Accepted Manuscript

Different partial volume correction methods lead to different conclusions: An ^{18}F-FDG PET Study of aging

Douglas N. Greve, David H. Salat, Spencer L. Bowen, David Izquierdo-Garcia, Aaron P. Schultz, Ciprian Catana, J. Alex Becker, Claus Svarer, Gitte Knudsen, Reisa A. Sperling, Keith A. Johnson

PII: S1053-8119(16)00151-8
DOI: doi: 10.1016/j.neuroimage.2016.02.042
Reference: YNIMG 12967

To appear in: *NeuroImage*

Received date: 13 November 2015
Accepted date: 15 February 2016

Please cite this article as: Greve, Douglas N., Salat, David H., Bowen, Spencer L., Izquierdo-Garcia, David, Schultz, Aaron P., Catana, Ciprian, Becker, J. Alex, Svarer, Claus, Knudsen, Gitte, Sperling, Reisa A., Johnson, Keith A., Different partial volume correction methods lead to different conclusions: An ^{18}F-FDG PET Study of aging, *NeuroImage* (2016), doi: 10.1016/j.neuroimage.2016.02.042

Different Partial Volume Correction Methods Lead to Different Conclusions: an ^{18}F-FDG PET Study of Aging

Douglas N. Greve[*,1,2]

David H. Salat[1,3]

Spencer L. Bowen[1]

David Izquierdo-Garcia[1]

Aaron P. Schultz[1]

Ciprian Catana[1]

J. Alex Becker[4]

Claus Svarer[5]

Gitte Knudsen[5,6]

Reisa A. Sperling[4,7]

Keith A. Johnson[1,4,7]

[1]Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Boston, MA, USA

[2]Harvard Medical School, Radiology Department, Boston, MA, USA

[3]Neuroimaging Research for Veterans Center, VA Boston Healthcare, USA

[4] Department of Radiology, Massachusetts General Hospital, Boston, MA, USA

[5] Neurobiology Research Unit and Center for Integrated Molecular Brain Imaging, Rigshospitalet, Copenhagen, Denmark

[6] Faculty of Health and Medical Sciences, Copenhagen, Denmark

[7] Department of Neurology, Brigham and Women's Hospital, Boston, MA, USA

*Corresponding author greve@nmr.mgh.harvard.edu, 617-312-9298, Fax 617-726-7422.

Abstract: A cross-sectional group study of the effects of aging on brain metabolism as measured with ^{18}F-FDG PET was performed using several different partial volume correction (PVC) methods: no correction (NoPVC), Meltzer (MZ), Müller-Gärtner (MG), and the symmetric geometric transfer matrix (SGTM) using 99 subjects aged 65-87 from the Harvard Aging Brain study. Sensitivity to parameter selection was tested for MZ and MG. The various methods and parameter settings resulted in an extremely wide range of conclusions as to the effects of age on metabolism, from almost no changes to virtually all of cortical regions showing a decrease with age. Simulations showed that NoPVC had significant bias that made the age effect on metabolism appear to be much larger and more significant than it is. MZ was found to be the same as NoPVC for liberal brain masks; for conservative brain masks, MZ showed few areas correlated with age. MG and SGTM were found to be similar; however, MG was sensitive to a thresholding parameter that can result in data loss. CSF uptake was surprisingly high at about 15% of that in gray matter. Exclusion of CSF from SGTM and MG models, which is almost universally done, caused a substantial loss in the power to detect age-related changes. This diversity of results reflects the literature on the metabolism of aging and suggests that extreme care should be taken when applying PVC or interpreting results that have been corrected for partial volume effects. Using the SGTM, significant age-related changes of about 7% per decade were found in frontal and cingulate cortices as well as primary visual and insular cortices.

Introduction: Positron emission tomography (PET) suffers from the partial volume effect (PVE) in which limited scanner resolution causes the activity to appear to spill out of one region and into another. This makes it difficult to quantify the effect in a given region because of loss of its own signal and contamination from nearby regions. The size of the PVE depends on many factors including the size and shape of the region and the size, shape of, and activity in nearby regions. This create a confound when studying aging or neurodegenerative diseases because it becomes unclear whether a difference between groups or across time is due to differences in tissue properties or is simply a side-effect of changes in size and shape due to atrophy. Atrophy-induced bias has been documented in simulations of an ^{18}F-FDG aging study (Meltzer et al., 1999); those simulations showed that the atrophic effects of aging caused a false enhancement of the decrease in measured ^{18}F-FDG uptake with age. It is therefore imperative that the PVEs be resolved before attempting to draw conclusions about neurodegenerative diseases.

Partial volume correction (PVC) methods have been developed to remove PVEs. The most popular are Meltzer (MZ, Meltzer et al., 1999), Müller-Gärtner (MG, Meltzer et al., 1996; Müller-Gärtner et al., 1992; Rousset et al., 1998b), and the geometric transfer matrix (GTM, Rousset et al., 1998a) and symmetric GTM (SGTM, Labbé et al., 1998; Sattarivand et al., 2012); see Erlandsson et al., 2012 for a general review of PVC methods. These methods require a second image of the brain from a modality that has substantially reduced PVEs compared to PET, such as MRI or CT. The GTM/SGTM is strictly for region-of-interest (ROI) analysis; MG and MZ provide voxel-wise results which can then be used in ROI analysis. All analyses performed in this paper are ROI-based using 3D PVC. The application of PVC to group ^{18}F-FDG studies of aging has produced conflicting results in terms of the biological conclusions about aging and metabolism. Some studies find regions with significant changes when not using PVC (NoPVC), few or none of which survive after applying PVC (Curiati et al., 2011; Ibanez et al., 2004; Kochunov et al., 2009; Yanase et al., 2005; Yoshii et al., 1988). Others report strong effects both with and without PVC (Knopman et al., 2014). Still others report strong aging results when using PVC without reporting the uncorrected results (Kalpouzos et al., 2009; Nugent et al., 2014a; Nugent et al., 2014b). Even the studies that do not correct at all are conflicting. For example de Leon et al., 1987; Hawkins et al., 1983; Kuhl et al., 1982 found no changes with age while Herholz et al., 2002; Loessner et al., 1995; Moeller et al., 1996; Petit-Taboue et al., 1998; Yoshizawa et al., 2014 found change with age. The studies that did not find

a change were performed when PET scanner resolutions were very poor which may account for the lack of results.

There are many differences in those studies that may account for the diversity of conclusions such as sample size, age ranges, scanner, PET reconstruction, scan duration, and whether kinetic modeling was performed or standardized uptake values used. However, the above studies also represent a diversity of PVC methods which may account for some of the variation. Several studies have compared PVC methods. Most have used simulations, phantoms, or just a few subjects (Boivin et al., 2014; Harri et al., 2007; Meltzer et al., 1999; Quarantelli et al., 2004; Rousset et al., 1998b) and do not address the question of whether different methods yield different biological conclusions. Uchida et al., 2011, studied aging and serotonin 2A receptor density using NoPVC, MZ, MG, and GTM and found that NoPVC and MZ showed age effects across the brain which went away entirely with GTM; MG was somewhere in between. Thomas et al., 2011 found that AD subjects differed from controls in ^{18}F-flumetamol uptake when MG was used but not with NoPVC or the GTM-derived region-based voxel-wise (RBV) method; their conclusions were limited to hippocampus.

Here we systematically study NoPVC, MZ, MG, and SGTM in ^{18}F-FDG data from an adult cross-sectional aging study. Each method was tested under a range of parameters attempting to not only document the differences in biological conclusions but to explain why these differences occur. While MG and SGTM/GTM can accommodate the modeling of extracerebral tissue such as cerebral spinal fluid (CSF), in practice they are almost never implemented in this way because CSF is not metabolically active and so is assumed to have no ^{18}F-FDG signal. Remarkably, this assumption does not appear to have ever been tested. So, a second aim of this study is to measure the signal in CSF and determine how its inclusion or exclusion in PVC affects the measured ^{18}F-FDG uptake. All PVC methods were implemented under the FreeSurfer neuroimage analysis software package (surfer.nmr.mgh.harvard.edu) and will be released with version 6 of this software.

Methods:

Subjects: Ninety-nine subjects were drawn from the Harvard Aging Brain (HAB) Study (Mormino et al., 2014). Study protocols were approved by the Partners Healthcare Institutional Review Board, and all participants provided written informed consent. Participants were

included if they had a score of less than 11 on the Geriatric Depression Scale, had a score of 0 on the Clinical Dementia Rating Scale, had a score of greater than 25 on the Mini-Mental State Examination, performed within education-adjusted norms on the Logical Memory delayed recall (>10 for ≥16 years of education, >6 for 8-15 years of education, and >4 for <8 years of education), and had no history of head trauma. Subjects aged from 66 to 87 years (mean 73.9y, s.d. 5.8y), 44 males and 55 females.

