Improved brain PET quantification using partial volume correction techniques

Benjamin Aneurin Thomas

UCL

A thesis submitted to University College London
for the degree of Doctor of Philosophy
in Medical Physics applied to Nuclear Medicine.

March 2012

I, Benjamin Aneurin Thomas, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Positron emission tomography (PET) suffers from a degradation in quantitative accuracy due to a phenomenon known as the partial volume effect (PVE). The effects are due to the limited spatial resolution of the scanner. Methods that correct for PVEs are known as partial volume correction (PVC) techniques and are either data-driven or make use of anatomical information from other modalities such as magnetic resonance (MR) imaging. This thesis reports investigations into PVC techniques for improving the quantification of brain amyloid PET tracers. These tracers image amyloid plaque aggregation in-vivo, which is a pathological hallmark of Alzheimer's disease.

An extension to existing anatomy-based PVC methods is reported. Region-based voxel-wise (RBV) correction has been shown to reduce PVE-induced regional bias and variance when compared to commonly applied PVC techniques. This has been proven in phantom studies and observed in clinical data. In addition, RBV has been used to demonstrate that white matter variability exists in two different amyloid tracers. This finding has implications for the application of PVC in amyloid imaging and also how scans should be normalised. Alternative reference regions were investigated in two amyloid PET tracers. The brain stem, in combination with PVC, was found to result in the strongest agreement between tracers.

Anatomy-based PVC techniques rely on parcellations of structural images. These parcellations are not necessarily representative of the PET data. A further extension to RBV is proposed which iteratively modifies the parcellations to find an optimal PVC in terms of the observed PET data. This novel technique reduces quantification errors due to PET-MR mismatch and has the potential to provide an additional parameter of 'functional volume change' in longitudinal studies.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my primary supervisor, Brian Hutton, whose ideas, discussions and encouragement were hugely important to the progress of this project. I would like to thank my secondary supervisor, Sebastien Ourselin, for his useful input into the project. Thank you also to my industrial supervisor, Lennart Thurfjell, for his contributions and for providing access to much of the data used throughout this work.

I would like to thank all members of the Institute of Nuclear Medicine, UCL for their support throughout my time in the department. Thank you also to everyone in the physics research group who have supported me during my time at the INM. In particular, I would like to thank Kjell Erlandsson for his help, advice and discussions surrounding my work and for the constructive feedback on my thesis. I am also very grateful to John Dickson and Ian Cullum for their assistance and support during this project.

Finally, a big thank you to my family and friends, in particular to my parents for their perpetual support and encouragement. And I must also express my deep gratitude to Annja, for her care, kindness and backing throughout.

Contents

List of Figures

List of Tables

List of Abbreviations

2D two dimensional

3D three dimensional

AC Attenuation correction

AD Alzheimer's disease

ADNI Alzheimer's Disease Neuroimaging Initiative

AIBL Australian Imaging Biomarkers and Lifestyle flagship study of ageing

AIR Adobe Integrated Runtime

ANOVA analysis of variance

APD avalanche photodiode

CC Cross-correlation

CDR Clinical Dementia Rating

CWM cerebellar white matter

CoV coefficient of variation

CSF cerebro-spinal fluid

CSV comma separated variable

CT Computed Tomography

DC Dice Coefficient

DICOM Digital Imaging and Communications in Medicine

DVR Distribution Volume Ratio

eHC elderly healthy control

EM Expectation-Maximisation

FBP Filtered Back Projection

FFT fast fourier transform

fMRI functional Magnetic Resonance Imaging

FOV field of view

FS FreeSurfer

FTD Fronto-temporal dementia

FWHM full-width at half-maximum

GE General Electric

GM grey matter

GTM Geometric Transfer Matrix

h-PVC Hyper-plane partial volume correction

ITK Insight Segmentation and Registration Toolkit

iY iterative Yang

HC healthy control

HU Hounsfield units

LOR line of response

MCI Mild Cognitive Impairment

MDD Minimum Detectable Difference

MI Mutual Information

MLEM Maximum Likelihood Expectation-Maximisation

MMSE Mini Mental State Examination

MRI Magnetic Resonance Imaging

MG Müller-Gärtner

mMG modified Müller-Gärtner

MRF Markov Random Field

MTC Multi-target correction

NEMA National Electrical Manufacturers Association

NFTs neurofibrillary tangles

NIfTI Neuroimaging Informatics Technology Initiative

NMI Normalised Mutual Information

NRR Non-rigid registration

NSEE normalised standard error of the estimate

OSEM Ordered Subset Expectation Maximisation

PET Positron Emission Tomography

PMT photomultiplier tube

PSF point-spread function

PVC Partial Volume Correction

PVE partial volume effect

RAMLA Row action Maximum-Likelihood algorithm

RBV region-based voxel-wise

RC recovery coefficient

RL Richardson-Lucy

ROI region of interest

RR Resolution Recovery

SD standard deviation

SMPVC segmentation-modifying partial volume correction

SPECT Single Photon Emission Computed Tomography

SPM Statistical Parametric Mapping

SSD sum of squared differences

SUVR standardised uptake value ratio

SWM sub-cortical white matter

TAC time-activity curve

TLS trimmed least squares

VC van Cittert

VOIs volumes of interest

WM white matter

WMR cerebellar WM to sub-cortical WM ratio

yHC young healthy control

Chapter 1

Introduction

Emission tomography is degraded by a phenomenon known as the partial volume effect (PVE). This effect is caused by the limited spatial resolution of the scanner and affects the quantitative accuracy of the observed images. The spatial resolution is determined by a number of factors relating to scanner design and the nature of radiation. It is this finite resolution that causes PVEs (see section 2.2, page 27), resulting in blurred images. PVEs change the apparent position of activity in an image.

The degree to which a structure will suffer PVEs depends on its size. Smaller structures tend to be more severely affected by the PVE than larger ones. Neurodegenerative diseases, such as Alzheimer's disease (AD), are characterised by atrophic changes, particularly in cortical grey matter (GM). Atrophy causes the thinning of cortical structures, which in turn become more susceptible to PVEs.

A diagnosis of AD is pathologically confirmed by the existence of amyloid-β plaques in GM. Positron Emission Tomography (PET) ligands are now available that can image amyloid-β *in-vivo*. These tracers have application both as diagnostic and prognostic tools. However, GM regions are known to suffer from large PVEs due to the relatively poor spatial resolution, typically 5 - 6mm full-width at half-maximum (FWHM). These PVEs may obscure patterns of disease that could be important in the search for AD biomarkers.

Partial Volume Correction (PVC) techniques aim to compensate for the effects of resolution, thereby removing PVEs and improving quantitative accuracy. These corrections are either data-driven or utilise additional information provided by structural imaging. Both data-driven and anatomy-based PVC approaches have been applied to PET studies, although PVC is not routinely applied in clinical practice. This is especially the case for amyloid PET tracers. The PVC techniques that have been applied are well documented however they make assumptions about regional uniformity that may not be appropriate for amyloid imaging.

The primary objective of this project was to develop and evaluate the performance of

PVC techniques with an emphasis on the post-reconstruction correction of amyloid PET data. To achieve this objective, it was necessary to investigate methods that can handle variability *within* tissue compartments as well as *between* tissues, while comparing their performance against common approaches. This has led to the creation of a novel PVC method which modifies Magnetic Resonance Imaging (MRI) region definitions and is driven by the observed PET data.

Amyloid PET images are typically normalised using the cerebellar GM as a reference region. The cerebellar GM is believed to be relatively free of amyloid-β deposits, even in diseased subjects. However, the structure is bordered by white matter (WM) which exhibited higher activity concentrations, potentially resulting in PVEs in the reference region. Therefore a further investigation into alternative reference regions, in combination with PVC, is reported.

The application of appropriate PVC to PET data will improve quantitative accuracy. This, coupled with the choice of reference region, may prove to be particularly important in longitudinal studies where functional changes can be masked by PVEs caused by structural changes.

1.1 Structure of the thesis

The experimental work in this thesis can be divided into three parts; the investigation of a modified PVC method compared with alternative PVC techniques using phantom data, evaluating PET normalisation with PVC in a clinical cohort and the development of a novel PVC technique with application in longitudinal studies.

The phenomenon of PVE in PET is discussed in chapter 2. PVC techniques used throughout the experimental chapters of the thesis are also introduced. In addition, the background chapter focuses on PET tracers that can image amyloid-β plaques, which are a hallmark of AD.

Chapter 3 describes two investigations. The first evaluates the effects of PVC using phantom amyloid PET data. The performance of anatomy-based and data-driven PVC methods, including the recently suggested refinement, are compared through regional analysis. The second investigation examines noise suppression methods for data-driven PVC approaches.

Clinical data from a multi-centre trial of a new amyloid tracer: [^{18}F]flutemetamol is assessed in chapter 4. The performance of a common anatomy-based PVC method is compared against the new technique. The quantitative reproducibility of [^{18}F]flutemetamol is also assessed. In addition, a comparison between two amyloid tracers: [^{11}C]PIB and

[^{18}F]flutemetamol is performed along with an investigation into alternative reference regions for amyloid PET normalisation.

Most anatomy-based PVC techniques make an assumption of regional uniformity. The regions defined from structural imaging modalities such as MRI do not necessarily represent the regions in the PET distribution. Chapter 5 describes the development and evaluation of a novel PVC technique in which the parcellation is iteratively modified in order to improve the PVC. This modification is driven by the PET data. The technique is evaluated using phantom data, clinical reproducibility and longitudinal data.

Finally, a summary of the findings of each chapter, along with potential areas for further investigation, are discussed in chapter 6.

Chapter 2

Background

This chapter provides a background to the main topics that are covered in this thesis. Section 2.1 introduces the acquisition of PET data. The issues of PVE in PET, which are central to this thesis, are then introduced in section 2.2. Existing approaches to PVC are discussed in section 2.3. Image processing applied to PET data is then summarised in section 2.4. An overview of AD and associated amyloid PET tracers are given in sections 2.5 and 2.6. The chapter concludes with a discussion in section 2.7.

2.1 PET image formation

2.1.1 Radioactive decay

PET scanners measure γ-rays that are the result of radioactive decay by positron emission. The spatial resolution of the scanner is limited by a number of factors due to behaviour of radioactive decay. As a nucleus undergoes decay (figure 2.1), a positron is emitted. The positron travels a distance before interacting with an electron and then annihilating. This distance is the minimum resolution achievable in PET and is typically quoted as the *mean positron range* of the nuclide. The positron range causes blurring and the extent of the blurring is dependent on which isotope is used and the medium in the proximity of the positron emission.

The annihilation of the positron and electron produces two γ-rays, each with an energy of 511 keV. The γ-rays are emitted at an angle of $180° \pm 0.5°$. The uncertainty in the angle also blurs the image. The degree of blurring due to the variation in the colinearity of the γ-rays depends on the distance between the pair of detectors. The closer the detectors are to the point of annihilation, the lower the error will be. The blurring is described as being due to the *non-colinearity*.

PET relies on the ability to localise annihilation events. This is acheived using detectors which typically consist of a scintillation crystal coupled to a photomultiplier tube (PMT). Photons interact with the crystal, producing visible light, which is then converted into an

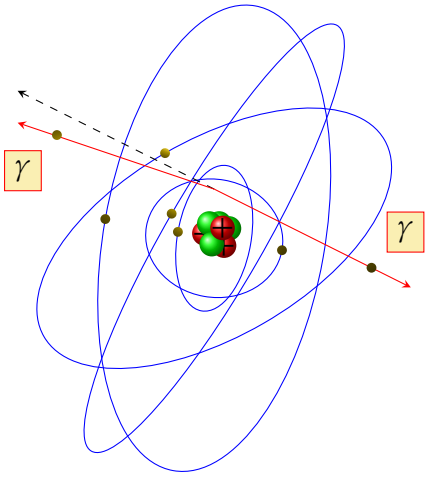


Figure 2.1: A nucleus undergoing positron decay. Red lines indicate the paths of the γ-rays. The dashed line represents the colinear path.

electrical signal. The detection of pairs of γ-photons allows the source of an event to be localised along a straight line between two detectors, referred to as the line of response (LOR). In order to completely sample an object, it is necessary for the detectors to cover 360°. As a result, detectors are typically arranged in a circular ring formation.

PET detectors often consist of a two dimensional (2D) array (or block) of scintillation crystals, coupled to PMTs. A PET gantry contains multiple rings of block detectors. In the case of 2D acquisition, events between conincident detectors, either in the same or closely neighbouring crystal ring(s), are recorded to form a single plane. Septa made of lead or tungsten are used to separate rings in order that only events which occured between these crystal rings are recorded. When perfoming three dimensional (3D) acquisition, the septa are removed and events which occur between detectors lying in any combination of rings can be recorded.

