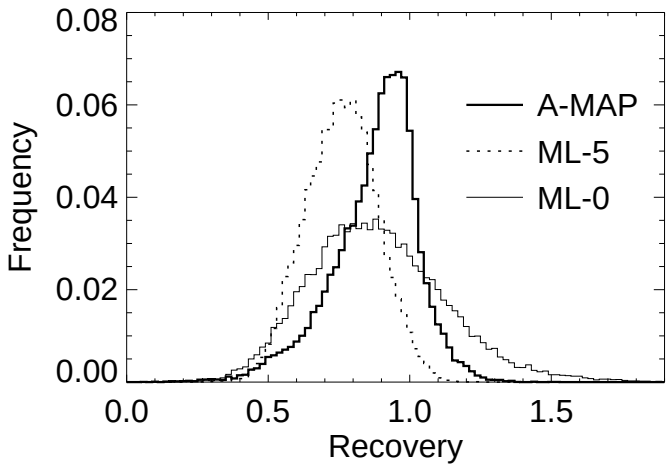
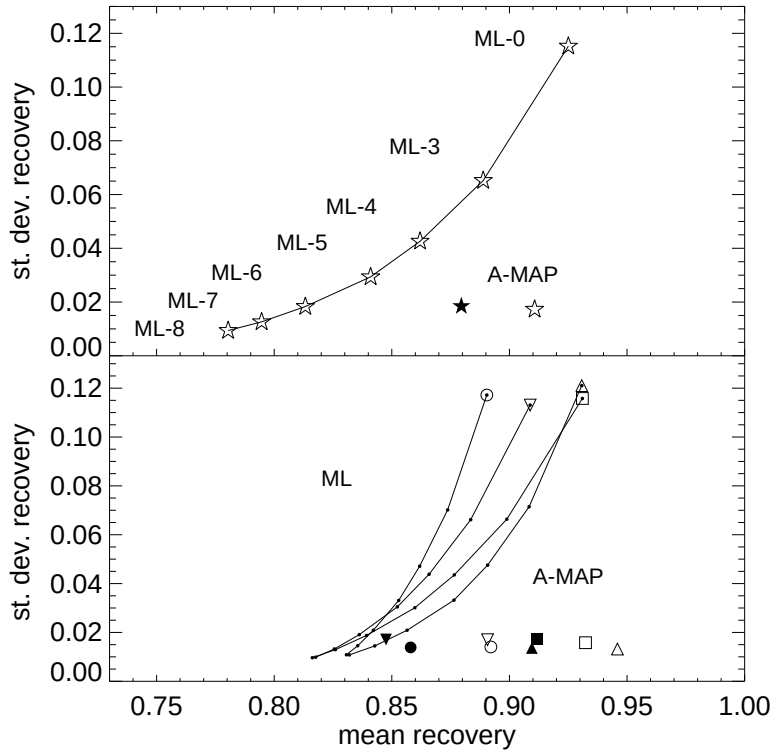


Figure 4. The distribution of recovery values within GM for the ML reconstruction without post-smoothing (ML-0, thin solid line), a post-smoothed ML reconstruction using 5 mm FWHM (ML-5, dotted line), and an A-MAP reconstruction ($\beta_{\mathbb{G}} = 10.0$, thick solid line), of a single noise realization of the baseline phantom projection data.

Table 1
Signal-to-noise ratio (SNR) of experiments 1–3, in four 3-D ROI around the hypometabolic regions. ML-X means post-smoothed ML using a 3-D Gaussian kernel with X mm FWHM.

Algorithm	ROI-1	ROI-2	ROI-3	ROI-4
ML-0	24.6	22.4	14.7	25.5
ML-4	38.0	31.5	21.1	39.5
ML-5	41.6	33.4	23.0	42.6
ML-6	45.5	35.4	25.2	45.3
ML-7	47.7	36.6	26.6	46.6
ML-8	48.9	37.5	27.4	47.2
A-MAP-1	42.3	37.6	23.9	40.8
A-MAP-2	44.9	37.8	23.5	41.0
A-MAP-3	45.4	36.9	23.2	41.2
A-MAP-2 (1 mm shift)	45.9	38.9	23.3	41.7
A-MAP-2 (1° rotation)	45.7	38.2	23.8	42.3