PET Acquisition: ^{18}F-FDG imaging was performed on a Siemens ECAT HR+ scanner (software version 7.2.2) without a neuroshield. Before injection, a 10 minute ^{68}Ge transmission scan was performed; this was segmented and used for attenuation correction. ^{18}F-FDG was administered intravenously with a 185-370 MBq bolus. After a 45-minute uptake period, ^{18}F-FDG-PET images were acquired for 30 minutes (six 5min frames) in 3-dimensional acquisition mode. The scanner room was quiet, and subjects had their eyes open and ears unplugged. The data were corrected for randoms, scatter, dead time, normalization, background, decay and attenuation, reconstructed with OSEM (3 iterations, 16 subsets, 2mm Gaussian filter) with voxel size 2.0594x2.0594x2.425mm, realigned, and summed. The point spread function (PSF) of the HR+ varies across space from about 4.5mm at the center to about 6.8mm at the edge of head (Jan et al., 2005); we modeled the PSF using a 6mm isotropic and space-invariant Gaussian kernel. Intensity normalization is described below.

MRI Acquisition: Magnetic resonance imaging was performed at the Massachusetts General Hospital Martinos Center for Biomedical Imaging on a 3-T imaging system (TIM Trio; Siemens) with a 12-channel head coil. Structural T1-weighted volumetric magnetization–prepared rapid-acquisition gradient echo images were collected (repetition time, echo time, and inversion time, respectively, 6400, 2.8, and 900 milliseconds; flip angle, 8°; and 1 × 1 × 1.2-mm resolution).

MRI Analysis: the MRIs were analyzed in FreeSurfer (FS) version 5.1 (surfer.nmr.mgh.harvard.edu) to provide detailed anatomical analysis for each subject (Dale et al., 1999; Fischl and Dale, 2000; Fischl et al., 1999; Segonne et al., 2007) including segmentation of subcortical structures including thalamus, hippocampus, amygdala, and ventricles (Fischl et al., 2002). The FS analysis also includes delineation of the surface separating cortical gray matter (GM) from white matter (WM) as well as segmentation of various

cortical areas (Fischl et al., 2004). The FS analysis in all subjects was manually inspected to assure quality of segmentation.

The FS segmentation was augmented to provide segmentations not available in the currently released FS analysis, namely extracerebral (i.e., sulcal) CSF, pons, skull, air cavities (e.g., sinuses), and other tissue in the head. To do this, a new atlas was created from an independent set of anatomical images from 79 subjects from the Brain Development IXI database (www.brain-development.org), ages 20 to 86 (mean 58, s.d. 17), with 34 males and 45 females. The pons belly was labeled by one of the authors (DNG) on the MNI305, an average of 305 brains distributed by the Montreal Neurological Institute (MNI), and mapped to the individual subject through a linear transform. The final pons label was created by intersecting the MNI305 label with the native FS brainstem label. Registration of the anatomical to the MNI305 is part of the FS analysis and is performed by the *mritotal* program distributed with the MINC software package (www.bic.mni.mcgill.ca). The tissue segmentation from the SPM software package (www.fil.ion.ucl.ac.uk/spm) was run on each subject as described in Izquierdo-Garcia et al., 2014 to create segmentations of skull, extracerebral CSF, air cavities, and remaining head tissue. These were all merged with the default FS segmentation to create a whole head segmentation with 101 regions-of-interest[1] (ROI), later reduced to 52 during PET analysis. The new segmentations from the 79 IXI subjects were used to create a new atlas in FS which was then applied to the 99 subjects of this study to provide individualized whole-head segmentation. An example segmentation is shown in Figure 1.

[1] A region-of-interest (ROI) is sometimes referred to as a volume-of-interest (VOI).

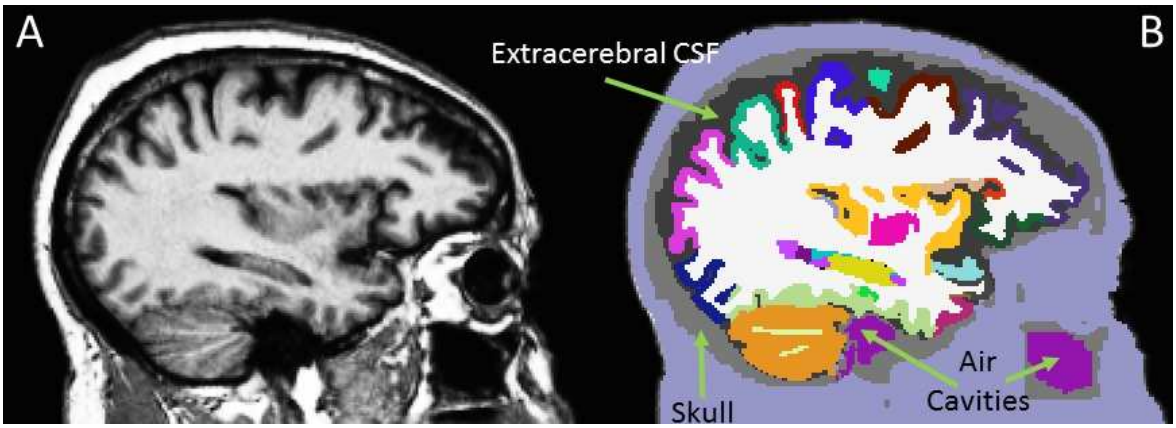

Figure 1. (A) Anatomical MRI (B) Whole head segmentation. Each color represents a different ROI including air cavities (e.g., sinuses), skull, and extracerebral CSF. Cerebral ROIs were created by the standard FreeSurfer analysis. The remainder of the head was modeled as a single ROI.

PET-MRI Integration: To perform PVC, one needs to map the segmentation information from the high-resolution MRI space into the PET volume. The PET was registered to the MRI using Boundary-based Registration (BBR; Greve and Fischl, 2009) using a 6 degree of freedom linear transform. The MRI segmentation was mapped into the PET space in a way that accounted for the tissue fraction effect (TFE; Erlandsson et al., 2012). The TFE is a type of PVE that occurs when multiple tissue types or ROIs occupy the same PET voxel. Its effect is generally much smaller than the PSF. Each PET voxel was divided into a uniform 3x3x3 grid. Each grid point was mapped into the (high-resolution) MRI space to determine the ROI for that grid point. This was done for all 27 grid points. The tissue fraction (TF) for a given ROI at a PET voxel was computed as the number of grid points for that ROI in that voxel divided by 27; the TF at a given PET voxel for a given tissue type (i.e., GM, WM, CSF) was computed by summing up the TFs for all ROIs of that tissue type in that voxel. To create a segmented volume in the PET space, each voxel was assigned to the ROI segment that had the highest TF at that voxel.

PET Analysis: the ^{18}F-FDG data were analyzed using 10 PVC methods (including without PVC) meant to reflect how data have been analyzed in the literature. To help keep track of the methods, the reader is directed to Table 1, which lists all the methods, parameters, and settings. Table 1 also includes the abbreviation with which we will refer to each method and parameter setting. In all methods, the ^{18}F-FDG data are reduced to 101 uptake values (one for each ROI). An initial analysis revealed no age-by-hemisphere interaction, so for lateralized ROIs, the values

from the left and right hemispheres were averaged together to yield a total of 52 ROI values. All methods used the same PET-MRI registration, segmentation, and PSF. All methods account for 3D partial volume effects. Though MG and MZ provided voxel-wise results, they were always reduced to ROI values for comparison with the SGTM. A relative standard uptake value (rSUV) was computed by dividing the ROI intensity by the intensity of pons for the given method; for MG, the SGTM value of pons was used. Pons was chosen because it is commonly used (Ewers et al., 2014; Kantarci et al., 2001; Knopman et al., 2014; Landau et al., 2011), easy to identify, has small PVEs, and is thought to be resistant to atrophy, at least in AD (Minoshima et al., 1995).

Without PVC (NoPVC): the measure for an ROI was simply computed as the mean uptake in the voxels in the ROI. For ventricular CSF, we also performed an analysis where voxels within 3mm of the edge of the ROI were removed so as to reduce spillover from adjacent GM or WM. This is for comparison against the ventricular CSF measure from the SGTM.

Meltzer (MZ): the MZ method (Meltzer et al., 1990) is relatively simple to implement. One obtains a binary mask of the brain (GM and WM structures), smooths it by the PSF, then divides the raw PET image by the smoothed mask on a voxel-by-voxel basis. Voxels outside of the smoothed mask are set to zero to avoid dividing by zero. Means for each ROI were then computed from this final image. MZ will have no effect for voxels well within the mask because the smoothed divisor will always be 1. Therefore, defining the edge of the mask becomes critical. We defined the brain mask by thresholding the unsmoothed GM+WM TF maps. To determine the sensitivity to brain mask definition, we chose thresholds of .10 (MZ-10), and .50 (MZ-50). In addition we dilated the 0.10 mask by one PET voxel (MZ-10-D1).

Symmetric Geometric Transfer Matrix (SGTM): For the SGTM analysis (Labbé et al., 1998; Rousset et al., 1998a), a linear model relating the ^{18}F-FDG intensities to the ROI means was established and solved using the following forward (1) and inverse equations (2):

$$y = X\beta \qquad (1)$$

$$\hat{\beta} = [X^T X]^{-1} X^T y \qquad (2)$$

where y is a vector of length equal to the number of voxels with the value of an element being the ^{18}F-FDG uptake value from that voxel, X is a matrix of size number of voxels by number of

ROIs (N_{ROI}) as described below, β is a vector of length N_{ROI} representing the true, but unknown, ROI means, and $\hat{\beta}$ is the estimate of the ROI means. Note that X is not the GTM itself. This linear model formulation of the GTM was originally developed in Labbé et al., 1998; Sattarivand et al., 2012 named this formulation the "symmetric GTM" (SGTM) and showed it has better noise and bias properties than the Rousset formulation. Accordingly, we refer to the method used here as the SGTM to emphasize the distinction with Rousset's GTM. Unlike MG and MZ, the SGTM does not have parameters.