The acquired planes are often stored in *sinograms*. In a sinogram, LORs are are plotted as a function of the angle of orientation against the displacement from the centre of the scanner. The value at each point in the sinogram represents the number of events detected by the pair of detectors which are associated with a given LOR. Each row of the sinogram is a *projection* through the object at a particular angle. The set of all projections constitutes the sinogram for a single plane. The sinograms from each plane are then reconstructed to generated a 3D volume.

When the γ-rays are emitted, the detected γ-ray pairs may be either *unscattered*, *scattered* or *random*. Unscattered (or true) γ-ray pairs are those where both γ-rays have reached the detectors with neither having been scattered during transmission to the point of detection. If

one or both γ-rays are scattered as they travel to the detectors, the LOR will no longer pass through the point at which the annihilation occurred. Scattered pairs degrade image quality as they cause an erroneous LOR to be recorded. The detection of random γ-ray pairs can also occur when two photons from different annihilations reach the detectors at the same time and therefore appear to be from the same event. Random events also cause an erroneous measurement of the LOR.

Each detected pair of γ-rays is called a *coincidence*. PET images are formed from millions of these events (counts). As the data is acquired, these counts are stored in sinograms (figure 2.2). Each projection represents the counts observed by pairs of detectors whose LORs are parallel to each other. When acquiring in 2D mode, to form a 3D data set, each plane in the axial direction is stored in a sinogram. The set of sinograms can then be used to reconstruct a 3D volume (see section 2.1.5). Alternatively, data may be acquired in 3D, where coincidences between detectors in different rings are recorded. The advantage of 3D over 2D is that sensitivity is substantially increased. However, the amount of both scatter and random coincidences also increases. Data acquired in 3D requires either a specific 3D reconstruction algorithm or for the data to be rebinned for use with 2D reconstruction.

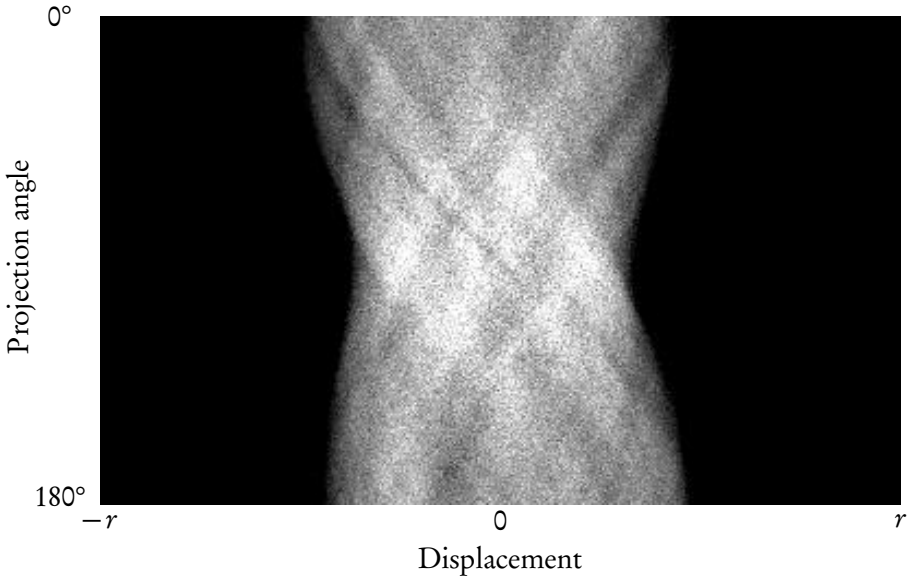


Figure 2.2: Example of a PET sinogram from a single slice. The y-axis shows the angle of the projection and x-axis is the distance from the centre of rotation, with r is the radius of the scanner.

2.1.2 Attenuation

The annihilations that are of interest are those which occur within human tissue. PET relies on the ability to detect coincident pairs of γ-rays. However, this requires both photons to

penetrate through tissue and reach the detectors. Photons can be absorbed by tissue(s) and other objects in the scanner field of view (FOV). This effect is known as *attenuation*. The effects of attenuation can dramatically affect the quantitative accuracy of PET. Attenuation correction (AC) is routinely applied in clinical practice and there are several approaches.

In PET, the attenuation along a LOR is independent of the depth of the annihilation. This independence can be used to correct for attenuation in the form of a *transmission scan*. The scan historically consisted of one or more external source(s) (typically ^{68}Ge or ^{137}Cs) being rotated about the patient, within the FOV. These data are used to form an attenuation map from which *attenuation correction factors* can be found. The correction factors are then applied to the measured projections, producing an attenuation-corrected sinogram.

Computed Tomography (CT) is now almost exclusively utilised for AC. A CT scan gives a quantitative measure of the attenuation coefficients of X-rays in tissue, given in Hounsfield units (HU). The attenuation map can be derived from the CT by converting the HU to suitable values for 511 keV. CT-based AC has advantages over rotating emission sources; it is much faster (a few seconds compared to 15-20 minutes with sources) and has less noise [Kinahan et al., 1998].

2.1.3 Poisson noise

Radioactive decay is a stochastic process. As a result, the number of counts observed in two measurements will almost always be different. The number of counts is proportional to both the acquisition time and the size of the voxels. This behaviour can be described by a Poisson distribution. The relative noise amplitude in PET data is $\frac{\sqrt{n}}{n}$, where n is the number of measured counts. Therefore, data acquired over a shorter time period will contain more noise. Smaller voxel sizes result in fewer counts-per-voxel, also leading to increased noise. The Poisson-distributed noise in the sinogram data will also affect the reconstructed volume, with noisier projection data producing noisier reconstructions.

2.1.4 Dectectors

γ-rays are measured by the detectors of the PET scanner. A detector consists of scintillator crystals coupled to a set of PMTs. The coupling of crystals and detectors can be configured in different ways. Anger [1958] proposed that an array of PMTs be coupled to a single crystal. This is the most common configuration in Single Photon Emission Computed Tomography (SPECT). A *block detector* consisting of a 2D array of scintillator crystals, coupled to PMTs [Casey and Nutt, 1986] is a common configuration in PET. The vast majority of detectors contain PMTs although avalanche photodiodes (APDs) are beginning to be used

in new scanners, especially in pre-clinical systems.

The size of the crystal also influences the spatial resolution of the scanner. The intrinsic resolution of the PET scanner is related to the size of the detector. This intrinsic resolution varies across the FOV. At the centre of the FOV the intrinsic resolution is usually assumed to be half the detector size, whereas at the face of the detector (the edge of the FOV) the intrinsic resolution increases to the size of the detector [Saha, 2010, p. 97].

2.1.5 Reconstruction

In order to produce a 3D image of activity, a reconstruction operation is required. Corrections for attenuation, scatter and randoms are applied to the sinograms prior to reconstruction. The reconstruction is performed on the projection data contained in the sinograms. Filtered Back Projection (FBP) can be used to reconstruct PET projection data. The algorithm back-projects the activity over the LORs. When used to reconstruct PET data, streaking artefacts tend to occur due to the Poisson distributed noise in the data and the implicit assumption in FBP of Gaussian noise [Schiepers, 2006, pp. 292–305].

Iterative reconstruction techniques allow the processes of image formation to be modeled. Maximum Likelihood Expectation-Maximisation (MLEM) is an iterative algorithm that updates an estimate of the reconstructed image by comparing the projections of the estimate with the observed projection data. The ratio of the two is used to update the estimate and this process is repeated until convergence or some stopping criterion is met. However, MLEM is slow, with a single MLEM iteration taking roughly twice as long as an FBP reconstruction [Schiepers, 2006, p. 295]. Hudson and Larkin [1994] proposed Ordered Subset Expectation Maximisation (OSEM) as an approach to accelerate MLEM reconstruction. OSEM considers a subset of the projection data at each iteration and has been shown to be computationally more efficient compared to MLEM.

Filtering, applied either during the reconstruction (as in FBP) or post-reconstruction, may degrade (increase) the resolution of the reconstructed images. A filter with a wide cutoff can blur the reconstructed image, whereas a filter with a narrow cutoff may permit large amounts of high frequency noise, resulting in visually unacceptable images. The choice of filter depends partly on application (quantification vs. visual assessment), but different filters will result in different reconstructed resolutions.

The studies used in this project consist of a mixture of data reconstructed using FBP and OSEM. All phantom data were reconstructed with FBP. The clinical datasets were acquired during multi-centre trials and were reconstructed according to the participating site's imaging protocol.

2.1.6 Spatial resolution

The spatial resolution of the PET scanner is a combination of the factors described above. The positron range, photon non-colinearity, detector size and reconstruction, all contribute to the overall reconstructed spatial resolution. The resolution of the PET scanner (R_t) is given by [Cherry et al., 2003, p. 333]:

$$R_t = \sqrt{R_i^2 + R_p^2 + R_a^2 + R_r^2} \tag{2.1}$$

where R_i is the resolution component due to the detector size, R_p is the resolution component due to the positron range, R_a is the resolution component due to the non-colinearity of the photons and R_r is the resolution component due to the reconstruction. The spatial resolution degrades the quantitative accuracy of the measured PET image. This work focuses on the correction of PET images for PVEs, which are a phenomenon caused by the limited spatial resolution of the scanner.

2.2 Partial volume effects

The quantitative accuracy of images in emission tomography is degraded by the limited spatial resolution of the tomograph [Aston et al., 2002]. As discussed previously, the reconstructed image resolution is limited by a number of factors. The mean positron range, non-colinearity of γ-rays, detector size and reconstruction parameters, all affect the reconstructed resolution. PET has a relatively poor resolution of between 5 and 6 mm FWHM and it is this poor spatial resolution that causes PVEs. These figures for resolution are based on the clinical data used in this work. Higher resolution brain imaging is reported with scanners such as the Siemens ECAT EXCAT HR+.

The term "partial volume effect" actually refers to two effects that degrade the quantitative accuracy. The first is the blurring due to resolution. The second, known as the *tissue-fraction effect*, is related to the grid of voxels on which the image is sampled. The finite spatial resolution causes a 3D blurring throughout the image. The blurring causes activity to spill-out of and into neighbouring regions [Soret et al., 2007]. Counts are not 'lost', but simply displaced from their true location due to the point-spread function (PSF) of the scanner. The spilling-out and spilling-in of activity is not usually balanced, further complicating estimation of the amount of PVE.

The degree to which an object is affected by PVEs also depends on its size. The maximum value of structures that have a spatial extent of less than $2^1/_2$ to 3 times the FWHM will suffer from PVEs [Quarantelli et al., 2004] (see figure 2.3). When investigating neurode-

generative dieases, these PVEs become important as apparent changes in signal can at least be partially attributed to PVEs.

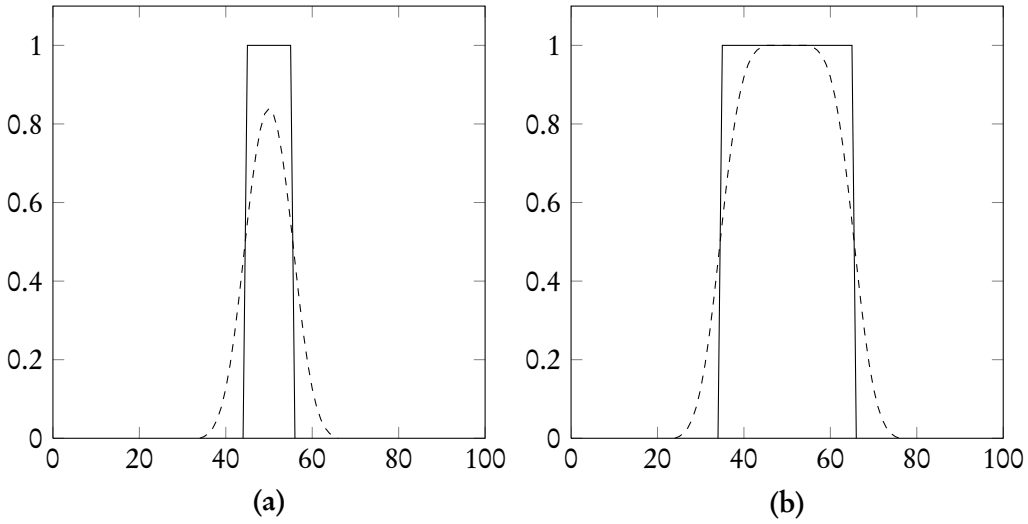


Figure 2.3: An example of the effect of convolution with a Gaussian PSF (10.5mm FWHM). The solid line represents the object before convolution and the dashed line represents the object after convolution. Plot (a) shows the effect on an object of size 10mm. The object in plot (b) is 30mm wide.