(a)



(b)

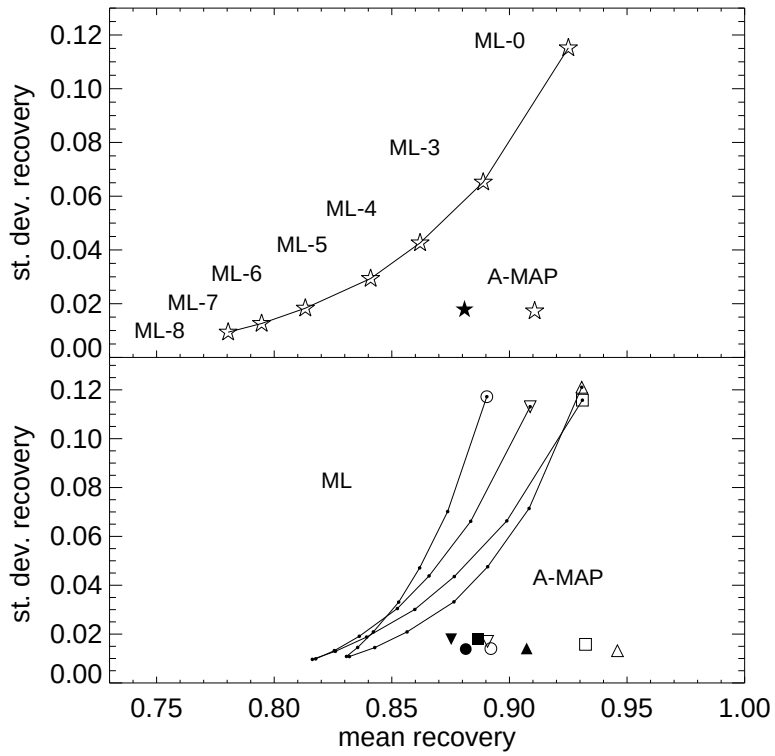


Figure 5. The plots show the results of the mean and the standard deviation of the recovery values of simulation experiment 2 for post-smoothed ML and for A-MAP when using perfectly aligned anatomical information (open symbols) and intentionally misaligned anatomical information (filled symbols). Panel (a) shows the results for shifting the anatomy over 1 mm, and panel (b) for rotating it over 1°. Same conventions as in Fig. 3.

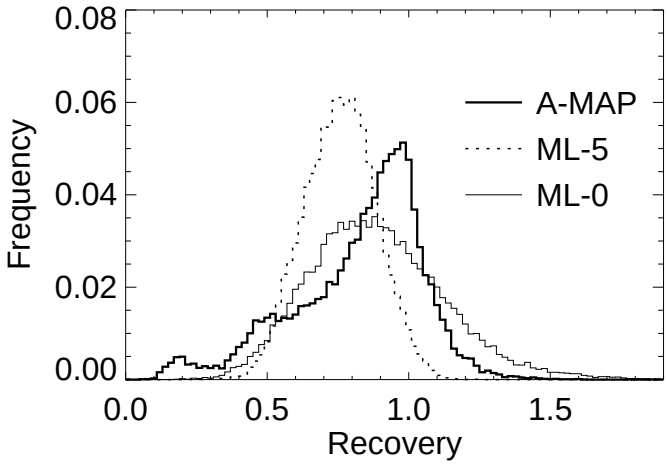


Figure 6. The distribution of recovery values within GM for the ML reconstruction without post-smoothing (ML-0, thin solid line), a post-smoothed ML reconstruction using 5 mm FWHM (ML-5, dotted line), and an A-MAP reconstruction using shifted anatomical information (thick solid line), of a single noise realization of the baseline phantom projection data.