The design matrix X was computed in the following way. For each ROI, an image was created in the PET space where the value in a voxel was the TF described above for that ROI in that voxel (this accounts for the TFE); this image will be very sparse. This image was then smoothed by the PSF of the scanner. The result (often called the regional spread function (RSF)) was then reshaped into a vector that then became the column in X for that ROI. This was repeated for each ROI. The X matrix depends on the PSF, the anatomy of the individual, and the PET voxel size; it is neither binary nor orthogonal. X may be badly conditioned or even non-invertible depending upon the PSF, and the number, size, and shape of the ROIs. If X is badly conditioned then it may amplify noise when inverted in Equation 2. We evaluated the SGTM using two sets of ROIs: (1) SGTM-Full: all ROIs, including extracerebral structures (i.e., structures that are neither GM nor WM), and (2) SGTM-NX: all ROIs excluding extracerebral ROIs (i.e., only considering GM or WM). The vast majority of SGTM implementations in the literature do not model extracerebral structures (i.e., SGTM-NX).

Müller-Gärtner (MG): The MG algorithm was originally formulated (Müller-Gärtner et al., 1992) to remove spill-in from the WM and CSF compartments as well as spill-out from GM compartment (the 3-compartment model). Meltzer et al., 1996 implemented a four compartment model. In the 3- and 4-compartment algorithms, GM, WM, and CSF are separately segmented from an MRI; the tissue masks are smoothed by the PSF. WM and CSF are multiplied by the expected value of the given tissue type. The resulting images represent the expected contribution of WM and CSF to each voxel in the PET image. These are then subtracted from the PET image to leave an estimate of the signal only from GM, which is then divided by the PSF-smoothed GM mask to give a final estimate of the GM signal without PVEs in each voxel. The smoothed GM mask has values between 0 (no GM) and 1 (full GM). A threshold must be chosen to prevent

noise amplification by dividing by a number close to 0. If the smoothed GM is not greater than this threshold at a voxel, then the voxel in the final image is set to 0; the ROI mean is computed excluding those voxels in the ROI. We analyzed data at mask-forming thresholds ranging from .01 to .8.

In the original formulation, Müller-Gärtner et al., 1992 extracted the mean WM signal from a hand-drawn ROI. Rousset et al., 1998b used the WM estimate from the GTM (the so-called "modified MG"). Our implementation extends this to use the SGTM estimate from all non-GM compartments, including CSF and other extracerebral structures. Most, perhaps all, implementations of MG simply ignore CSF. As with the SGTM, we analyzed the data with all structures (MG-Full) or excluding extracerebral structures (MG-NX). In both cases, the corresponding SGTM (Full or NX) was used to produce non-GM compartment means.

Group Analysis of Age Effects: at this point we have 10 different PET analysis methods (see Table 1 for a complete list) each with 52 ROIs for all 99 subjects. For each method and ROI, we regressed the ^{18}F-FDG value against subject age, including an offset term (see Figure 2). The age was relative to 75 years, so the offset term can be interpreted as the mean ^{18}F-FDG uptake of a hypothetical 75 year old. A two-tailed t-test was performed to compute the significance of the age slope. The same test was applied to the anatomical volume of all ROIs as well as the surface area and mean thickness of all cortical ROIs. Volume and surface area were normalized by intracranial volume prior to group analysis. We also performed an analysis where we tested the product of the SGTM uptake for a region times the volume for that region (SGTM*Volume). A Bonferroni correction was applied to correct for multiple comparisons across ROIs. With 52 ROIs, this required that the uncorrected p-value for an ROI be less than 0.001 for a corrected p-value of 0.05.

Volume Regression (NoPVC-VolReg): some researchers attempt to account for PVEs at the group analysis level by including a subject-specific anatomical measure as a nuisance regressor (Bauer et al., 2013; Walhovd et al., 2010). We have replicated this method by including the volume of the ROI as computed by FreeSurfer, i.e., we added another column to the group age analysis design matrix where the value at a given row is the ROI volume corresponding to the subject indicated by the row. A similar analysis was performed for surface area and thickness.

Partial Volume Bias Adjusted (PVBA): We performed another analysis in which we assess the effects of anatomical bias on the NoPVC results by estimating and subtracting out the atrophy-induced bias. The mean uptake in each ROI for a hypothetical 75yo was computed from the group SGTM analysis. A subject-specific synthetic ^{18}F-FDG image was created in each subject's anatomical/segmentation space by filling each voxel in each ROI for that subject with the uptake value of that ROI from the hypothetical 75yo. This anatomical-space ^{18}F-FDG synthetic image was resampled into the PET space (including TFEs) and smooth by an isotropic, space-invariant 6mm Gaussian PSF. There is no effect of age on the simulated uptake values since the underlying values are fixed for all subjects at that of a 75yo, so any change with age detected in the simulations indicates a systematic change with age (i.e., a bias) and so reflects subject differences in anatomy and not true changes in uptake. The simulated NoPVC group slope was subtracted from the actual NoPVC group slope and the p-values recomputed. This measures the ability of the NoPVC to detect an actual age change in metabolism free from the bias that arises when not performing PVC. An example of this analysis is worked out in the results section. PVBA is not a true correction method because it does not attempt to recover the signal.

Results. The Results section is divided into methodological results and neurobiological results. We first describe the tabulated results then investigate each method more fully. The reader is directed to Table 1 to help keep track of the methods, parameter settings, and overall results of each. We will use Figure 2 to demonstrate how the performance metrics were computed. Figure 2 shows a plot of rSUV versus age for SGTM-Full, NoPVC, and simulated NoPVC for the isthmus cingulate (IC) along with best-fit lines; each data point represents a subject. The "offset" mentioned in the tables refers to the value of the best-fit line at age 75y, and the "slope" is the slope of the line as a percentage of the offset per decade. The p-value is computed from a t-test testing the null hypothesis that the slope=0. Table 1 shows the number of ROIs whose $p<.001$ for each method (N ROIs).

Table 2 shows a breakdown of which GM ROIs were significant for each method as well as some anatomical results. It also has a column showing the significant results for the NoPVC

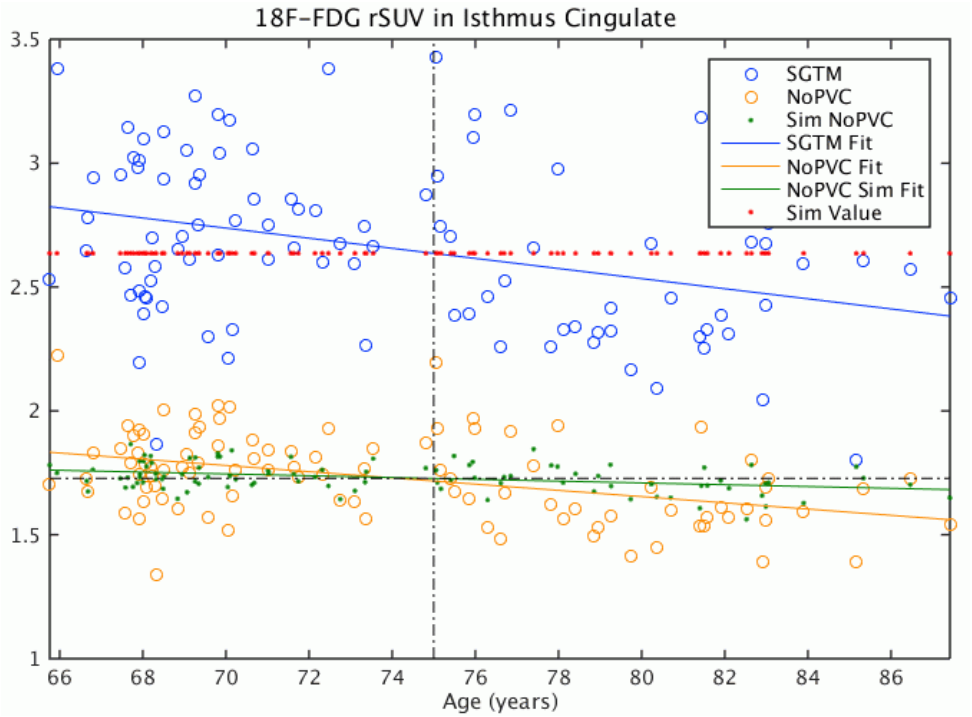

Figure 2. ^{18}F-FDG uptake vs age in the isthmus cingulate (IC) cortex for SGTM and NoPVC along with best-fit lines from the group age analysis. The NoPVC under simulation conditions (NoPVC Sim) is also shown. Sim Value is the value of the simulated uptake for IC, fixed at 2.636 for all subjects. The black dashed line is a horizontal (slope=0) for reference. The SGTM slope is -0.204±0.056/decade (p=.00042). The NoPVC slope is -0.126±0.027/decade (p=.00001). The NoPVCSim slope is -0.036±0.009/decade (p=0.00017). After adjustment (PVBA), the NoPVC slope declines to -0.090± 0.027/decade (p=.00137). ± indicates standard deviation of the mean.

simulation. Supplementary Table S1 details the average ^{18}F-FDG uptake for a 75yo in each ROI, the slope of the ^{18}F-FDG-age association, and the uncorrected p-value for each ROI for NoPVC and SGTM-Full. Supplementary Table S2 shows the detailed anatomical results (volume, thickness, cortical surface area) for each ROI.