The effect of resolution can be thought of as a convolution of the true image with a 3D PSF [Tohka and Reilhac, 2008]. The resolution within the PET FOV is spatially-varying, that is, the size of the scanner PSF is different at distinct locations in the FOV. The PSF is normally smallest at the centre of the scanner, increasing towards the detectors. The effects of space variance tend to be more pronounced in the transaxial and radial vs. tangential than the axial dimensions [Mawlawi et al., 2004]. For the purposes of PVC in PET, the PSF tends to be approximated by an isotropic Gaussian.

The tissue-fraction effect is due to voxel size. PET images are sampled using a grid of voxels. This sampling can result in multiple tissue types existing within a single voxel. Each tissue type within a voxel may contribute differently to the tracer concentration. The value at a given voxel with mixed tissue types becomes the mean activity for all types that are present. Thus, the mean activity at the voxel does not necessarily describe the true tracer distribution accurately.

PVC techniques have been proposed which seek to correct for PVEs due to the spatial resolution. The tissue-fraction effect can only be accounted for with the inclusion of structural information from high resolution modalities such as CT or MRI [Aston et al., 2002]. A selection of different proposed correction techniques are described in the following section.

2.3 Approaches to partial volume correction

PVC approaches can either be included in the reconstruction or applied as a post-reconstruction correction. These can then be further sub-divided into region of interest (ROI)-based or voxel-based methods. In addition, some approaches are purely data-driven, while others utilise information from structural imaging modalities. This section describes PVC techniques that have been proposed for neurological applications.

2.3.1 Post-reconstruction PVC methods

2.3.1.1 Iterative deconvolution

Iterative deconvolution approaches are purely data-driven methods that correct for PVEs caused by the spatial blurring. The Richardson-Lucy (RL) [Lucy, 1974, Richardson, 1972] and van Cittert (VC) [van Cittert, 1931] are common deconvolution techniques. Both approaches are often used in the field of astronomy and more recently have been applied in PET imaging [Teo et al., 2007, Tohka and Reilhac, 2006, 2008]. RL was proposed by Richardson [1972] and independently by Lucy [1974], hence the name given to the algorithm. It uses a Bayesian approach and assumes a Poisson noise model. A non-negativity constraint is also required in order that a log-likelihood formulation can be derived [Tohka and Reilhac, 2008]:

$$\begin{aligned} \ell(t) &\propto \sum_x \left[f^+(x) \times \log(t \otimes h)(x) - (t \otimes h)(x) \right], \\ f^+(x) &= \max(f(x), 0). \end{aligned} \tag{2.2}$$

where h is the scanner PSF, f is the observed image, and t is the true image. Expectation-Maximisation (EM) is then used to maximise the log-likelihood with the following iterative rule:

$$\begin{aligned} t_{k+1}(x) &= t_k(x) \left[\frac{f^+(x)}{(t_k \otimes h)(x)} \right], \\ t_0(x) &= f^+(x). \end{aligned} \tag{2.3}$$

where t_k is the estimate of the deconvolved image at iteration k. This will converge to the maxiumum likelihood solution.

The VC deconvolution is an iterative least squares formulation to the problem of image restoration. Unlike the RL method, VC assumes that the noise process is Gaussian. When the VC approach is applied to PET imaging, it tends to be the *reblurred* VC that is applied rather than the classic VC. The VC technique uses a gradient decent approach to find the minimiser of a least squares criterion. The iterative process is additive rather than multiplicative as is the

case with RL. A full description of the VC method is given in section 3.2.1.2 (page 66).

However, as is discussed in greater detail in section 3.2 (page 65), deconvolution approaches suffer from noise amplification issues. The RL method is often prematurely terminated at 10-12 iterations due to noise. This results in incomplete recovery. Noise suppression, such as the inclusion of priors and regularisation have been applied during deconvolution [Boussion et al., 2009, Dey et al., 2006, Nunez and Llacer, 1990, Rudin et al., 1992]. Overcoming the noise amplification issues related to deconvolution remains a topic of research.

2.3.1.2 Meltzer's method

Anatomical information from structural imaging can also be utilised for the purposes of PVC. Patient-specific MRI or CT data can be registered to the emission data (or vice versa). Knowledge of the anatomy, such as the the location of different structures, can then be used to correct for PVEs. Videen et al. [1988] proposed a method for correcting a single target region, based on anatomical data. The approach corrects cerebral tissue for the effects of the 'non-brain' region. It is assumed that the non-brain region contains no activity. The correction is performed by convolving a binary mask of the cerebral tissue with the scanner PSF in order to calculate the correction factors at each voxel. The observed PET image is then divided by the correction factors on a voxel-by-voxel basis. This results in PVC of the cerebral tissue only. The observed image is thought of as the union of two compartments, convolved with a PSF:

$$f_O(x) = [f_{GM+WM}(x) + f_{NB}(x)] \otimes h(x), \tag{2.4}$$

where f_O is the observed image, f_{GM+WM} is uptake in the cerebral tissue, f_{NB} is the non-brain compartment and h is the PSF of the scanner. A binary map of the tissues is created:

$$p(x) = \begin{cases} 1, & \text{if } x \text{ is brain tissue} \\ 0, & \text{otherwise.} \end{cases} \tag{2.5}$$

The amount of uptake in the non-brain compartment is assumed to be zero, so the equation becomes:

$$f_O(x) = [f_{GM+WM}(x)p(x)] \otimes h(x). \tag{2.6}$$

The technique assumes that tissue concentration consists of small fluctuations about a larger mean value [Meltzer et al., 1999], allowing the approximation:

$$f_O(x) \approx f_{GM+WM}(x)[p(x) \otimes h(x)], \tag{2.7}$$

The uptake values can then be found by rearranging equation 2.7 into the form:

$$f_{GM+WM}(x) \approx \frac{f_O(x)}{p(x) \otimes h(x)}. \tag{2.8}$$

The implementation reported by Videen et al. [1988] operated in 2D. Meltzer et al. [1990] applied this approach in 3D and it has become known as Meltzer's method, although it is mathematically equivalent to that of Videen et al. [1988]. The drawback of this method is that PVEs between GM and WM regions are not taken into account.

2.3.1.3 Müller-Gärtner

Meltzer's method was extended to a three compartment model by Müller-Gärtner et al. [1992]. The approach uses a registered MRI image that is segmented into GM, WM and cerebro-spinal fluid (CSF). A single target region is corrected, typically the GM. The technique is performed by scaling the 'background' regions (WM and CSF) to an estimate of their mean values, convolving thems with the PSF and subtracting them from the observed PET image. This produces the recovery coefficients and the GM voxels are corrected using the same step as Meltzer's method. The approach is known as the Müller-Gärtner (MG) method. The three compartment correction is given by the equation:

$$f_{GM}(x) = \frac{f_O(x) - [C_{WM} p_{WM}(x)] \otimes h(x) - [C_{CSF} p_{CSF}(x)] \otimes h(x)}{p_{GM}(x) \otimes h(x)}. \tag{2.9}$$

where f_O is the observed image, f_{GM} is uptake in the grey matter, p_{GM}, p_{WM} and p_{CSF} are masks (where $p(x) \in [0,1]$) of the GM, WM and CSF space, C_{WM} is the estimated WM mean value, C_{CSF} is the estimated CSF mean value (usually assumed to be zero) and h is the PSF of the scanner. The masks p_{GM}, p_{WM} and p_{CSF} are typically defined as binary masks, although it is be possible to use probabilstic masks which could account for PVEs in the MRI data. When using probablistic masks each voxel must sum to less than or equal to 1 over the set of masks. C_{WM} is typically estimated from an eroded WM region thought to be relatively free of PVEs, such as the centrum semiovale. It is necessary to erode the WM to ensure that the C_{WM} is estimate is not biased due to spill-over from other regions.

The MG correction produces a voxel-wise correction in GM voxels only. The images can be difficult to interpret as it consists of the cortical ribbon, which is at most a few voxels thick. For visual assessment, MG-corrected images are somtimes overlaid onto the registered MRI data to provide the viewer with familiar anatomical references. Alternatively, a 'false' WM region can be added to the image, by applying the C_{WM} value at every voxel in the p_{WM}

mask.

2.3.1.4 Geometric Transfer Matrix

Rousset et al. [1998a] proposed an MRI-based method that performs a volume-based correction which can correct for multiple regions. The Geometric Transfer Matrix (GTM) technique requires that the MRI is parcellated (see section 2.4.2, page 41) into a set of non-overlapping ROIs. Each ROI is assumed to contain uniform activity which can be described accurately by the mean value within it. The mean value of each region is found and held in a vector $\mathbf{t}$ of length N, where N is the number of ROIs. The fractional contribution of each region to its neighbours is then calculated. This is performed in the following way: first the binary mask of region p_i is convolved with the scanner PSF ($p_i' = p_i \otimes h$). The contribution of p_i into p_j is calculated by summing the voxel values in p_i' belonging to the masked region p_j. The sum is then divided by the total number of voxels in p_j to give the fractional contribution $\omega_{i,j}$. The contributions are held as an $N \times N$ matrix G, known as the Geometric Transfer Matrix (GTM):

$$G = \begin{pmatrix} \omega_{1,1} & \omega_{1,2} & \cdots & \omega_{1,N} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{N,1} & \omega_{N,2} & \cdots & \omega_{N,N} \end{pmatrix} \tag{2.10}$$

Each row (and each column) of the GTM should always sum to 1 in order to account for all activity from a given region. Each element on the diagonal represents the spill-out of each region, with each element being less than 1. Thus, the regional activity in the observed image can be described by:

$$\begin{pmatrix} t_1 \\ \vdots \\ t_N \end{pmatrix} = \begin{pmatrix} \omega_{1,1} & \omega_{1,2} & \cdots & \omega_{1,N} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{N,1} & \omega_{N,2} & \cdots & \omega_{N,N} \end{pmatrix} \begin{pmatrix} T_1 \\ \vdots \\ T_N \end{pmatrix} \tag{2.11}$$

where $\mathbf{T}$ is the vector of true mean values. The true regional mean values are then found by calculating:

$$\mathbf{T} = G^{-1}\mathbf{t}. \tag{2.12}$$

The GTM does not however produce an image and therefore is only useful when analysis focuses on regional mean values. The MG is arguably easier to implement as it is performed on a three-class segmentation rather than a more detailed parcellation required for GTM. Regional uniformity is an assumption made by the GTM. Therefore, regions should be defined

in such a way that this assumption can be considered appropriate. For example, different tissues are contained in separate regions. As the GTM can correct for regional spill-over rather than just between-tissue PVEs, including more regions in the GTM (ignoring registration and segmentation errors) should improve the accuracy of the correction.

In theory, each voxel could be considered as a region. The result of performing the GTM would produce the exact solution to the problem. This solution would be a PV-corrected image. However, as the GTM is of size $N \times N$, the matrix is too large to invert and would be ill-conditioned. Segobin et al. [2010] proposed a method that uses GTM as part of a RL deconvolution and reported that the matrix became ill-conditioned when too many regions were included.

The GTM, as originally described by Rousset et al. does not account for the spatial variation in scanner resolution. Du et al. [2005] proposed the perturbated GTM in order to apply the correction to SPECT data using OSEM. This was performed by adding small perturbations to the ROIs and then projecting and reconstructing both the original and perturbed image, while accounting for the spatial resolution. The difference of the reconstructed images will represent the desired distribution.

2.3.1.5 Modified Müller-Gärtner

A variation on the MG method has been proposed by Rousset et al. [1998b] which performs the calculation of the WM mean value using an alternative approach. First a GTM correction is performed on a set of regions. The set will normally consists of $3 \leq N \ll M$ regions, where N is the number of regions and M is the total number of voxels. The value found for the WM region is used as the mean value for the WM compartment during a MG correction. This correction method is known as the modified Müller-Gärtner (mMG) [Quarantelli et al., 2004] due to the alternative way the WM value is found. Note that while the GTM can correct for multiple regions, the mean values are unused when performing the voxel-based correction.

2.3.1.6 Multi-target correction

MG (and mMG) performs a voxel-based correction, but is unable to correct for multiple regions. Whereas, GTM corrects for multiple regions, although cannot generate an image. Erlandsson et al. [2006] reported a hybrid PVC technique referred to as Multi-target correction (MTC). The method first calculates PV-corrected regional mean values using the GTM approach. These mean values are then used to perform a voxel-based correction for

multiple regions. The voxel-based correction step is given by the following equation:

$$f_C(x) \approx \sum_{j=1}^{N} p_j(x) \frac{f_O(x) - \sum_{i \neq j} \mathbf{T}_i p_i(x) \otimes h(x)}{p_j(x) \otimes h(x)}. \tag{2.13}$$

where f_C is the corrected image, f_O is the observed image, p_i and p_j are the masks of regions i and j respectively, $\mathbf{T}_i$ is the PV-corrected mean value of region i calculated by equation 2.12 and h is the scanner PSF.