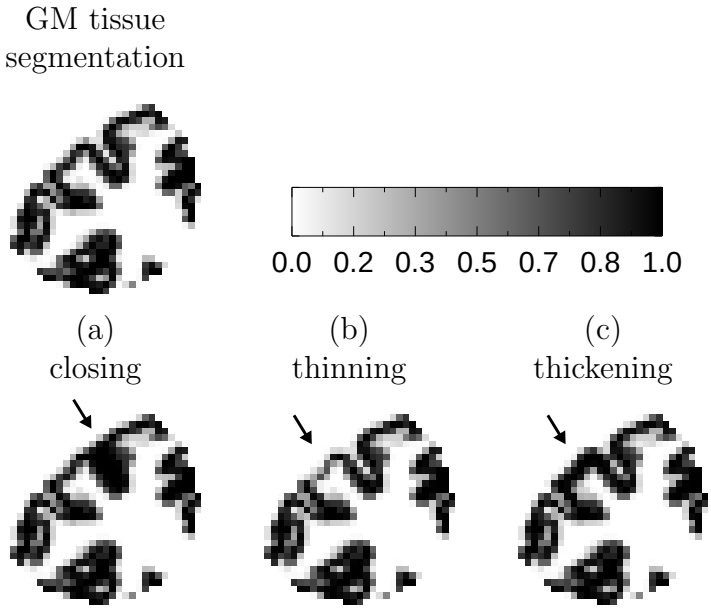


Figure 7. The GM fuzzy tissue segmentation data used for the construction of the emission data in experiment 3, is shown in the upper-left. The three types of simulated segmentation errors are: (a) the closing of a sulcus, (b) the thinning, and (c) the thickening of a gyrus. Only the central planes of the 3-D data are shown.

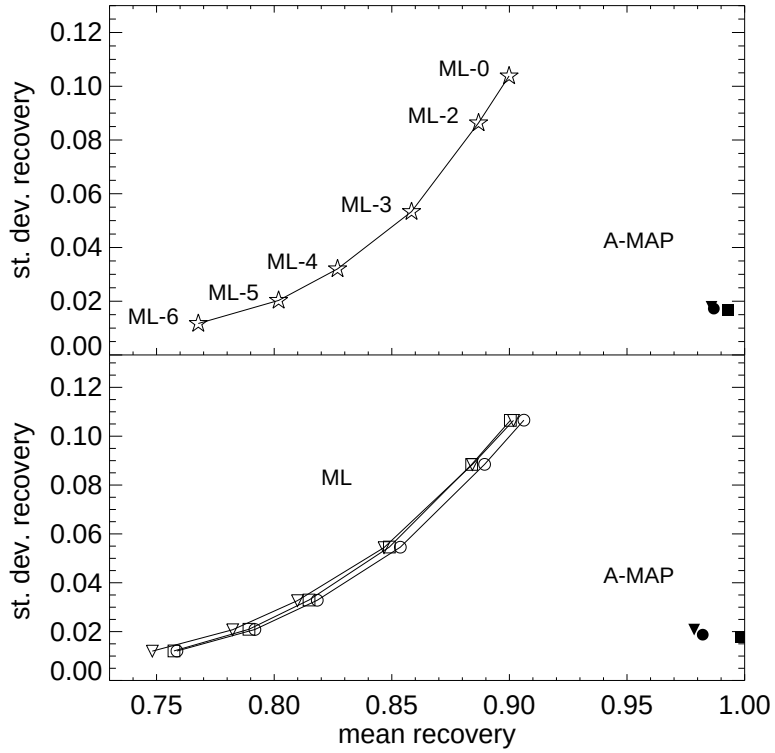


Figure 8. The plots show the results of the mean and the standard deviation of the recovery values of experiment 3 for post-smoothed ML and for A-MAP using modified anatomical information. The upper panel shows the results computed in an overall 3-D ROI. The lower panel shows the results computed within 3 different 3-D ROI around the sites of the corresponding induced segmentation errors. The filled symbols indicate the use of the modified anatomical information: (▼) for the "closing" of a sulcus, (■) for the "thinning" of a gyrus, and (●) for the "thickening" of a gyrus. Same conventions as in Fig. 3.