NoPVC Method. In NoPVC over half (27/41) of the GM ROIs had significant age correlations; these were mostly cortical with the exception of thalamus and ventricular CSF. The partial volume effect manifest itself as substantial differences between NoPVC and PVC. The offset (mean uptake) of SGTM-Full was about 1.652 times that of NoPVC (Table S1). Focusing on IC

(Figure 2), the offset was 1.717 for NoPVC and 2.636 for SGTM-Full, a factor of 1.54. The raw slope was also much more for SGTM (-0.204/decade, blue line) than NoPVC (-0.126/decade, yellow line), a factor of 1.62 (1.579 for all ROIs). While the SGTM slope is 1.62 times that of NoPVC, the standard error in slope increases by a factor of 2.1 relative to NoPVC, making the SGTM less significant (p=0.00042) than NoPVC (p=0.00001). However, the NoPVC simulations show that some of the NoPVC slope is due to anatomical changes with age and not metabolism. For example, the simulated NoPVC in Figure 2 (green line/dots) has a slope of -0.036/decade. By construction, there were no metabolic age changes in the simulation, implying that about a third (-0.036/-0.126) of the measured NoPVC slope is related to anatomy, not metabolism (i.e., a slope bias). The slope bias was highly significant (p=0.00017) in IC. When the IC slope bias was subtracted from the slope of the real data (i.e., slope of green line subtracted from the slope of the yellow line in Figure 2) and the p-value recomputed (i.e., the PVBA method), the adjusted slope (-0.090/decade) was not significant by our standard (p=.00137).

Across cortical ROIs, the slope bias averaged 26% of the NoPVC slope, and overall there were 18 ROIs where the slope bias was significant. In hippocampus, the slope bias was 79%. After adjustment, 17 NoPVC ROIs lost their significance, and the remainder matched the SGTM and MG methods quite well (Table 2). Computing the slope as a percent of offset reduces the impact of PVEs because both slope and offset are affected. For example, the IC percent slopes are not so different: -7.3%/decade (NoPVC) and -7.7%/decade (SGTM). However, this correction is not perfect because PVEs affect offset and slope differently as can be seen for most of the other ROIs (Table S1). Also, the percentages can only be computed after statistical analysis so the percentage calculation cannot be used to remove the bias from the p-values. The ROIs showing significant bias tended to be around the rolandic operculum, medial temporal lobe, and visual cortex. These areas do not match well with the regions that had significant anatomical changes, and inclusion of ROI volume as a nuisance variable (NoPVC-VolReg) had virtually no effect on the number or identity of the significant ROIs (Tables 1 and 2). Surface area and thickness were no more effective than volume (results not shown).

SGTM Method. SGTM-Full had 12 significant ROIs mainly in frontal and cingulate areas. When extracerebral structures were removed from the SGTM model (SGTM-NX), the number of

significant ROIs dropped from 12 to a subset of 5 (Tables 1 and 2). We also tested the SGTM excluding only CSF from the model, and the results were nearly identical to SGTM-NX.

Müller-Gärtner Method. The MG method was evaluated at various thresholds. We chose to report the results at a threshold of 0.3 for several reasons. At 0.3 and less, all ROIs for all subjects had at least some voxels that survived the thresholding. Beyond 0.3, some subjects/ROIs would have to be excluded due to data loss. At 0.3, ROIs kept about 95% of their voxels; at 0.5, ROIs maintained only 40% of their voxels. Reflecting the anatomical results, the total number of 0.3 threshold voxels significantly decreased with age. The number of surviving voxels in each GM ROI always decreased. However, none of the ROIs individually were significant; this was not the case at higher thresholds. The MG results at lower thresholds did not change much. This justification for 0.3 cannot be generalized to other studies with different age groups and PSFs.

With full head model (MG-Full), there were 12 significant ROIs in the real data, and they overlapped almost perfectly with the ROIs in SGTM-Full. When the extracerebral compartments were not included, the number of significant ROIs in the real data dropped from 12 down to 5; these were identical with the ROIs in SGTM-NX.

Meltzer Method. The MZ results were highly dependent on how the brain mask was defined. The most generous brain mask (MZ-10-D1) had 26 significant ROIs in the real data. This dropped to 12 without dilation (MZ-10) and to only 2 with a threshold of 0.5 (MZ-50). MZ-10-D1 was nearly identical to the NoPVC. The ROIs in MZ-10 were nearly identical to those in SGTM-Full.

Metabolic Changes with Age. The neurobiological results are only reported and discussed for the SGTM-Full (NoPVC shown in Table S1). The SGTM-Full had 12 significant ROIs, all in cortex and largely in frontal areas but also in insula, primary visual cortex and lingual gyrus (a higher visual area). The magnitude of the decline had a range of -9.9 to -6.2%/decade. None of the subcortical regions were close to having a significant change with age. In terms of simple uptake, the primary visual and auditory cortices had the largest values, perhaps due to functional activation from inevitable visual and auditory stimulation. Cortical areas had an rSUV of around 2.0-3.0 whereas subcortical GM was lower at around 1.5-2.0. The uptake in cerebral WM was 0.8, about one third that in cortical GM. Interestingly, CSF had an uptake of about 0.35 ($p<10^{-10}$), 15% that of GM, and was nearly identical in ventricular and extracerebral ROIs. When the NoPVC average for the ventricular CSF was recomputed after eroding the voxel by 3mm, the

result of 0.42 was nearly identical to that of the SGTM. The measured uptake in air cavities was 0.33; the signal was not nearly this high in air outside of the head. The skull had the least rSUV at .08. The remainder of the head had an uptake of 0.23. When considering the SGTM uptake times ROI volume, many ROIs (29) became significant; these were mostly cortical ROIs but also included hippocampus and thalamus.

Global Signal Changes with Age. The total GM SGTM signal significantly decreased with age (-4.5%/decade, p<.003) as did the whole-brain signal (-4.8%/decade, p<.0005). We evaluated which ROIs were responsible for the reduction in total GM signal with age as shown in Table 3. Cerebellum cortex topped the list but this was partially due to the fact that it is a very large ROI, so even small changes affect the total GM signal. We also compared the global (i.e., whole head) signal in NoPVC to SGTM and found them to be the same. This indicates that our software is working properly since the SGTM only redistributes activation and so should have the same global signal as NoPVC. The global signals for MG and MZ are not the same as NoPVC; this is expected because MG and MZ are rescaling methods whereas the SGTM is a redistribution method.

Pons Reference Region: The pons volume did not change significantly with age (-1.3%/decade, p=0.5). The pons SGTM uptake was only 1.06 times that of NoPVC, and this ratio did not change significantly with age (-0.3%/decade, p=0.4). In the NoPVC simulations, synthetic pons uptake did not change significantly with age (0.2%/decade, p=.17). This gives us some confidence that the reference region uptake is not changing with age and is not highly influenced by PVEs.

Anatomical Changes with Age. Anatomically (Table S2), virtually all GM ROIs decreased in volume, surface area, and thickness with age, but there were only 8 ROIs with significant volume decreases: parahippocampal gyrus, superior parietal gyrus, post central gyrus, hippocampus, amygdala, accumbens, thalamus, ventral diencephalon, and cerebral WM. Significant surface area decreases were found in only 3 cortical ROIs: inferior temporal, fusiform, and superior parietal. No ROIs showed significant thickness changes. Significant increases in volume were found in extracerebral (i.e., sulcal) CSF. Ventricular CSF also increased in size but just missed the cutoff for multiple comparisons correction.

Discussion:

The results presented here show that the way PVC is performed (i.e., the PVC method and parameters used) is critical to the conclusions that one draws about the effects of aging on metabolism. Depending upon the method, parameters, and anatomical modeling, one could report that nearly all brain areas have a significant drop in metabolism between the ages of 66-87 or one could report that virtually no areas are affected or one could report that only frontal regions are affected. At some parameter settings, the methods give similar results. MG and MZ show strong sensitivity to parameter settings. This variety of conclusions from this single data set reflects the diversity of results reported in the literature as described in the Introduction. Below we attempt to explain why there is such a diversity by examining the assumptions of each method in detail.