As MTC uses the GTM to obtain the mean values, the assumption of regional uniformity is still made. Like the GTM, MTC requires parcellated MRI data and the parcellation should reasonably satisfy the assumption of uniformity. However, the key difference is that MTC produces an image. The results of MTC can be analysed in the image domain, providing the ability to calculate image statistics that would otherwise by unavailable when performing the GTM alone. A more general extension of this method is described in chapter 3.

2.3.1.7 Wavelet-based correction

An anatomy-based PVC technique that uses wavelets was originally proposed by Boussion et al. [2006]. The method uses a wavelet transform to perform *multi-resolution analysis*. This involves the incorporation of high-resolution anatomical information into the lower resolution PET image. The original proposal operated in 2D, although this was extended to 3D [Le Pogam et al., 2008] in order to account for the 3D nature of PVEs. When there is a correlation between the anatomical image and the structural image, the regions will be PV-corrected. However, in situations where there is no correlation, for example, a defect that cannot be seen on MRI but can with PET, the new method will not introduce artefacts, although at the same time PVC will not occur either. This would result in an image where some regions may be PV-corrected, while others may not be.

The wavelet method has been further modified to incorporate an atlas [Shidahara et al., 2009] to segment the anatomy. The results suggest that the introduction of the atlas improved quantification. The issue of a lack of correlation between modalities has not been addressed and in the situation where structural information is not reflected in the functional data, the authors state that artefacts will occur in the recovered image. They suggest however that this situation is unlikely in brain imaging.

2.3.2 Reconstruction-based PVC methods

An alternative approach to PVC is to perform the correction inside the reconstruction as opposed to a post-reconstruction step. This can be advantageous when modeling scanner

resolution as this can be more complicated in the image domain. In addition, the noise in the projection domain is known to be Poisson distributed. This is not the case in the image domain. The remainder of this section describes some approaches to reconstruction-based PVC.

2.3.2.1 Resolution recovery

Resolution Recovery (RR) is the process of including knowledge of the scanner response function into the system matrix used for reconstruction [Comtat et al., 2008, Reader et al., 1998]. It is possible to model the complete image acquisition process during the reconstruction and to include corrections for image degradation caused by effects such as: resolution [De Bernardi et al., 2007], scatter, attenuation and patient motion [Carson et al., 2003]. Correction for the PSF can be achieved by modifying the projector. The forward projection is computed and then the projections are convolved with the scanner PSF which is typically approximated by a Gaussian.

Comtat et al. [2008] modelled the PSF as a space-invariant Gaussian. The results demonstrated that modeling the resolution reduced the convergence rate of OSEM and increases activity in small structures. The authors also stated that over-estimation of the PSF induces Gibbs artefacts and that an under-estimated PSF is preferable. While incorporating knowledge of the PSF into the reconstruction should eliminate PVEs, RR does not achieve perfect recovery [Reader et al., 2003].

2.3.2.2 Anatomical priors

Anatomical priors have been included in the reconstruction in order to suppress noise. Smoothing priors are commonly applied during iterative reconstruction for this purpose. These have the effect of reducing the variance of voxels in a neighbourhood. Segmented (or parcellated) anatomical data can be used to increase or decrease the smoothness constraint, based on the location of voxels with respect to anatomical regions. A typical approach is to relax the smoothness constraint near the borders of regions, while smoothing more throughout the interior. These approaches often lead to pleasing qualitative results.

Bowsher et al. [1996] proposed a method for reconstructing emission data while also segmenting it, using anatomical *a priori* information. This was achieved by using a Bayesian approach where prior probabilities were modified depending on whether a particular segmentation of a region was within boundaries observed in the anatomical segmentation. While this method requires MRI segmentation, Bowsher et al. [2004] presented an algorithm that was driven by the MRI intensity alone. The *Bowsher prior* encourages a larger degree of smoothing in neighbourhoods of voxels where the MRI intensities are similar, evaluating a *clique*

of MRI voxels when determining the smoothing constraint. The prior can produce visually impressive images. However, little evaluation of the quantitative accuracy, or whether the assumption that PET (or SPECT) activity relates to MRI intensities, has been carried out.

Bataille et al. [2007] applied a prior based on anatomical brain regions that were blurred to compenstate for PVEs. This approach was originally proposed for whole-body reconstruction [Comtat et al., 2002]. The technique was shown to perform similarly to the (post-reconstruction) GTM method. The blurring of the labels was also shown to reduce the effects of PET-MR mis-registration. The application of anatomical priors remains the focus of much research, as the results of the reconstructions can be very non-linear, potentially affecting quantitative accuracy. In addition, the reconstructed data often have spatially variant resolution, making the application of post-reconstruction PVC non-trivial.

2.3.2.3 PVC in the projection domain

A novel approach to PVC in SPECT was proposed by Erlandsson and Hutton [2010]. PVC was performed in the projection domain, using anatomical information. The authors perform PVC using an 'iterative' FBP, which was later modified for OSEM reconstruction [Erlandsson et al., 2011]. The algorithm calculates regional mean values based on the region definitions of parcellated MRI data. A PVC step similar to that of Yang et al. [1996] is performed in the projection space. The mean calculation, followed by PVC step is iterated until convergence. This corrected projection data is then used within the reconstruction.

By applying the correction to the projections, the spatially-varying resolution in SPECT can be accounted for. The OSEM version (OSEM-PVC) was compared to OSEM with RR. Both were found to significantly improve contrast compared to standard OSEM. OSEM-PVC was reported to have lower regional variability than with RR, due to the superior structural definition provided by the anatomical data.

2.4 Image processing

Image processing is often a necessary pre-processing step in medical analysis. This may be performed for the purposes of analysis or in order to apply corrections to an image. The term *image processing* refers to a broad field of research, however this section focuses on the topics of image registration and brain parcellation, as these are necessary to be able to perform the PVC techniques described in this thesis.

2.4.1 Registration

Image registration is the process of geometrically aligning two or more images of the same object measured at different times and/or using different instruments [Maintz and Viergever,

1998, Zitova, 2003]. In terms of medical imaging, registration is applied to combine data sources in order to improve the available information about a patient. Medical image registration is either intra-modality or inter-modality. Examples where intra-modality registration is necessary are: the alignment of dynamic PET frames to account for patient motion [Andersson and Thurfjell, 1997], the comparison of pre- and post-operative images [Ferrant et al., 2002] and monitoring disease progression [Scahill and Fox, 2007, Villemagne and Rowe, 2010].

Inter-modality registration is the task of aligning images from different imaging modalities, for example PET and CT or PET and MRI. This may be performed for the purposes of correcting one modality using the information of the other, such as CT-based AC [Kinahan et al., 1998] or MRI-based PVC [Rousset et al., 2007] in emission tomography. Registration can also be used to assess functional images in terms of their anatomy, improving the interpretation of the functional data. This interpretation may either be visual, such as image fused PET-CT or provide quantitative measures such as functional tumour volume in radiotherapy planning.

The parameters that align one image to another by image registration is called a *transformation*. There are three types of transformation that are commonly applied in medical image analysis: rigid, affine and non-rigid. The appropriate type of transformation to use depends on what is being registered. Each transformation is briefly described below:

2.4.1.1 Rigid transformation

A rigid transformation allows the target image (the one being transformed) to be rotated and translated only. This can be described using homogeneous coordinates [Foley et al., 1996, Maintz and Viergever, 1998]:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \left[\begin{array}{ccc|c} & R & & \begin{matrix} t_x \\ t_y \\ t_z \end{matrix} \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \tag{2.14}$$

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix},$$

$$R_y = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix},$$

$$R_z = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

where x,y and z represent the 3D coordinates of the point to be transformed, t_x, t_y and t_z represent the translation to be applied each axis, R_x, R_y and R_z are rotation matrices around the x-,y- and z-axis respectively and α, β and γ are the angles of rotation about the x-,y- and z-axis respectively.

A rigid transformation is appropriate for applications where images of the same subject are co-registered together, such as PET-PET frame registration.

2.4.1.2 Affine transformation

An affine transformation permits rotation and translation of the target image as in the rigid case, but additionally the image may be scaled and/or sheared. Scaling is applied to a point in the following way:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{2.15}$$

where s_x, s_y and s_z are the scaling factors to be applied in the x,y and z directions. Shearing is typically applied when the some distortion is known to exist in one of the images, for example geometric distortion in MRI images due to non-linear gradient fields [Caramanos et al., 2010]. In matrix form, shearing can be applied in the following way:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & k_{xy} & k_{xz} & 0 \\ k_{yx} & 1 & k_{yz} & 0 \\ k_{zx} & k_{zy} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{2.16}$$

where k_{xy} is the shear due to x along the y direction and k_{xz} is the shear due to x along the z direction, and so forth.

In the case of brain imaging, affine transformations can be used to perform inter-modality registrations, such as PET to MR. It should be noted however that re-sampling the target image to the same voxel dimensions as the reference image, and then performing a rigid registration, will often suffice, assuming there is no distortion.

2.4.1.3 Non-rigid transformation

Both rigid and affine transformations are global as they are applied to the complete image space. A non-rigid (also referred to as elastic) transformation allows local warping of the target image in order to align it with the reference. Non-rigid transforms are non-linear and therefore cannot be represented by a matrix. There are several ways that non-rigid transformations have been parameterised, including B-splines [Modat et al., 2010, Rueckert et al., 2006], statistical shape models [Cootes et al., 2004], and discrete cosine transformations [Ashburner and Friston, 1999]. The result of a non-rigid transformation is that the target image is deformed to align with the reference image, while (normally) preserving topology. An example use of non-rigid transformations is to warp subject images to a common template space for the purposes of inter-subject comparison. For a review of the most widely used non-rigid registration algorithms see Klein et al. [2009].

2.4.1.4 Optimisation schemes and cost functions

In order to accurately register two images, the parameters of the transformation, whether rigid, affine or non-rigid, have to be determined. An optimisation scheme is required to search for the parameters. These schemes typically operate on the result of a metric (or cost function) based on the similarity of the transformed target image and the reference image. Common optimisation techniques such as Powell [Powell, 1964] or Simplex [Dantzig, 1998] can be applied to find these parameters. Often multi-resolution pyramids and/or multi-scale methods are used in conjunction with an optimiser to reduce the search space and therefore speed up convergence, while avoiding local minima [Maintz and Viergever, 1998].

The choice of cost function depends upon application. For intra-modality registration, the sum of squared differences (SSD) or cross-correlation are often used [Zitova, 2003]. Both these methods assume that the intensities in the target and reference images are in some way correlated to one another. This assumption is not valid for inter-modality registration. For example, the measured activity in a PET image may bear no relation to the MRI intensity values. In this instance, cost functions based on statistical dependencies in the two images are often applied. Mutual Information (MI), Normalised Mutual Information (NMI) and Cross-correlation (CC) are commonly used for PET-MR and PET-CT registrations [Studholme et al., 1999, Viola and Wells III, 1997].

2.4.1.5 Registration software used in this work

All PET-MR registrations in the work were performed using the software package Nifty Reg (University College London, UK) [Modat et al., 2010]. The software is open source and

can be downloaded from: `http://sourceforge.net/projects/niftyreg/`. The package includes the applications `reg_aladin` and `reg_f3d`.

`reg_aladin` performs rigid and/or affine registrations using the intensity-based block matching technique of Ourselin et al. [2001]. The algorithm has two stages. First, a displacement field is computed by matching blocks between the target and reference image. The block matching stage uses normalised CC to match corresponding blocks, given by the following equation:

$$\mathrm{NCC}(x,y) = \frac{1}{n-1} \sum_x \sum_y \frac{(I(x,y)-\bar{I})(J(x,y)-\bar{J})}{\sigma_I \sigma_J} \tag{2.17}$$

where $I(x,y)$ is a block in the target image, $J(x,y)$ is a block in the reference image, $\bar{I}$ and $\bar{J}$ are the means and σ_I and σ_J are the standard deviation (SD) of target and reference block respectively.

Second, a trimmed least squares (TLS) regression is used to find the transformation that provides the best fit to the displacement field. This process of block matching, followed by TLS regression is then iterated until convergence. In addition, a multi-resolution approach is also applied. The registration is first performed on down-sampled images that have been Gaussian smoothed. The degree of smoothing is reduced as the algorithm iterates until the final registration is performed using the original image resolution. For a full description of the algorithm, see Ourselin et al. [2001, 2002].

`reg_f3d` is an accelerated version of the non-rigid B-spline registration technique proposed by Rueckert et al. [2006]. Non-rigid registration (NRR) is not used during the processing or analysis of data in this work, although a description of the algorithm is given for completeness. An alternative method of analysis could have been applied using NRR to a place subject scans in a common template space. However, this was not applied as the key aim of this work is improving individual subject quantification rather than group analyses or classification tasks.