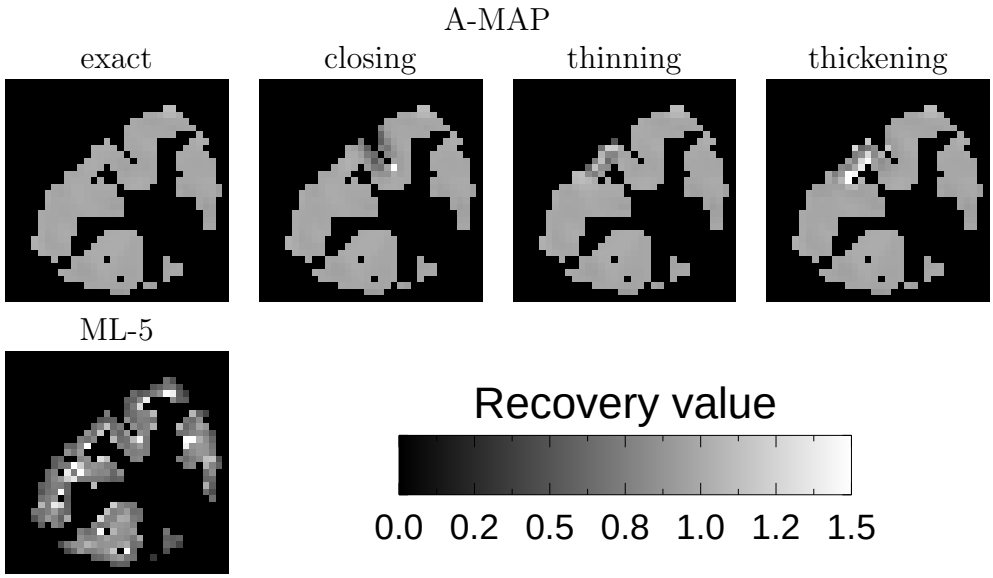
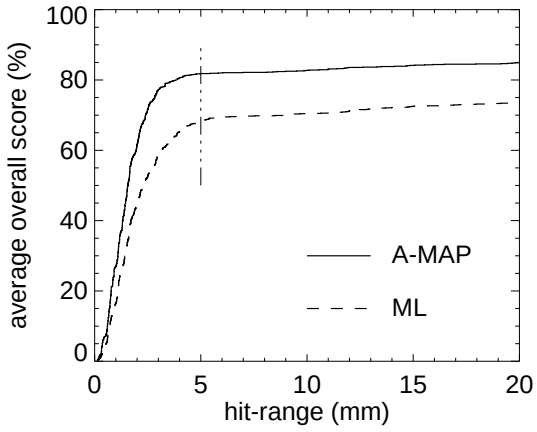
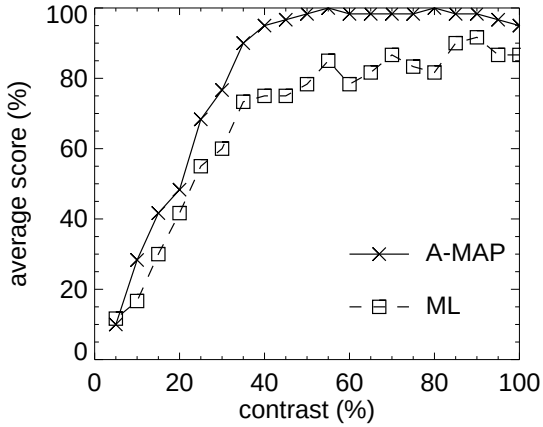


Figure 9. The images show the recovery values within GM for the central plane of the A-MAP reconstruction, using exact and modified anatomical information, and the post-smoothed ML reconstruction, using a Gaussian kernel with 5 mm FWHM, of a single noise realization of the projection data.

(a)



(b)



(c)