Not performing PVC (NoPVC) yielded the most number of significant ROIs, and so it is tempting to conclude that it is the most powerful method. However, simulations showed that NoPVC has a bias that exaggerates the age effect; 18 ROIs tested positive for an age effect on metabolism where no age effect was present. This happened in ROIs where there were no significant changes in volume, surface area, or thickness. The reason for this has to do with a complex interaction between atrophy and PVEs. Figure 3, left, shows brain images in a coronal slice through primary visual cortex for a 20yo (top) and 80yo (bottom). A cartoon of the uptake across space is shown in the right panels. There are two ROIs (A and B) separated by CSF space; the gray dashed lines indicate the total uptake. The size of the ROIs is set to be fixed across age. The brain undergoes both large and small structural changes with age. For example, one can see that the sulcal CSF spaces have expanded due to loss of GM and WM. However, the locus of tissue loss does not need to be proximal to the locus of CSF expansion. For example, the primary visual area (V1, calcarine sulcus) may retain most of its volume, thickness, and surface area, yet the sulcus itself may expand greatly due to atrophy in other parts of the brain. When the banks of a sulcus are close to each other (as in a younger subject), the PVE causes signal from both banks to reinforce each other. When the banks are pushed apart (as with an older subject), the adjacent high-activity GM is replaced by low-activity CSF. Thus, the two banks reinforce each other less, which causes a loss of measured signal when PVEs are not corrected (the peak of the gray dashed line in Figure 3 drops from 0.80 to 0.75). This is how NoPVC can detect false decreases in metabolism in ROIs where the GM volume, surface area, or thickness do not change with age. This also explains why using an anatomical measure as a nuisance regressor (NoPVC-VolReg) is

ineffective for PVC. Another NoPVC regression method (untested here) is to use a subject-specific voxel-wise nuisance regressor generated from a smoothed GM mask (Oakes et al., 2007). Even in the case where including a volume covariate is effective, it will likely be collinear with regressors of interest causing reduced statistical efficiency and possibly bias. When the slope from the simulated data was subtracted (PVBA), the number of ROIs dropped from 27 down to 10 and became much more consistent with PVC results. Thus, it appears that PVC is not inherently less powerful than NoPVC. An important caveat is that this conclusion depends on the size or the ROIs studied here -- there will be some ROI size below which PVC noise amplification will become important.

The MZ method showed extreme sensitivity to variation in the brain mask definition. If the brain mask was conservative, the age effect largely went away (only 2 ROIs). If the brain mask was liberal, then there were large detected age effects that were nearly identical to NoPVC (26 ROIs out of 41). The explanation for this is relatively simple: MZ only corrects near the edge of the brain mask; away from this edge, MZ is identical to NoPVC; as the mask edge is dilated, more ROIs become like their NoPVC counterparts. The amount of dilation needed is very small, in our case on the order of one PET voxel. This is sufficient to explain why some researches (Knopman et al., 2014; Uchida et al., 2011) find virtually no differences between MZ and NoPVC while others (Ibanez et al., 2004) see their NoPVC results go away after MZ correction. MZ makes the assumption that GM and WM activity are equal. This assumption is obviously inaccurate (the rSUV of GM is about 3 times that of WM). MZ cannot model extracerebral structures and so cannot compensate for CSF spillover described below. Others (Harri et al., 2007; Meltzer et al., 1999; Quarantelli et al., 2004) have already shown in simulations that the MZ is inferior to SGTM and MG in the cases where the registration, segmentation, and assumed PSF are inaccurate. The main advantage of MZ is that it is very simple to implement with off-the-shelf tools. MZ also provides voxel-wise results. However, even when mean results comparable to the SGTM are obtained, there is no guarantee that the voxel-wise results will be accurate given that voxels near the edge of the brain will be overcorrected and those near the GM/WM border will be undercorrected.

Under the SGTM assumptions (see below), MG will give exact results at the level of the GM compartment (i.e., when all GM has the same uptake). The near-exact nature of MG is reflected by the significant MG-Full ROIs being nearly identical to those in SGTM-Full. However, MG

cannot account for the spillover between ROIs within the GM compartment. MG also requires choosing a threshold to exclude low-GM voxels. This is problematic because this threshold must be chosen carefully. In our data, a threshold of 0.3 seemed optimal; higher thresholds caused a loss of data. MG has the advantage of allowing for voxel-wise analysis.

In the SGTM and MG, the modeling of extracerebral structures was a major factor influencing in the number of significant ROIs. When these structures were removed from the SGTM (SGTM-NX) and MG (MG-NX), the number of significant ROIs dropped from 12 down to 5. The results were nearly identical when only CSF was excluded from the model indicating that it is the modeling of CSF and not skull or head that is important.

CSF was found to have a surprisingly high signal relative to GM and WM even when only

considering a ventricular ROI whose voxels were at least 3mm away from GM and WM. The

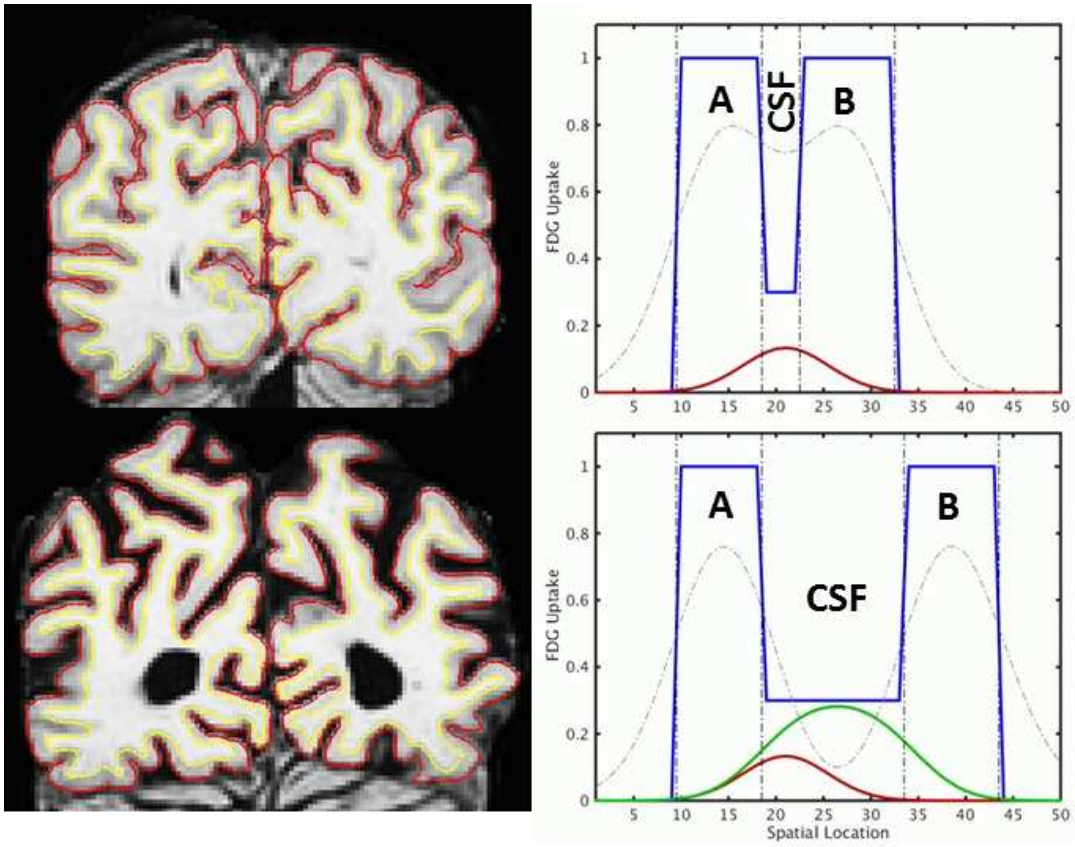

Figure 3. Left: Coronal slice through visual cortex of a 20yo and 80yo. The yellow line is the boundary between white matter and cortical gray matter; the red line is the cortical gray matter and extracerebral CSF. Both the ventricular and extracerebral CSF spaces are expanded in the 80yo. Right: Simulations showing how an expansion of the CSF space causes an increase in spillover from CSF into regions A and B (i.e., green line > redline). The gray dashed lines show the spillover of A/B into the CSF spaces; this spill-over will be removed from CSF if the GTM is operating properly.

vast majority of papers that use PVC ignore CSF because CSF has no metabolism and so is assumed to have no signal; however, this assumption does not seem to have ever been tested. While CSF is not metabolically active, there are biological reasons to expect ^{18}F-FDG to be present in CSF. CSF is a fluid that is nearly indistinguishable from brain insterstitial fluid (Redzic et al., 2005). It is mainly created from blood by the choroid plexus (CP) in the ventricles but can also be created in the sulcal space by direct communication with pial blood vessels (Redzic et al., 2005). The glucose content in human CSF is roughly 60-70% of that in blood (Roos, 2004). Deane and Segal, 1985 found that various forms of sugar crossed the blood-CSF barrier in the CP of sheep. Rmeily-Haddad et al., 2011 used PET ^{18}F-FDG measures in CSF to make inferences about the function of the choroid plexus; they report the ^{18}F-FDG signal in the

CSF monotonically increased up to the end of their imaging time of 40min. Note that Rmeily-Haddad et al., 2011 used PET to measure the CSF signal and did not sample the CSF directly. We are not aware of any other studies on ^{18}F-FDG imaging in CSF.

Despite biological plausibility, it is possible that our measurement of CSF signal is an artifact possibly caused by the non-negativity constraint inherent in OSEM reconstruction (Reilhac et al., 2008), incorrect segmentation, inaccurate scatter correction, and/or incomplete PVE removal from neighboring high concentration regions. Evidence of such an artifact was found in air cavities. The detected air signal may have air-specific explanations such as the difficulty in segmenting air, bone, and soft tissue solely from a T1-weighted MRI or inaccurate attenuation correction in low-attenuation air. We cannot resolve whether CSF does or does not have signal from this data. A more definitive test would be to draw and test CSF immediately after imaging. In any event, including or excluding the CSF ROI from the PVC model has a strong effect.