The algorithm operates by moving a lattice of control points defined over the target image. Displacing control points changes the mapping between the target image and the reference image. The deformation field is represented as cubic B-splines [Rueckert et al., 2006]. A cost function comprised of NMI and a penalty term, the bending energy, is used to evaluate the accuracy of the mapping [Modat et al., 2010].

The MI between two images I and J is calculated by [Zitova, 2003]:

$$\begin{aligned} MI(I,J) &= H(J) - H(J|I) \\ &= H(I) + H(J) - H(I,J) \\ &= H(I,J) - H(I|J) - H(J|I) \end{aligned} \tag{2.18}$$

where $H(I)$ and $H(J)$ are the *marginal* entropies of I and J respectively. The marginal entropy can be thought of as the uncertainty associated with a given image. $H(I|J)$ and $H(J|I)$ are the *conditional* entropy, that is, a measure of what J does not explain about I, and what I does not explain about J, respectively. $H(I,J)$ is then the *joint* entropy, or the uncertainty of both I and J. NMI can then be defined as:

$$\text{NMI} = \frac{H(I) + H(J)}{H(I,J)} \tag{2.19}$$

A higher NMI value implies that there is a greater degree of shared information between the two variables. A conjugate gradient ascent optimisation scheme is applied to find the optimal mapping between the target and reference. When the scheme has converged, the result is a non-linear registration of the target to the reference image.

2.4.2 Anatomical parcellation

The anatomy-based PVC techniques described in this chapter require structural information derived from MRI data. Parcellation is the process of dividing an image into a set of relevent labels that describe the image in some way. In this context, the term 'parcellation' is used to refer to the division of the brain into a set of anatomical labels, including the sub-division of GM into multiple regions. Whereas, 'segmentation' refers to the labelling of GM, WM and CSF as distinct regions only.

Parcellation of the MRI is a necessary step for the purposes of both PVC and analysis. Manual parcellation is time-consuming, can suffer from intra- and inter-observer variability and can result in parcellations that are more consistent in one axis than the others [Fischl et al., 2002]. Parcellation can be achieved by registering the subject MRI to an atlas. Typically, an atlas consists of a structural image (usually MRI) and an associated (template) image of anatomical labels, both in the same coordinate system. Registration to the atlas space is normally performed between the subject scan and the structural image of the atlas. When scans from different modalities exist for the same subject, they are usually co-registered to each other before a single subject to atlas registration is performed. The registration is then

propagated to the images of the other modalities. Figure 2.4 depicts how a PET image could be analysed in a template space using rigid registration and NRR.

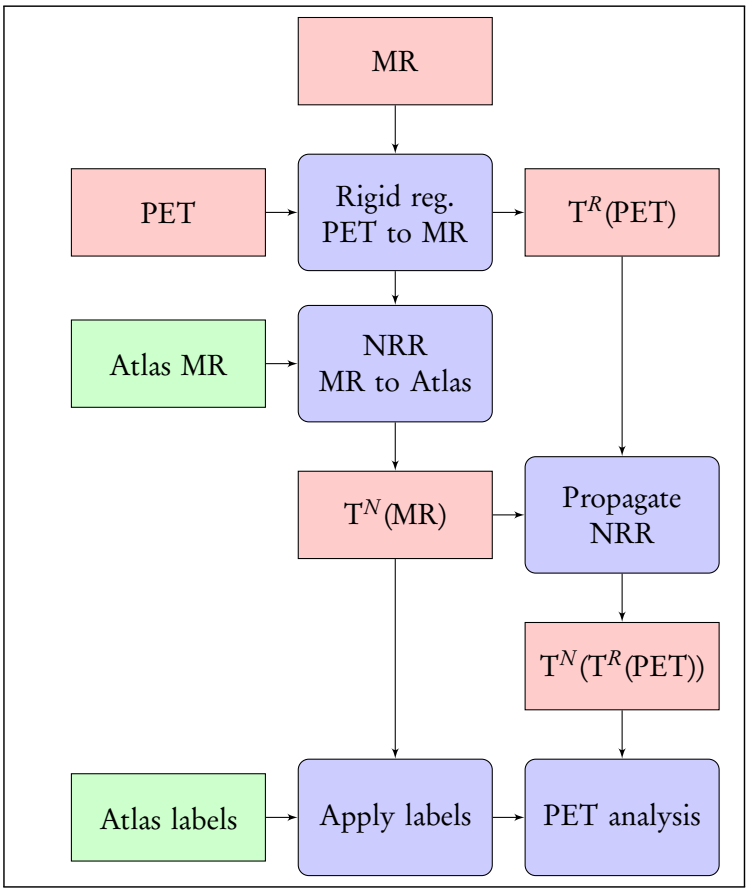


Figure 2.4: An example of how a subject PET image could be analysed in a template space through MR-MR registration.

The atlas may either be based on the manual parcellation of an individual [Talairach and Tournoux, 1988] or a probabilistic atlas derived from the anatomy of many subjects [Evans et al., 1993, Mazziotta et al., 1995]. The advantage of individual atlases such as the Talairach-Tournoux is that detailed anatomical labeling can be applied. However, anatomical variability of subjects, particularly in those with some neurological abnormality [Thompson et al., 1997], will result in parcellation errors. Whereas probabilistic atlases, being the product of multiple subjects, reflect the most common features of the population, while losing the finer anatomical detail.

More recently, multi-atlas approaches have been proposed for brain parcellation [Aljabar et al., 2009, Fischl et al., 2002, Klein et al., 2005]. Voxels are classified depending on the labels of multiple atlas images. The aim of this approach is to maintain the anatomical detail in the individual atlas image, while reducing errors due to anatomical variability.

Klein et al. [2005] proposed an algorithm called *Mindboggle* which parcellates a MRI

image according to the labels of (up to) 19 individual atlases. The voxels are classified by a majority vote system, selecting a label at random when votes are tied. The authors reported that the accuracy of parcellation significantly improved when multiple atlases were used.

Another multi-atlas approach was proposed by Aljabar et al. [2009] (see also Heckemann et al. [2010]). A database of 275 manually segmented brains was created. The parcellation of a new image was performed by first registering the MRI to the atlas space. The most 'similar' images from the database were then used to classify the voxels of the new image. Two similarity measures were investigated by Aljabar et al. [2009]: an image-based metric and age.

The image-based similarity measure used NMI. While NMI is often used for inter-modality registrations, the authors felt that NMI is appropriate for comparing MRI images from different subjects as contrast and intensity would vary across the cohort. 20 of the 275 atlases with the highest NMI values were used for the subsequent parcellation of the MRI image. The accuracy of the parcellation when using atlases based on NMI was found to be significantly better than random selection.

Selection based on age was also investigated. 20 of 224 atlases (age was unavailable for 51 subjects) with the closest age to that of the subject undergoing parcellation, were chosen. Performance was similar to that of the image-based selection, although slightly inferior. The authors report that the largest improvements in parcellation, when using age selection, were seen in older subjects. This may be due to atrophy effects altering the appearance of the imaging data.

2.4.2.1 FreeSurfer

The PVC techniques described in this work require parcellated MRI data. The data are also required to perform regional analysis of the PET data. The FreeSurfer (FS) package (Martinos Center for Biomedical Imaging, Harvard, MA, USA) [Fischl et al., 2002, 2004] is a suite of automated software tools for analysing and visualising MRI data. FS is currently considered state-of-the-art in terms of automated cortical parcellation [Desikan et al., 2006]. Although FS is used throughout this work, it should be noted that the PVC methods described do not rely on any one particular parcellation algorithm. As FS is used in all of the subsequent experimental chapters, a brief description of the parcellation process is given below:

Pre-processing: First an affine registration is applied to place the MRI in Talairach space. Initial volume labels are then assigned. An intensity correction is applied to the registered MRI. A NRR between the intensity corrected image and Talairach atlas is then performed.

Parcellation: The parcellation labels are modeled by a first order anisotropic non-stationary Markov Random Field (MRF), created from manually segmented images. Information about the spatial relationship between labels can be expressed as a function of their location in addition to their local anisotrophy [Fischl et al., 2004]. The parcellation procedure operates on both a probabilistic atlas defined by the MRF, which is independent of the subject, and intensity values measured from the registered subject MRI.

At each point, the probability of a given label is calculated as the probability of that class occurring in the atlas, multiplied by the likelihood of observing the measured MRI intensity in that class. An initial segmentation is computed by assigning every point to the class with the highest probability. A *neighbourhood function* is then calculated using the segmentation, which re-estimates the probability for each point depending on the classification of neighbouring points. The segmentation is recomputed using the new set of probabilities. This process is repeated until the segmentation can no longer change. The complete derivation of this process is given in Fischl et al. [2004].

The registration process performed on the subject MRI is invertible, therefore the parcellation can be propagated back to the original subject MRI space. The result is a detailed, subject-specific parcellation. There are several hundred labeled regions in the parcellated image. These are combined into a set of larger regions for the purposes of PVC and analysis. The parcellation itself takes approximately 30 minutes, although it should be noted that the complete FS processing pipeline takes roughly 24 hours (Intel Q8300 2.5 GHz, 4GB RAM) per subject.

2.5 Alzheimer's disease

2.5.1 Overview

AD is the most common form of dementia, characterised by progressive cognitive and functional impairment, ultimately resulting in death. As the average age of the global population increases, the prevalence of dementia is expected to continue rising, with an estimated 81 million sufferers by 2040 [Sloane et al., 2002]. Early detection of AD, before clinical features are apparent, is highly desirable. A patient with a reduction in cognitive ability, but not yet meeting the criteria for AD, is often described as having Mild Cognitive Impairment (MCI). This MCI state is believed by some to be the pre-clinical stage of AD. This has yet to be conclusively proved, however, a large proportion of MCI subjects convert to AD within one year [Drzezga et al., 2003], making the MCI population particularly interesting. The ability to identify pre-clinical AD, also referred to as the prodromal phase [Jagust, 2008], is highly

desirable from both a patient management perspective and for clinical trials assessing the efficacy of disease modifying drugs such as anti-amyloid therapies [Aisen, 2002].

Definitive diagnosis of AD can currently only be made with the histopathologically confirmed presence of amyloid-β plaques and neurofibrillary tangles (NFTs). For neuroimaging clinical trials, probable AD patients are typically screened by a neuropsychological battery. The Mini Mental State Examination (MMSE) is often performed to give a measure of cognitive function. In addition, the Clinical Dementia Rating (CDR) assessment can quantify dementia severity. However, there are difficulties distinguishing pre-clinical AD from normal ageing using neuropyschological testing alone. Thus, the imaging modalities of MRI, functional Magnetic Resonance Imaging (fMRI) and emission tomography are being investigated as potential solutions.

The ability to diagnose AD *in vivo* using non-invasive imaging modalities is currently the focus of much research. MRI investigations have centered on cerebral atrophy. Analysis of CSF changes has been shown to be potential biomarker [Shaw et al., 2009]. Volumetric measurement of specific brain structures implicated in AD pathology, such as the hippocampus and entorhinal cortex [Barnes et al., 2004, Fox and Schott, 2004, Frisoni et al., 1999, Jack, 1997, Juottonen et al., 1998, van de Pol et al., 2007], can discriminate between AD patients and normal controls. Volumetric information has also been used to identify patterns of atrophy in MCI subjects that predict conversion to probable AD [McEvoy et al., 2009].

PET and SPECT have also been investigated for potential AD biomarkers. One of the most extensively studied PET tracers is [^{18}F]FDG [Foster et al., 2008, Ishii et al., 2001, Mosconi et al., 2007, Yakushev et al., 2009]. A [^{18}F]FDG biomarker is highly desirable because it is the most widely available PET ligand. Tracers that can image amyloid-β plaque aggregation *in vivo* are also now available.

2.5.2 Atrophy in AD

Cortical atrophy is a feature of normal ageing [Fox and Schott, 2004]. GM and WM losses have been found to correlate with age [Resnick et al., 2003]. Resnick et al. found atrophy rates of 2.4 ± 0.4 and 3.1 ± 0.4 cm^3 per year for GM and WM volumes, respectively. However, AD patients are known to exhibit increased global rates of atrophy when compared to healthy control (HC) subjects [Thompson et al., 2003]. Atrophic changes associated with AD are believed to occur early [Frisoni et al., 1999] and before clinical features are present. MRI studies have shown that atrophy can be identified in the medial-temporal lobe, in particular the hippocampus and entorhinal cortex [Frisoni et al., 1999, Jack, 1997, Jack et al., 2005]. It is interesting to note that there is little [^{11}C]PIB uptake in these regions [Li et al., 2008], which

may be due to the existence of NFTs rather than plaques and/or could be caused by PVEs due to atrophy.