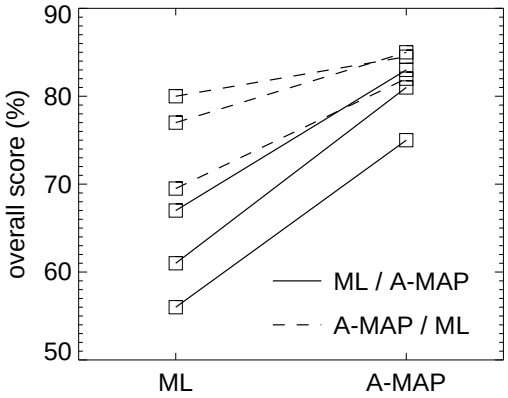
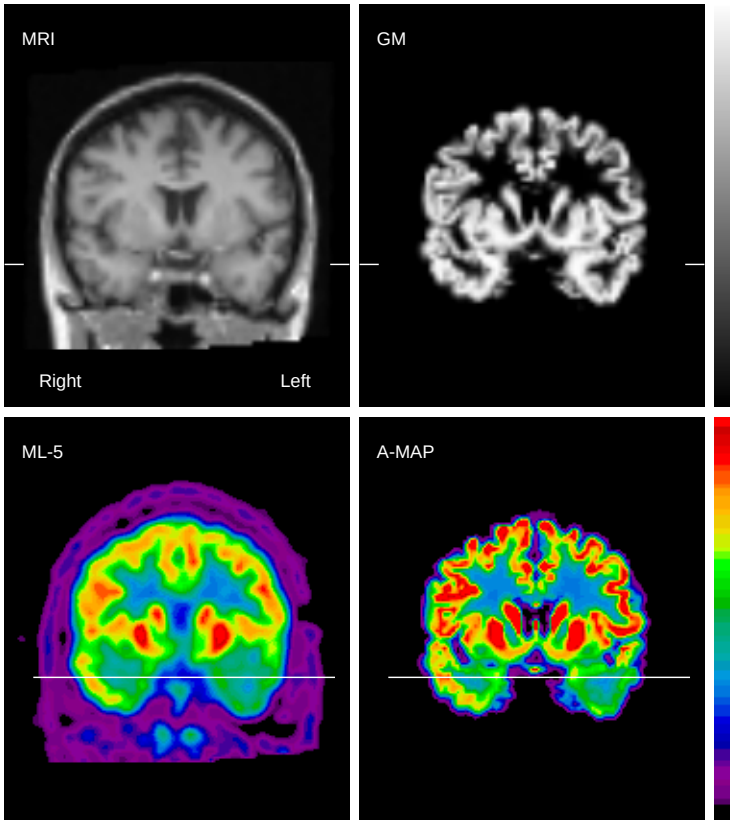


Figure 10. Results of the human observer study. Panel (a) shows the total score, i.e. over all contrast levels, for the mean observer as function of the distance between the true lesion location and the location indicated by the observer, for A-MAP (solid line) and ML (dashed line). Panel (b) shows the score for the mean observer as function of the contrast level of the hypometabolic regions, for A-MAP (solid line) and ML (dashed line). Panel (c) shows the total score, i.e. over all contrast levels, for each observer, and for ML and A-MAP. In (c), the solid lines are for the observers who started with the observation of the ML reconstructions, the dashed lines are for those who started with the A-MAP reconstructions.

A



B

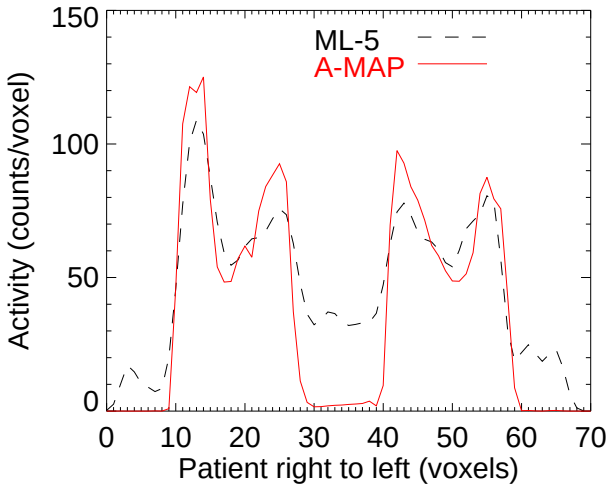


Figure 11. Panel A shows coronal views of the mean MR image, the derived GM segmentation, the ML reconstruction post-smoothed using an isotropic Gaussian kernel with 5 mm FWHM, and the A-MAP reconstruction. The A-MAP reconstruction is only shown for brain voxels, i.e. with a GM, WM, or CSF tissue fraction value exceeding ε. Panel B shows the activity distribution along a line profile through the temporal lobes for ML and A-MAP. The line profile is indicated in the coronal views.