If there is signal in CSF, then the effect of not modeling it can be explained with the help of Figure 3 where, as mentioned above, one can see that the sulcal CSF spaces have expanded with age but the size of the ROIs does not change. The true uptake (blue line) in the CSF is set to be non-zero. The red line shows the spillover from CSF into A and B in the younger subject. For the older subject, all the uptake values are the same, but the space between A and B has enlarged to allow more CSF. The spillover from CSF is shown in green. One can see that the amount of spillover into ROI A in the older subject is more than that of the younger subject (i.e., green>red). Though the difference in spillover is small, it is systematic. When this spillover is not accounted for, PVC will cause an apparent increase in ^{18}F-FDG uptake in the ROIs. Of course in real data, there will be decreases in ^{18}F-FDG with age. However, when CSF is not accounted for in PVC, these decreases will be partially negated by the uncompensated increases in CSF spillover and make the PVC decreases appear to be weaker than they truly are. If this is true, then the vast majority of PVC ^{18}F-FDG aging results in the literature are probably weaker than they should be because modeling of CSF is so rare. This effect would only be present when the CSF signal is non-zero, which may not be the case with other radiotracers.

The assumptions of the SGTM/GTM are (1) the PVE is accurately represented by linear PSF smoothing[2], (2) the PSF is known, (3) the segmentation is accurate, (4) the segmentation is complete (i.e., all structures with uptake are represented), (5) the MRI-PET registration is accurate, and (6) the uptake is constant over the ROI. When these assumptions are met, the SGTM will give the exact mean in the ROI (under noiseless conditions). MZ and MG make these assumptions as well as others described above. From this study alone, we cannot unambiguously state that SGTM is better than the other methods. However, Quarantelli et al., 2004 performed realistic simulations and found the GTM to be better than MZ and MG even after simulating errors in segmentation, registration, and PSF. Harri et al., 2007 showed GTM to have better recovery coefficient (RC) performance than MG and MZ on physical phantoms. Meltzer et al., 1999 and Labbé et al., 1998 showed that the MG RC was superior to that of MZ on physical phantoms. Meltzer et al., 1999 also showed that MG RC was superior to that of MZ on simulations even when including segmentation, registration, and PSF errors. However, Meltzer et al., 1999 found that MG was more sensitive to these errors in basal ganglia and temporal cortex (but not frontal or occipital cortices), defining sensitivity as the change in RC per error in registration, segmentation, or PSF. This sensitivity may be important for group studies and must be balanced against raw RC performance, especially in multisite studies where the PSF differs across site. Based on this body of evidence, we think that SGTM with full head model is the preferred method for ROI analysis. MG-Full also works well and can be used for voxel-wise analysis (see Greve et al., 2014 for caveats), however care needs to be taken when masking. MZ should be used and interpreted with extreme care due to its sensitivity to how the brain mask is constructed. When PVC is performed, researchers should report both PVC and NoPVC results for completeness.

As an investigation of the neurobiology of aging, the SGTM results showed that metabolism drops between ages 66 and 87 by about 7% per decade in the frontal and cingulate cortices. Hypometabolism in the frontal lobe is one of the most consistent results in the literature (Curiati et al., 2011; Ibanez et al., 1998; Ibanez et al., 2004; Kalpouzos et al., 2009; Knopman et al.,

[2] This assumption can be relaxed with perturbation GTM (Bowen, S.L., Byars, L.G., Michel, C.J., Chonde, D.B., Catana, C., 2013. Influence of the partial volume correction method on (18)F-fluorodeoxyglucose brain kinetic modelling from dynamic PET images reconstructed with resolution model based OSEM. Phys Med Biol 58, 7081-7106, Du, Y., Tsui, B.M., Frey, E.C., 2005. Partial volume effect compensation for quantitative brain SPECT imaging. IEEE Trans Med Imaging 24, 969-976.).

2014; Kochunov et al., 2009; Nugent et al., 2014a; Nugent et al., 2014b; Yanase et al., 2005; Yoshii et al., 1988), though it did not always survive PVC. Frontal regions also accounted for much of the reduction in global signal with age (Table 3). We also found strong effects in visual and insular cortices. On the whole, the entire cortex showed hypometabolism, and it may be that changes in other cortical areas would become significant with more subjects. Knopman et al., 2014 found significant changes across all of cortex; however, they had a much larger sample size (806) with a much larger age range (30-95) and used an implementation of MZ that was no different than NoPVC. Subcortically, there were virtually no areas that were close to showing any significant metabolism changes in our results.

The SGTM*Volume represents the total metabolism for an ROI. This quantity can be interpreted as a measure total "functionality" for the ROI. For example, if the volume of the hippocampus drops by 10% but its uptake per unit tissue remains the same, then one can infer that hippocampus has lost 10% of its functional capacity. The SGTM*Volume measure showed changes across cortex as well as thalamus and hippocampus. This is somewhat expected given that SGTM uptake and volume are both decreasing, so their product will decrease even more with age. Note that this does not necessarily mean that metabolism per unit tissue is changing.

Conclusions: we have analyzed the effects of aging on ^{18}F-FDG uptake in a cross-sectional set of subjects aged 66-87y without PVC and with three popular PVC methods (GTM/SGTM, MG, and MZ) to examine how the results changed with method and how sensitive the conclusions were to changes in method parameters. The results without PVC were highly contaminated by anatomical changes causing ^{18}F-FDG in the entire brain to appear to drop with age; the results became consistent with PVC when this bias was removed. MZ was found to be highly sensitive to changes in the definition of the brain mask. A generous mask caused MZ to be identical to NoPVC; a conservative mask caused almost all significant ROIs to go away. MG could perform well but was sensitive to a PSF-dependent threshold and cannot account for spillover between ROIs. Based on these results, we suggest that researchers use the SGTM or GTM methods when doing an ROI analysis and MG or possibly the GTM-based RBV when performing voxel-wise analysis. GTM software is available through PVEOut (pveout.ibb.cnr.it) and PMOD (www.pmod.com); all methods in this manuscript will also be available in FreeSurfer 6.0. Researchers should carefully weigh the tradeoffs in MZ before using that method. SGTM was our PVC method of choice because it makes the fewest assumptions and has been found by

others to be robust. We found a surprisingly high ^{18}F-FDG signal in CSF; excluding CSF from the PVC model caused a precipitous drop in the power of PVC to detect age effects due to the expansion of CSF spaces with age. The diversity of results with different PVC methods in our study mirrors the confusion found in the literature about aging and metabolism. Our results suggest that the exact method and exact parameters are crucial to the conclusions one draws from the data. Therefore, researchers should be diligent in reporting their methodological details and readers should not assume that different PVC methods are comparable. These results show that one should be extremely careful in comparing results across studies that use different PVC methods. The anatomical changes in our cohort were very mild; one can expect that the problems and sensitivities with PVC found in this study to be exacerbated in cases of more extreme atrophy. Based on these data, we conclude that brain metabolism does decrease between the ages of 66 and 87 mainly in frontal and cingulate regions. We found no evidence of decline in subcortical regions, including hippocampus.

Acknowledgements: Support for this research was provided in part by the National Institutes of Health grants 5R01EB006758-04, 5R01NS052585-05, 5R21NS072652-02, P41-RR14075, R01RR16594-01A1, 1R21EB018964-01, 1S10RR023043, and 1S10RR023401. This work was also supported by EU 7th Framework Program: INMiND (HEALTH-F2-2011-278850).

References:

Bauer, C.M., Cabral, H.J., Greve, D.N., Killiany, R.J., 2013. Differentiating between Normal Aging, Mild Cognitive Impairment, and Alzheimer's disease with FDG-PET: Effects of Normalization Region and Partial Volume Correction Method. Journal of Alzheimer's Disease & Parkinsonism 3.

Boivin, G., Vassilis Genoud, Zaidi, H., 2014. MRI-guided partial volume correction in brain PET imaging: Comparison of five algorithms. Biomedical Technologies 1, 73-81.

Bowen, S.L., Byars, L.G., Michel, C.J., Chonde, D.B., Catana, C., 2013. Influence of the partial volume correction method on (18)F-fluorodeoxyglucose brain kinetic modelling from dynamic PET images reconstructed with resolution model based OSEM. Phys Med Biol 58, 7081-7106.