Atrophy in AD is thought to begin in the temporal and limbic cortices, spreading into the frontal and finally the occipital lobe [Singh et al., 2006]. The sensorimotor cortices tend to be spared during the progression of atrophy. In addition, the course of cortical atrophy in both healthy ageing and AD has been found to be asymmetric in the two hemispheres of the brain [Frisoni et al., 2007, Thompson et al., 2003], with atrophy occurring first in the left hemisphere. [^{18}F]FDG scans of AD patients have indicated a similar pattern, with hypometabolism in the temporal and parietal lobes and posterior cingulate, with subsequent deficits in frontal areas. The sparing of the sensorimotor corticies has also been observed with amyloid PET [Villemagne et al., 2011a].

As previously discussed, the degree of PVE is related to size of the structure. Atrophy causes a loss of tissue and therefore reduces the spatial extent of brain structures, particularly in cortical GM. The PVEs due to atrophy are likely to be more severe in AD subjects than controls. This suggests that PVC should be applied to brain PET images when investigating neurodegenerative disorders that have an atrophic component.

2.5.3 Amyloid

Alois Alzheimer first discovered the presence of amyloid plaques in 1907. It is these amyloid plaques that are believed to define AD. The amyloid cascade hypothesis [Hardy and Higgins, 1992] suggests that the deposition of amyloid-β peptides, in the form of plaques, is the prime event in AD pathogenesis. Plaque aggregation is believed to then be followed by synaptic dysfunction, neurodegeneration, cognitive and functional decline and ultimately death. Pathologically, AD is identified by extracellular plaques and intracellular tangles formed of phosphorylated tau protein, called NFTs.

The amyloid cascade hypothesis is the most widely accepted model of AD pathogenesis and while much research supports this hypothesis, the model may be a simplification of the complicated relationships between ageing, amyloid-β aggregation and tau deposition [Weiner et al., 2010]. High amyloid burden has been observed with [^{11}C]PIB in cognitively normal elderly subjects [Klunk et al., 2004] and it remains unclear as to what this means. While plaque burden may imply that a subject is at risk of developing AD, burden itself does not necessarily imply memory impairment [Aizenstein et al., 2008]. While amyloid may (or may not) be the root cause of AD, imaging amyloid burden *in-vivo* could prove to be a useful diagnostic tool.

The accumulation of amyloid-β is know to begin before the clinical indicators of AD

are present [Braak and Braak, 1995, Långström, 2007]. The first amyloid PET tracer was [^{11}C]PIB [Klunk et al., 2004], developed by the University of Pittsburg. Klunk et al. modified an amyloid-binding histological dye called thioflavin-T to create the tracer. [^{11}C]PIB exhibits a high-binding affinity to amyloid-β plaques in GM regions and non-specific binding in WM. Quantification of amyloid burden is possible and this has been shown to discriminate HC from AD subjects [Boxer et al., 2007, Kemppainen et al., 2006, Pike et al., 2007, Rowe et al., 2007]. While [^{11}C]PIB appears to be a robust when separating HC and AD groups, MCI groups tend to exhibit a bi-modal distribution [Li et al., 2008]. Some MCI subjects exhibit high levels of amyloid burden when imaged with [^{11}C]PIB, while others appear normal.

More recently, fluorine-labelled amyloid tracers have been undergoing evaluation. General Electric (GE) are assessing [^{18}F]flutemetamol [Vandenberghe et al., 2010b], which is chemically very similar to [^{11}C]PIB, where the carbon has been replaced with fluorine. [^{18}F]florbetaben [Barthel et al., 2011] is currently being evaluated by Bayer and Avid Radiopharmaceuticals have developed [^{18}F]florbetapir [Clark et al., 2011]. As opposed to carbon-based tracers, fluorine-labelled ligands do not require an on-site cyclotron and are preferable as the tracers can then be used at centres other than specialised research PET facilities. This at least partly influenced the decision to perform amyloid PET scans in the Alzheimer's Disease Neuroimaging Initiative (ADNI) Grand Opportunities trial [Weiner et al., 2010] with [^{18}F]florbetapir. Amyloid PET imaging is considered an important area of research into the development of treatments to slow down or halt the progression of AD.

2.6 Emission tomography in AD

Emission tomography modalities have been extensively investigated in order to find suitable biomarkers for AD. In terms of PET, this section focuses on [^{18}F]FDG and amyloid tracer. The SPECT tracer, [^{99m}Tc]-HMPAO is also reviewed. [^{18}F]FDG is tracer used extensively in PET for many clinical applications including, but not exclusive to, neurological imaging. Whereas, amyloid agents are specifically designed to label amyloid-β plaques in the brain, implicated in AD pathogenesis.

2.6.1 Single Photon Emission Computed Tomography

It has been suggested that brain perfusion imaging with SPECT could differentiate between AD patients and controls [Jagust et al., 2001]. Studies with [^{99m}Tc]-HMPAO SPECT have identified hypoperfusion in the temporal and / or parietal lobes as a strong indicator of AD. Jagust et al. [2001] performed a [^{99m}Tc]-HMPAO study with clinically diagnosed AD pa-

tients and age-matched normal controls. The images were visually inspected by experts for hypoperfusion. The authors reported that SPECT imaging improved the sensitivity of a diagnosis of probable AD by 7-8%, to 92% and doubled the sensitivity when clinical diagnosis of 'possible AD' had been given.

The performance of SPECT when compared with other modalities appears to be lower [Messa et al., 1994]. Messa et al. compared hypoperfusion using [^{99m}Tc]-HMPAO with hypometabolism using [^{18}F]FDG PET, reporting that significant hypoperfusion and glucose hypometabolism occur in cases of suspected early AD. Abnormalities were found to be significant in the parietal and temporal regions with [^{99m}Tc]-HMPAO. Frontal, parietal and temporal lobe abnormalities were observed with [^{18}F]FDG. Messa et al. report that the detection rate of abnormalities with [^{18}F]FDG and [^{99m}Tc]-HMPAO were 100% and 90% respectively. However, it should be noted that a high-resolution SPECT scanner (8.4mm FWHM resolution) was used for the study, which may account for the high detection rate.

Imran et al. [1999] applied voxel-based analysis to [^{99m}Tc]-HMPAO images of AD and HC subjects. Cerebral blood flow was compared with psychiatric evaluation scores. Patients were diagnosed as probable AD according to the standard diagnostic criteria. Disease severity was then assessed using MMSE and CDR. Voxel-based analysis was carried out using Statistical Parametric Mapping (SPM). The authors found parieto-temporal abnormalities, consistent with those of Messa et al. [1994]. However, the authors also report that the abnormalities were larger on the left side. This could be due to greater volume losses in the left hemisphere [Thompson et al., 2003], as discussed earlier. The hypoperfusion of the left parietal region was found to correlate strongly with age, MMSE scores and CDR.

The spatial resolution limits the achievable results with SPECT. Brain SPECT is additionally limited by the tracer kinetics of the ligand used. In addition, SPECT images suffer from large PVEs due to the resolution. Although SPECT is a lower cost technology, most research in this field is now carried out using PET.

2.6.2 FDG PET

PET imaging with [^{18}F]FDG has been researched extensively [Foster et al., 2008, Ishii et al., 2001, Kono et al., 2007, Mosconi et al., 2007] in the search of AD biomarkers, [^{18}F]FDG PET describes cerebral glucose metabolism and has been shown to detect early changes in AD patients [Foster et al., 2008].

The degree of hypometabolism in a [^{18}F]FDG image has been shown to correlate with disease progression. However, the metabolic changes during the early stages of the disease can be difficult to discriminate from the normal effects of ageing [Ishii et al., 2006, Sakamoto

et al., 1997]. The metabolic change could also be either too small to detect and/or confounded by PVEs.

[^{18}F]FDG has been applied in the discrimination between dementias. Fronto-temporal dementia (FTD) and AD can be difficult to differentiate clinically [Foster et al., 2008] as both dementias follow a similar path of decline. The patterns of metabolism found in both dementias with [^{18}F]FDG are quite different. In AD, the posterior temporoparietal cortex and posterior cingulate are affected, whereas in FTD, it is the frontal lobes, anterior temporal and anterior cingulate that display hypometabolism [Foster et al., 2007].

Drzezga et al. [2003] performed a longitudinal study using [^{18}F]FDG in AD, MCI and HC groups. 36% of subjects diagnosed as MCI at baseline converted to AD within one year. Statistical analysis of the baseline scans of MCI converters found significant differences in the temporoparietal and posterior cingulate regions when compared to controls. Hypometabolism in the posterior cingulate of non-converting MCI subjects also tended to significance compared to controls. Observations of further prefrontal hypometabolism were made when assessing the follow-up scans of converters. This study is particularly interesting as the authors claim that it is possible to delineate the abnormal regions at baseline, predicting which subjects will convert to AD.

2.6.3 Amyloid PET

With the advent of new PET ligands that can image amyloid-β plaques *in-vivo*, research in this field has now focused on amyloid PET imaging. [^{11}C]PIB was developed by the University of Pittsburg [Klunk et al., 2004]. It is currently the most extensively studied radioligand for the imaging of amyloid-β plaques in AD patients. [^{11}C]PIB has been shown to be both highly sensitive and specific for AD diagnosis [Price et al., 2005, Rowe et al., 2007], as opposed to [^{18}F]FDG which exhibits high sensitivity but lower specificity [Edison et al., 2007].

Studies using [^{11}C]PIB report that binding is robust in the vast majority of AD subjects [Kemppainen et al., 2006, Klunk et al., 2004, Rowe et al., 2007]. [^{11}C]PIB uptake is high in the prefrontal and lateral temporoparietal cortex, posterior cingulate and precuneus [Klunk et al., 2004, Rowe et al., 2008]. Low uptake can be seen in the medial-temporal lobe and sensorimotor cortex [Rowe et al., 2008]. However, Rowe et al. [2007] found no correlation between the [^{11}C]PIB uptake and the onset of cognitive decline or MMSE scores. AD subjects tend to exhibit increased [^{11}C]PIB binding regardless of disease severity [Rowe et al., 2007].

In terms of multimodality imaging systems, Archer et al. [2006] and Jack et al. [2009] have both shown that there is a correlation between cerebral atrophy and [^{11}C]PIB uptake, making a diagnostic system using both [^{11}C]PIB and MRI plausable. Also, Price et al. [2005]

report that [^{11}C]PIB uptake inversely correlates with [^{18}F]FDG glucose metabolism in areas such as the parietal cortex.

PVC is not regularly applied, despite the likelihood of large PVEs being present in diseased subjects. Bourgeat et al. [2010] did perform PVC as part of a study into the relationship of amyloid burden and hippocampal atrophy. They applied a mMG correction and found a significant relationship between [^{11}C]PIB uptake in the inferior temporal lobe and hippocampal volume. The authors do however perform the same analysis on the data without PVC and state that the results are broadly similar.

Mormino et al. [2008] applied Meltzer's method for PVC. Strong correlations between episodic memory deficits and [^{11}C]PIB uptake were found. However, when normalised for intercranial volume, [^{11}C]PIB uptake was found to no longer be correlated with memory deficits. The authors suggest that it is amyloid-β deposition that induced volume loss in the hippocampus.

Fluorine-labelled amyloid tracers are now being assessed. Rowe et al. [2008] have investigated the use of [^{18}F]florbetaben, which binds to plaques with an uptake distribution almost identical to that of [^{11}C]PIB. The authors report that the dynamic range of [^{18}F]florbetaben is lower than [^{11}C]PIB. However, even with reduced range, the discrimination of AD subjects from controls and from FTD subjects was acheived with an accuracy similar to that of [^{11}C]PIB.

Clark et al. [2011] performed a study with [^{18}F]florbetapir in 29 subjects (*in-vivo*) for which post-mortem amyloid-β immunohistochemistry was later available. The authors report 96% agreement between the visual interpretation of the [^{18}F]florbetapir images and immunohistochemistry results. This study shows that the compound is clearly binding to amyloid-β. However, the diagnostic and prognostic value, as with all amyloid tracers, remains unclear.

[^{18}F]flutemetamol was compared to [^{11}C]PIB by Vandenberghe et al. [2010b]. The main purpose of the study was to evaluate the efficacy of blinded visual reading with [^{18}F]flutemetamol. High sensitivity and specificity (93.1% and 93.3% respectively) was observed when classifying AD and HC subjects as either amyloid-positive or amyloid-negative. The secondary purpose of the investigation was to evaluate the agreement between [^{18}F]flutemetamol and [^{11}C]PIB for a sub-set of the cohort (AD: $n = 20$; MCI: $n = 20$) who had undergone PET scans with both tracers.

Linear regression analysis found strong agreement between the two tracers in cortical regions in both subject groups. The authors state that the observed activity in the WM was

higher in [^{18}F]flutemetamol images than [^{11}C]PIB. This is believed to be due to different washout characteristics of the two ligands. No PVC was applied in the study, although the differences in WM uptake could potentially affect results due to varying degrees of spill-over. It should be noted that [^{18}F]flutemetamol, [^{18}F]florbetaben and [^{18}F]florbetapir all exhibit higher WM signal than [^{11}C]PIB.

2.6.3.1 Longitudinal assessment

Amyloid tracers have been shown to discriminate between AD and control subjects with high sensitivity and specificity. As previously discussed, amyloid deposition occurs before AD manifests clinically. Longitudinal studies have been reported which show increases in tracer uptake over time for AD and MCI groups [Villemagne et al., 2011b]. Resnick et al. [2010] also observed greater longitudinal decline in subjects with high amyloid-β deposition. Villemagne et al. [2011b] performed mMG PVC and noted 30% increases in GM signal after correction. However, they also report that the longitudinal percentage increases in uptake were similar to that of the uncorrected data. The authors do suggest that while PVC did not change the conclusions of the study, it may be necessary in extended longitudinal studies.

2.6.4 Quantitaive analysis of amyloid PET

This section explains how quantitative analysis of amyloid PET data has been performed in the literature. The two most common approaches are graphical analysis and the calcluation of ratio images. Reference is made to [^{11}C]PIB, although the approaches are also applicable to the fluorine-labelled amyloid tracers.

2.6.4.1 Graphical analysis

The binding of [^{11}C]PIB to fibrillar amyloid-β deposits in GM is believed to be specific and reversible [Mintun et al., 2006]. Graphical analysis is a technique that is often used in PET quantification. These methods typically use either arterial blood sampling or an image-derived input function. As arterial sampling is invasive, an image-derived input function is desirable. Reversible uptake can be analysed through Logan plot analysis [Logan et al., 1996].

The cerebellar GM is a commonly selected reference region for amyloid PET analysis [Li et al., 2008]. It is chosen because the cerebellar GM exhibits little fibrillar amyloid plaque aggregation [Klunk et al., 2004], even in the late stages of disease, although diffuse deposits have been observed in *in vitro* studies [Braak and Braak, 1991, Joachim et al., 1989, Larner, 1997].

The Logan plot is a graphical analysis tool used to derive the Distribution Volume Ratio (DVR) based on the integral transformation of time-activity curves (TACs) of the target and

reference regions [Logan, 2000]. Price et al. [2005] validated Logan analysis for [^{11}C]PIB PET by comparing the results of graphical analysis to compartmental modeling using an artertial input function. The ratio was calculated from 35-90 minutes post injection. This was found to give stable results in ROIs and produce robust parametric images.

2.6.4.2 Ratio images

An alternative approach to the quantification of [^{11}C]PIB PET images is to calculate a ratio image based on the uptake in a reference region, using late scan frames. The amyloid tracer signal is thought to be stable from around 40 minutes post-injection. These ratio images are calculated by dividing the image on a voxel-basis by the mean (or median) uptake in the reference region. The images are referred to as standardised uptake value ratio (SUVR) images. This approach has been regularly applied to [^{11}C]PIB data [Bourgeat et al., 2010, Fripp et al., 2008, Kemppainen et al., 2006, Lopresti et al., 2005, Mormino et al., 2008].

The advantage of this approach over the Logan analysis is that the scan time is relatively short (20 - 30 minutes), as opposed to a dynamic 90 minute acquisition required for Logan analysis. There is some debate about the ideal duration and start time. For example, Kemppainen et al. [2006] used a 60 to 90 minute scan, Bourgeat et al. [2010] used a 40 to 70 minute protocol and Mormino et al. [2008] chose a 50 to 70 minute acquisition. Similar scan times have been used for [^{18}F]flutemetamol quantification [Vandenberghe et al., 2010b]. The ratio image approach is used throughout this work as it is straight-forward to apply and dynamic data was not available for the multi-centre clinical datasets.

2.7 Discussion

PET images suffer from a degrading phenomenon known as the partial volume effect. This chapter has introduced PVEs and associated PVC techniques in the context of PET imaging. Various approaches to PVC exist. Data-driven techniques are desirable as anatomical information is not always available outside of a research context. The performance of iterative deconvolution methods tends to be inferior to those based on anatomy. The limiting factor with deconvolution-type approaches appears to be noise amplification, and noise suppression techniques continue to be investigated.

The anatomical data from high resolution modalities such as MRI has been utilised for the purposes of PVC. Quantitative accuracy of brain PET can be significantly improved by the application of anatomy-based PVC methods. They do however possess some potential disadvantages. The PVC techniques rely on registration and segmentation methods. Errors in either or both will degrade the PVC performance. In addition, most methods make as-

sumptions about the regional uniformity of PET uptake. This assumption may be invalid for some tracers and depends on how the regions are defined. Errors in the quantification of small structures due to PVEs are known to be large in brain PET [Soret et al., 2007]. PVC is a step that is often overlooked and rarely applied in clinical practice. Improvements in registration, segmentation and image reconstruction make PVC feasible.

PET tracers that can image amyloid-β plaques *in-vivo* may provide useful biomarkers for detecting and monitoring AD. PVEs may obfuscate these biomarkers, especially in subjects with severe tissue loss. Atrophy, along with amyloid-β deposition, is a hallmark of AD. Reduction in the size of a structure will result in more severe PVE. Longitudinal studies in particular could be affected as increasing amyloid deposition could be attenuated by the effects of atrophy. This thesis reports investigations into PVC techniques to improve quantitative accuracy for single scans and longitudinal studies.

Chapter 3

Evaluating partial volume correction techniques using phantom data

This chapter describes the evaluation of PVC strategies in amyloid brain PET using digital phantom data. Section 3.1 describes a comparative study of four PVC techniques. The performance of three anatomy-based techniques and a data-driven approach, the VC deconvolution, are assessed. Section 3.2 describes a further investigation into improvements to the VC deconvolution through the use of filters and priors.

3.1 Assessing the performance of partial volume correction techniques

PET suffers from PVEs which degrade the quantitative accuracy of the images. PVC techniques have been proposed that aim to reduce the influence of PVEs and thereby improve quantitative accuracy. This study evaluates the performance of four PVC approaches: the MG method, an extension to MG known as mMG, region-based voxel-wise (RBV) correction and the VC iterative deconvolution.

The MG approach [Müller-Gärtner et al., 1992] is an anatomy-based technique that performs a voxel-wise correction using MRI data (see section 2.3, page 29 for a full description). The MG method assumes regional uniformity in WM and CSF compartments, correcting the GM for PVEs only. Rousset et al. [1998b] proposed an alternative method for estimating the WM mean value used during the MG correction that uses the GTM [Rousset et al., 1998a]. This alternative is referred to as the mMG. The third anatomy-based method that is evaluated in this study is RBV correction [Thomas et al., 2011]. It performs a voxel-wise correction for multiple regions, based on parcellated MRI data. In addition to the anatomy-based methods, the VC deconvolution is a data-driven approach that has recently been proposed for use in PET imaging [Teo et al., 2007, Tohka and Reilhac, 2008]. The remainder of this section

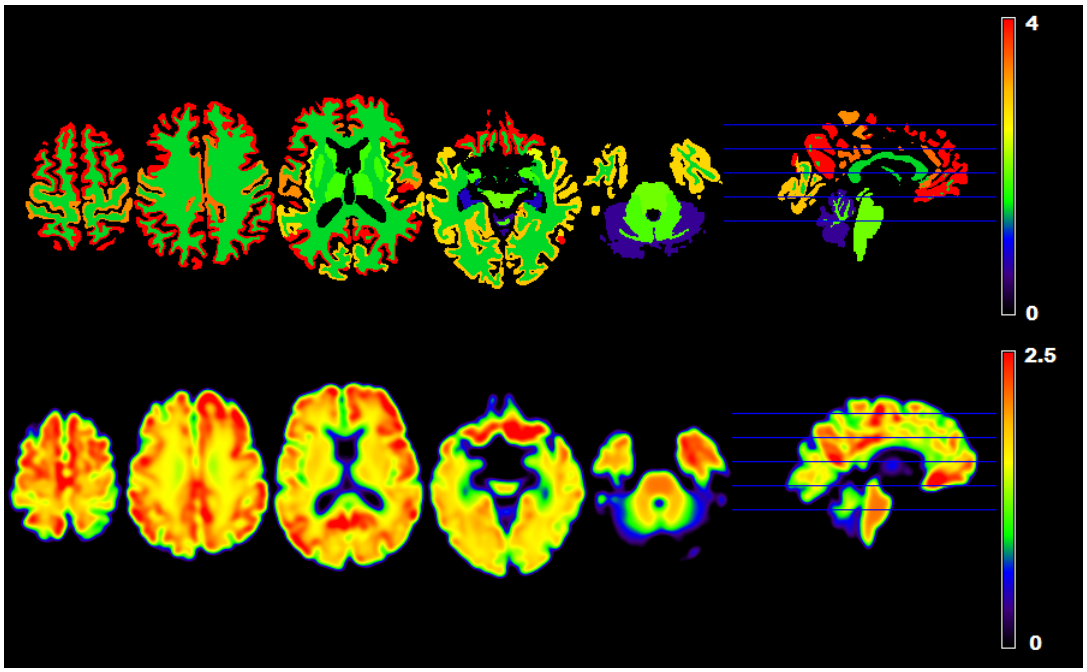


Figure 3.1: Example of a simulated [^{18}F]flutemetamol SUVR image of the AD phantom. Images were normalised using the cerebellar GM and are shown on a scale of 0 to 4 for the ground truth (top row) and 0 to 2.5 after the application of a 6.7 mm Gaussian PSF (bottom row).

investigates the performance of these PVC techniques using simulated amyloid PET images.

3.1.1 Materials and methods

3.1.1.1 Phantom images

Digital phantom datasets were created to evaluate the performance of the PVC techniques. The images were generated using the parcellated MRI from an AD subject and a HC subject. Ten AD-like distributions and ten control-like distributions were generated using PV-corrected SUVR from subject scans. The ten AD-like distributions were applied to the parcellated regions of the AD subject MRI (referred to as the 'AD phantom'), with the ten normal distributions applied to the control subject MRI (subsequently referred to as the 'HC phantom'). The images were then convolved with a Gaussian PSF of 6.7 mm FWHM. An example of one of the AD phantom images can be seen in figure 3.1.

3.1.1.2 Region-based voxel-wise correction

The MTC technique (section 2.3.1.6, page 33) was modified, using a similar approach to that of Yang et al. [1996]. This method is referred to as RBV [Thomas et al., 2011]. As with MTC, RBV utilises parcellated MRI data to perform a voxel-wise correction for multiple regions. Regional mean values (**T**) are calculated using the GTM (section 2.3.1.4, page 32), which are then used for the voxel-wise correction. The correction step in equation 2.13 (page

34) is instead given by:

$$f_C(x) \approx f_O(x)\left[\frac{s(x)}{s(x)\otimes h(x)}\right], \qquad (3.1)$$
$$s(x) = \sum_{i=1..N}[\mathbf{T}_i p_i(x)].$$

where s is a piece-wise constant image of the regional mean values found by the GTM. The ratio of s to the convolution of s with the PSF ($s \otimes h$) are the correction factors required to correct the observed PET image f_O. When f_O is multiplied by the correction factors, the result is a PV-corrected image almost identical to that produced by MTC.

The modification to the correction step was proposed for two reasons. First, it is more straight-forward to implement than MTC as the convolution of each region mask by the PSF, calculated during the GTM step, does not need to be stored for reuse in the voxel-wise correction. Second, the RBV formulation is less susceptible to large edge effects that could be caused by noise in the image. In the noiseless case, RBV and MTC produce identical results. However, when noise (other than the spatial resolution blurring) exists in the observed PET, it is possible that at a given voxel the denominator in the correction step could be very small on the region edge. This situation could result in a very large corrected voxel value. As MTC corrects on a region-by-region basis, this would be more likely to occur.

RBV is evaluated in this section and then used throughout this work as it provides the ability to correct for multiple compartments, while producing an image for both visual and quantitative assessment.

3.1.1.3 Partial volume correction

Four approaches to PVC were evaluated: VC, MG, mMG and RBV correction, where MG and mMG are variations on the same technique. The VC approach is a data-driven technique that operates on the PET data alone. The three other techniques require anatomical information typically provided by segmented MRI data. RBV correction was performed with the same regions as those used to create the phantom images. Binary masks of the GM, WM and CSF were generated from the parcellated MRI. These segmentations were produced for the purposes of MG and mMG correction. The four PVC methods were applied assuming a resolution of 6.7 mm, matching the resolution of the phantom images. In addition, the sensitivity of RBV correction to errors in the measure of the PSF was carried out as a separate analysis. The AD phantom was corrected with RBV correction using 4.7 - 8.7 mm at increments of 0.2mm. The frontal, parietal, cingulate and sub-cortical white matter (SWM) were assessed in terms of mean RC. These regions were chosen as the GM regions tend to exhibit high uptake

in AD subjects, with the frontal and parietal regions exhibiting severe atrophy effects in the AD phantom image. SWM was also assessed as this region would be likely to demonstrate under- and over-correction for GM spill-over effects.