Curiati, P.K., Tamashiro-Duran, J.H., Duran, F.L., Buchpiguel, C.A., Squarzoni, P., Romano, D.C., Vallada, H., Menezes, P.R., Scazufca, M., Busatto, G.F., Alves, T.C., 2011. Age-related metabolic profiles in cognitively healthy elders: results from a voxel-based [18F]fluorodeoxyglucose-positron-emission tomography study with partial volume effects correction. AJNR Am J Neuroradiol 32, 560-565.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical Surface-Based Analysis I: Segmentation and Surface Reconstruction. Neuroimage 9, 179-194.

de Leon, M.J., George, A.E., Tomanelli, J., Christman, D., Kluger, A., Miller, J., Ferris, S.H., Fowler, J., Brodie, J.D., van Gelder, P., et al., 1987. Positron emission tomography studies of normal aging: a replication of PET III and 18-FDG using PET VI and 11-CDG. Neurobiol Aging 8, 319-323.
Deane, R., Segal, M.B., 1985. The transport of sugars across the perfused choroid plexus of the sheep. J Physiol 362, 245-260.
Du, Y., Tsui, B.M., Frey, E.C., 2005. Partial volume effect compensation for quantitative brain SPECT imaging. IEEE Trans Med Imaging 24, 969-976.
Erlandsson, K., Buvat, I., Pretorius, P.H., Thomas, B.A., Hutton, B.F., 2012. A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. Phys Med Biol 57, R119-159.
Ewers, M., Brendel, M., Rizk-Jackson, A., Rominger, A., Bartenstein, P., Schuff, N., Weiner, M.W., Alzheimer's Disease Neuroimaging, I., 2014. Reduced FDG-PET brain metabolism and executive function predict clinical progression in elderly healthy subjects. Neuroimage Clin 4, 45-52.
Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proceedings of the National Academy of Sciences 97, 11044-11049.
Fischl, B., Salat, D.H., Albert, M., Dieterich, M., Haselgrove, C., Kouwe, A.v.d., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341-355.
Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. Neuroimage 9, 195-207.
Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cereb Cortex 14, 11-22.
Greve, D., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48, 63-72.
Greve, D.N., Svarer, C., Fisher, P.M., Feng, L., Hansen, A.E., Baare, W., Rosen, B., Fischl, B., Knudsen, G.M., 2014. Cortical surface-based analysis reduces bias and variance in kinetic modeling of brain PET data. Neuroimage 92, 225-236.
Harri, M., Mika, T., Jussi, H., Nevalainen, O.S., Jarmo, H., 2007. Evaluation of partial volume effect correction methods for brain positron emission tomography: Quantification and reproducibility. J Med Phys 32, 108-117.
Hawkins, R.A., Mazziotta, J.C., Phelps, M.E., Huang, S.C., Kuhl, D.E., Carson, R.E., Metter, E.J., Riege, W.H., 1983. Cerebral glucose metabolism as a function of age in man: influence of the rate constants in the fluorodeoxyglucose method. J Cereb Blood Flow Metab 3, 250-253.
Herholz, K., Salmon, E., Perani, D., Baron, J.C., Holthoff, V., Frolich, L., Schonknecht, P., Ito, K., Mielke, R., Kalbe, E., Zundorf, G., Delbeuck, X., Pelati, O., Anchisi, D., Fazio, F., Kerrouche, N., Desgranges, B., Eustache, F., Beuthien-Baumann, B., Menzel, C., Schroder, J., Kato, T., Arahata, Y., Henze, M., Heiss, W.D., 2002. Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. Neuroimage 17, 302-316.
Ibanez, V., Pietrini, P., Alexander, G.E., Furey, M.L., Teichberg, D., Rajapakse, J.C., Rapoport, S.I., Schapiro, M.B., Horwitz, B., 1998. Regional glucose metabolic abnormalities are not the result of atrophy in Alzheimer's disease. Neurology 50, 1585-1593.
Ibanez, V., Pietrini, P., Furey, M.L., Alexander, G.E., Millet, P., Bokde, A.L., Teichberg, D., Schapiro, M.B., Horwitz, B., Rapoport, S.I., 2004. Resting state brain glucose metabolism is not reduced in normotensive healthy men during aging, after correction for brain atrophy. Brain Res Bull 63, 147-154.
Izquierdo-Garcia, D., Hansen, A.E., Forster, S., Benoit, D., Schachoff, S., Furst, S., Chen, K.T., Chonde, D.B., Catana, C., 2014. An SPM8-based approach for attenuation correction combining segmentation

and nonrigid template formation: application to simultaneous PET/MR brain imaging. J Nucl Med 55, 1825-1830.
Jan, S., Comtat, C., Strul, D., Santin, G., trebossen, R., 2005. Monte Carlo Simulation for the ECAT EXACT HR+ System Using GATE. IEEE Transactions on Nuclear Science 52, 627-633.
Kalpouzos, G., Chetelat, G., Baron, J.C., Landeau, B., Mevel, K., Godeau, C., Barre, L., Constans, J.M., Viader, F., Eustache, F., Desgranges, B., 2009. Voxel-based mapping of brain gray matter volume and glucose metabolism profiles in normal aging. Neurobiol Aging 30, 112-124.
Kantarci, K., Jack, C.R., Jr., Xu, Y.C., Campeau, N.G., O'Brien, P.C., Smith, G.E., Ivnik, R.J., Boeve, B.F., Kokmen, E., Tangalos, E.G., Petersen, R.C., 2001. Mild cognitive impairment and Alzheimer disease: regional diffusivity of water. Radiology 219, 101-107.
Knopman, D.S., Jack, C.R., Jr., Wiste, H.J., Lundt, E.S., Weigand, S.D., Vemuri, P., Lowe, V.J., Kantarci, K., Gunter, J.L., Senjem, M.L., Mielke, M.M., Roberts, R.O., Boeve, B.F., Petersen, R.C., 2014. 18F-fluorodeoxyglucose positron emission tomography, aging, and apolipoprotein E genotype in cognitively normal persons. Neurobiol Aging 35, 2096-2106.
Kochunov, P., Ramage, A.E., Lancaster, J.L., Robin, D.A., Narayana, S., Coyle, T., Royall, D.R., Fox, P., 2009. Loss of cerebral white matter structural integrity tracks the gray matter metabolic decline in normal aging. Neuroimage 45, 17-28.
Kuhl, D.E., Metter, E.J., Riege, W.H., Phelps, M.E., 1982. Effects of human aging on patterns of local cerebral glucose utilization determined by the [18F]fluorodeoxyglucose method. J Cereb Blood Flow Metab 2, 163-171.
Labbé, C., Koepp, M., Ashburner, J., Spinks, T., Richardson, M., Duncan, J., Cunningham, V., 1998. Absolute PET Quantification with Correction for Partial Volume Effects with Cerebral Structures. Quantitative Functional Brain Imagaing with Positron Emission Tomography, pp. 59-66.
Landau, S.M., Harvey, D., Madison, C.M., Koeppe, R.A., Reiman, E.M., Foster, N.L., Weiner, M.W., Jagust, W.J., Alzheimer's Disease Neuroimaging, I., 2011. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. Neurobiol Aging 32, 1207-1218.
Loessner, A., Alavi, A., Lewandrowski, K.U., Mozley, D., Souder, E., Gur, R.E., 1995. Regional cerebral function determined by FDG-PET in healthy volunteers: normal patterns and changes with age. J Nucl Med 36, 1141-1149.
Meltzer, C.C., Kinahan, P.E., Greer, P.J., Nichols, T.E., Comtat, C., Cantwell, M.N., Lin, M.P., Price, J.C., 1999. Comparative evaluation of MR-based partial-volume correction schemes for PET. J Nucl Med 40, 2053-2065.
Meltzer, C.C., Leal, J.P., Mayberg, H.S., Wagner, H.N., Jr., Frost, J.J., 1990. Correction of PET data for partial volume effects in human cerebral cortex by MR imaging. J Comput Assist Tomogr 14, 561-570.
Meltzer, C.C., Zubieta, J.K., Links, J.M., Brakeman, P., Stumpf, M.J., Frost, J.J., 1996. MR-based correction of brain PET measurements for heterogeneous gray matter radioactivity distribution. J Cereb Blood Flow Metab 16, 650-658.
Minoshima, S., Frey, K.A., Foster, N.L., Kuhl, D.E., 1995. Preserved pontine glucose metabolism in Alzheimer disease: a reference region for functional brain image (PET) analysis. J Comput Assist Tomogr 19, 541-547.
Moeller, J.R., Ishikawa, T., Dhawan, V., Spetsieris, P., Mandel, F., Alexander, G.E., Grady, C., Pietrini, P., Eidelberg, D., 1996. The metabolic topography of normal aging. J Cereb Blood Flow Metab 16, 385-398.
Mormino, E.C., Betensky, R.A., Hedden, T., Schultz, A.P., Amariglio, R.E., Rentz, D.M., Johnson, K.A., Sperling, R.A., 2014. Synergistic Effect of β-Amyloid and Neurodegeneration on Cognitive Decline in Clinically Normal Individuals. JAMA Neurology.
Müller-Gärtner, H.W., Links, J.M., Prince, J.L., Bryan, R.N., McVeigh, E., Leal, J.P., Davatzikos, C., Frost, J.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, 571-583.

Nugent, S., Castellano, C.A., Goffaux, P., Whittingstall, K., Lepage, M., Paquet, N., Bocti, C., Fulop, T., Cunnane, S.C., 2014a. Glucose hypometabolism is highly localized, but lower cortical thickness and brain atrophy are widespread in cognitively normal older adults. Am J Physiol Endocrinol Metab 306, E1315-1321.

Nugent, S., Tremblay, S., Chen, K.W., Ayutyanont, N., Roontiva, A., Castellano, C.A., Fortier, M., Roy, M., Courchesne-Loyer, A., Bocti, C., Lepage, M., Turcotte, E., Fulop, T., Reiman, E.M., Cunnane, S.C., 2014b. Brain glucose and acetoacetate metabolism: a comparison of young and older adults. Neurobiol Aging 35, 1386-1395.

Oakes, T.R., Fox, A.S., Johnstone, T., Chung, M.K., Kalin, N., Davidson, R.J., 2007. Integrating VBM into the General Linear Model with voxelwise anatomical covariates. Neuroimage 34, 500-508.

Petit-Taboue, M.C., Landeau, B., Desson, J.F., Desgranges, B., Baron, J.C., 1998. Effects of healthy aging on the regional cerebral metabolic rate of glucose assessed with statistical parametric mapping. Neuroimage 7, 176-184.

Quarantelli, M., Berkouk, K., Prinster, A., Landeau, B., Svarer, C., Balkay, L., Alfano, B., Brunetti, A., Baron, J.C., Salvatore, M., 2004. Integrated software for the analysis of brain PET/SPECT studies with partial-volume-effect correction. J Nucl Med 45, 192-201.

Redzic, Z.B., Preston, J.E., Duncan, J.A., Chodobski, A., Szmydynger-Chodobska, J., 2005. The choroid plexus-cerebrospinal fluid system: from development to aging. Curr Top Dev Biol 71, 1-52.

Reilhac, A., Tomei, S., Buvat, I., Michel, C., Keheren, F., Costes, N., 2008. Simulation-based evaluation of OSEM iterative reconstruction methods in dynamic brain PET studies. Neuroimage 39, 359-368.

Rmeily-Haddad, M., Baledent, O., Stoquart-ElSankari, S., Serot, J.-M., Baily, P., Meyer, M.-E., 2011. The kinetics of 18f-fluorodeoxyglucose uptake in the choroid plexus. International Journal of Imaging Systems and Technology 21, 107-144.

Roos, K.L., 2004. Principles of Neurologic Infectious Diseases: Principles and Practice. . McGraw-Hill Professional.

Rousset, O.G., Ma, Y., Evans, A.C., 1998a. Correction for partial volume effects in PET: principle and validation. J Nucl Med 39, 904-911.

Rousset, O.G., Ma, Y., Wong, D.F., Evans, A.C., 1998b. Pixel versus Region-based Partial Volume Correction in PET. Quantitative Functional Brain Imaging with Positron Emission Tomography, pp. 67-75.

Sattarivand, M., Kusano, M., Poon, I., Caldwell, C., 2012. Symmetric geometric transfer matrix partial volume correction for PET imaging: principle, validation and robustness. Phys Med Biol 57, 7101-7116.

Segonne, F., Pacheco, J., Fischl, B., 2007. Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. IEEE Trans Med Imaging 26, 518-529.

Thomas, B.A., Erlandsson, K., Modat, M., Thurfjell, L., Vandenberghe, R., Ourselin, S., Hutton, B.F., 2011. The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease. Eur J Nucl Med Mol Imaging 38, 1104-1119.

Uchida, H., Chow, T.W., Mamo, D.C., Kapur, S., Mulsant, B.H., Houle, S., Pollock, B.G., Graff-Guerrero, A., 2011. Effects of aging on 5-HT(2A) R binding: a HRRT PET study with and without partial volume corrections. Int J Geriatr Psychiatry 26, 1300-1308.

Walhovd, K.B., Fjell, A.M., Dale, A.M., McEvoy, L.K., Brewer, J., Karow, D.S., Salmon, D.P., Fennema-Notestine, C., Alzheimer's Disease Neuroimaging, I., 2010. Multi-modal imaging predicts memory performance in normal aging and cognitive decline. Neurobiol Aging 31, 1107-1121.

Yanase, D., Matsunari, I., Yajima, K., Chen, W., Fujikawa, A., Nishimura, S., Matsuda, H., Yamada, M., 2005. Brain FDG PET study of normal aging in Japanese: effect of atrophy correction. Eur J Nucl Med Mol Imaging 32, 794-805.

Yoshii, F., Barker, W.W., Chang, J.Y., Loewenstein, D., Apicella, A., Smith, D., Boothe, T., Ginsberg, M.D., Pascal, S., Duara, R., 1988. Sensitivity of cerebral glucose metabolism to age, gender, brain volume, brain atrophy, and cerebrovascular risk factors. J Cereb Blood Flow Metab 8, 654-661.

Yoshizawa, H., Gazes, Y., Stern, Y., Miyata, Y., Uchiyama, S., 2014. Characterizing the normative profile of 18F-FDG PET brain imaging: sex difference, aging effect, and cognitive reserve. Psychiatry Res 221, 78-85.

Table 1. Comparison of various partial volume correction (PVC) methods based on the number of GM ROIs (out of 41) with significant age effect (N ROIs)

	Method	N ROIs	Notes
1	NoPVC	27	No Partial Volume Correction
2	NoPVC-VolReg	25	No Partial Volume Correction with Volume Regression
3	PVBA	10	Partial Volume Bias Adjustment (subtract slope of age bias)
4	SGTM-Full	12	Symmetric Geometric Transfer Matrix (SGTM) with full head model
5	SGTM-NX	5	SGTM without CSF or other extracerebral ROIs
6	MG-Full	12	Müller-Gärtner with full head model thresholded at .30
7	MG-NX	5	MG without CSF or other extracerebral ROIs; thresholded at .30
8	MZ-10-D1	26	Meltzer; brain mask threshold=0.10 and dilated by 1
9	MZ-10	12	Meltzer; brain mask threshold=0.10
10	MZ-50	2	Meltzer; brain mask threshold=0.50

Table 2. Comparison of methods on an ROI basis based on significant age effect. ROI No matches that in Tables S1 and S2. Count is the number of times the ROI was significant over all methods. Method numbers and names match those in Table 1; the "Total Number of ROIs" is the number of significant ROIs for that method. NoPVC Sim is the simulation of the NoPVC method. * Significant negative age correlation (p<.001 uncorrected). No ROIs in this table had significant positive activation. Volume, Surface Area, and Thickness are from the anatomical analysis in Table S2. X Not applicable. See Table S3 for actual p-values for each cell.

ROI No	Region-of-Interest	Count	1 NoPVC	2 NoPVC-VolReg	3 PVBA	4 SGTM-Full	5 SGTM-NX	6 MG-Full-30	7 MG-NX-30	8 MZ-10-D1	9 MZ-10	10 MZ-50	NoPVC Sim	Volume	Surface Area	Thickness	SGTM*Volume
1	Medialorbitofrontal	9	*	*	*	*	*	*	*	*	*						
2	Superior Frontal	7	*	*	*	*		*		*	*						*
3	Rostral Middle Frontal	10	*	*	*	*	*	*	*	*	*	*					*
4	Caudal Middle Frontal	6	*	*	*	*		*		*							*
5	Lateralorbitofrontal	9	*	*	*	*	*	*	*	*	*						*
6	Pars Orbitalis	5			*	*	*	*	*								
7	Pars Triangularis	0															
8	Pars Opercularis	7	*	*	*	*		*		*	*						*
9	Rostral Anterior Cingulate	5	*	*	*			*		*							
10	Caudal Anterior Cingulate	9	*	*	*	*		*	*	*	*	*					
11	Posterior Cingulate	4	*	*						*	*						*

12	Isthmus Cingulate	6	*	*		*		*		*	*		*				
13	Insula	6	*	*		*		*		*	*						*
14	Supramarginal	2	*							*			*				*
15	Primary Auditory	3	*	*						*			*				*
16	Banks of the STS	4	*	*						*	*		*				
17	Superior Temporal	3	*	*						*			*				*
18	Middle Temporal	3	*	*						*							*
19	Inferior Temporal	1		*											*		*
20	Fusiform	3	*	*						*			*	*	*		*
21	Entorhinal	1	*										*				*
22	Parahippocampal	0											*	*			*
23	Superior Parietal	3	*	*						*				*	*		*
24	Inferior Parietal	3	*	*						*							*
25	Precuneus	3	*	*						*			*				*
26	Primary Visual	8	*	*	*	*	*	*		*	*		*				*
27	Cuneus	3	*	*						*			*				*
28	Lingual	5	*	*		*				*	*		*				*
29	Lateral Occipital	1		*									*				*
30	Precentral	3	*	*						*			*				*
31	Postcentral	2	*							*			*	*			*
32	Paracentral	0															
33	Hippocampus	0											*	*	X	X	*
34	Amygdala	0												*	X	X	
35	Caudate	0													X	X	
36	Putamen	0													X	X	
37	Pallidum	0													X	X	
38	Accumbens	0												*	X	X	
39	Thalamus	2	*							*			*	*	X	X	*
44	Cerebellum Cortex	0													X	X	
45	Vermis	0											*		X	X	
	Total number of ROIs		27	25	10	12	5	12	5	26	12	2	18	8	3	0	29

Table 3. List of ROIs accounting for half of the change in total GM ^{18}F-FDG changes with age. The % column indicates the percentage of the GM changed accounted for by that ROI. The Cum% column is the cumulative percent accounted for.

ROI	%	Cum%
Cerebellum Cortex	8.3%	8.3%
Superior Frontal	8.1%	16.3%
Superior Parietal	7.1%	23.5%
Rostral Middle Frontal	6.4%	29.8%
Lateral Occipital	4.7%	34.5%

Inferior Parietal	4.4%	39.0%
Precuneus	4.3%	43.3%
Precentral	4.3%	47.5%

Highlights

- A study of partial volume correction (PVC) methods for FDG of aging is performed.
- Methods: no PVC, Meltzer, Müller-Gärtner, symmetric geometric transfer matrix.
- Testing was performed with and without modeling of cerebral spinal fluid (CSF).
- Different methods yielded a wide range of results for the same data.
- CSF had a surprisingly high FDG signal and strongly influenced PVC.