3.1.1.4 Analysis

Regional analysis was performed on the PV-corrected and -uncorrected images. Regions were evaluated in terms of their mean values. The regional mean values were recorded using a 10% trimmed mean. The trimmed mean is found by removing the highest and lowest 5% of voxel values and then recalculating the mean over the remaining elements. The ratio of a given regional mean to the ground truth (the true value in the phantom) is called the recovery coefficient (RC). A RC of 1 indicates perfect recovery. A value of greater than 1 is observed in over-corrected regions, with values of less than 1 in under-corrected regions. The coefficient of variation (CoV) was calculated for each ROI as the standard deviation divided by the trimmed mean. The inter-subject CoV is denoted as CoV_s. All images were analysed using FSL v4.1 (FMRIB, Oxford University, UK).

3.1.2 Results

Reductions were observed in all cortical GM regions of both phantoms. These reduction are caused by PVEs. The AD phantom exhibited more severe reductions in cortical regions than the HC phantom due to atrophy. The RC in cortical regions of the HC phantom was >0.75, compared to the AD phantom which exhibited RC as low as 0.52.

Errors reduced when the four PVC techniques were applied. However, bias was observed in the hippocampus of the AD phantom. Both of the MG techniques induced a positive bias in the hippocampus. The MG method produced a bias of +11.7% and +8.1% with the mMG approach. The bias was a result of spill-over from neighbouring GM regions. The performance of the four PVC algorithms is shown in terms of mean RC across the group for the HC phantom (figure 3.2) and the AD phantom (figure 3.3).

The MG methods also induced a small positive bias in the cerebellar GM of the AD phantom. The same bias was not observed in the HC phantom. RBV correction produced the most accurate results with an absolute error of less that 0.5% in all GM regions, for both of the phantoms. The MG and mMG methods performed similarly with absolute errors in the HC phantom of 0.9% and 1.1%, respectively. When applied to the AD phantom, the MG approaches produced absolute errors of 3.0% (MG) and 4.4% (mMG). The VC deconvolution technique improved recovery compared to the uncorrected data, although performance was lower than that of the anatomy-based methods. For the AD phantom, an absolute error of

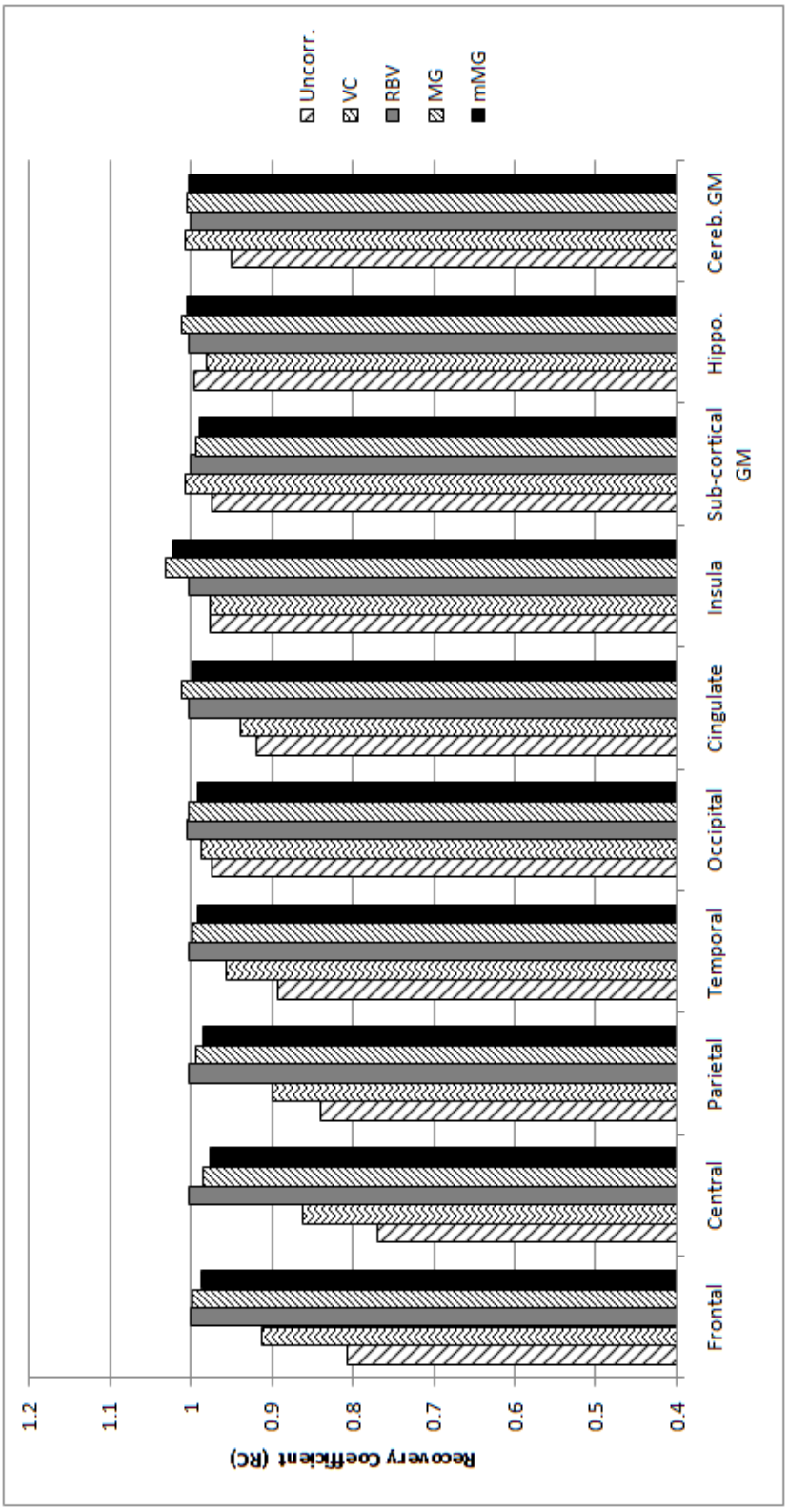


Figure 3.2: The effects of partial volume correction on recovery of grey matter regions of the healthy control phantom.

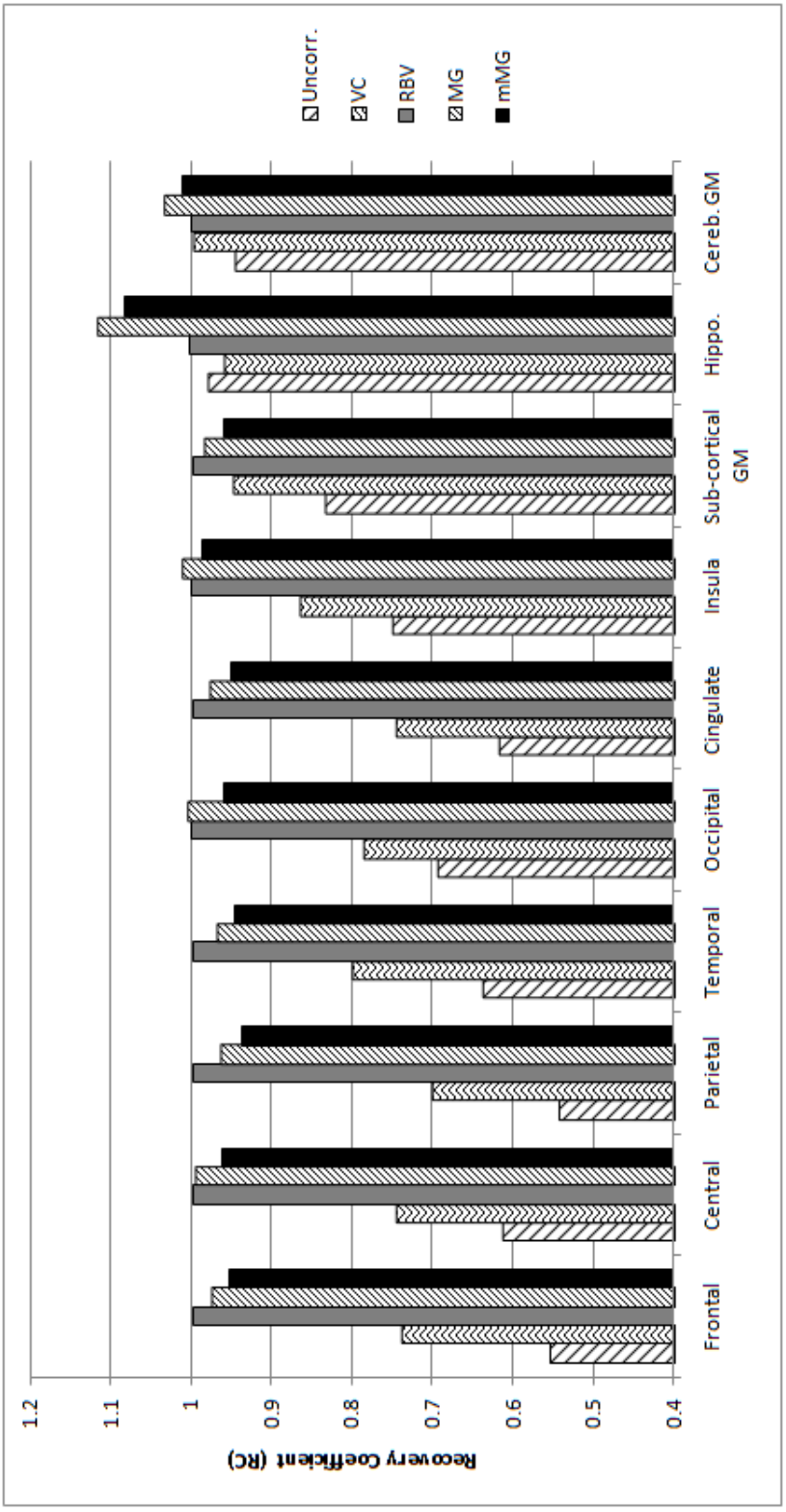


Figure 3.3: The effects of partial volume correction on recovery of grey matter regions of the Alzheimer's disease phantom.

22.5% remained after correction using the VC deconvolution.

The inter-subject variability is shown in figures 3.4 and 3.5. An increase in CoV_s was observed for all PVC techniques. When applying RBV correction the CoV_s was almost identical to that of the ground truth. This observation suggests that the phantom images are accurately corrected by RBV correction and that an increase in the CoV_s is not a PVC-induced error.

The results of the effects of PSF measurement errors on RBV correction can be seen in figure 3.6 in terms of FWHM against RC. Under-estimation of the PSF resulted in incomplete correction for PVEs. This can be seen in the over-estimation of SWM and under-estimation in the GM regions. When applying the correct PSF, RBV correction results in near perfect recovery in all regions. The cortical GM (frontal and parietal) were more sensitive to PSF errors than the cingulate. This is probably due to the thickness of cortical GM being less than or equal to the FWHM of the PSF and therefore suffering larger PVEs. At 6.2mm (0.5mm less than the true resolution), the RC is 0.95 for the cortical GM. This reduces to 0.92 at 5.7mm (1mm less).

3.1.3 Discussion

Digital phantom simulations of [^{18}F]flutemetamol PET data were used to evaluate the performance of PVC techniques. The simulations show that PVEs cause large quantification errors. These errors are particularly high in the AD phantom where the GM signal is attenuated due to the effects of atrophy.

The application of PVC resulted in a more accurate quantification of activity in cortical GM regions. In all GM regions other than sub-cortical GM, the CoV_s was lower in the blurred images compared to the ground truth. This apparent reduced variability is due to WM spill-over effects creating a smaller spread in the inter-subject variability.

The results show that RBV was the most accurate of the PVC methods, correcting for both atrophy and spill-over between multiple regions. The MG techniques rely on the assumption of regional uniformity. This assumption results in an inability to compensate for intra-regional PVEs.

Given that RBV appeared to out-perform the other PVC methods in idealised conditions (perfect registration, segmentation and PSF estimation), the sensitivity of RBV correction to errors in the measure of the PSF was carried out as a separate analysis. The spatial resolution increases by about 1mm, from 1cm to 10cm off-centre when estimated for the GE Discovery ST using the National Electrical Manufacturers Association (NEMA) NU 2-1994 protocol [Mawlawi et al., 2004]. Given that the cortex (assuming the brain is in the centre of the FOV) is probably between 6-8cm from centre, the assumption of a space invariant PSF for

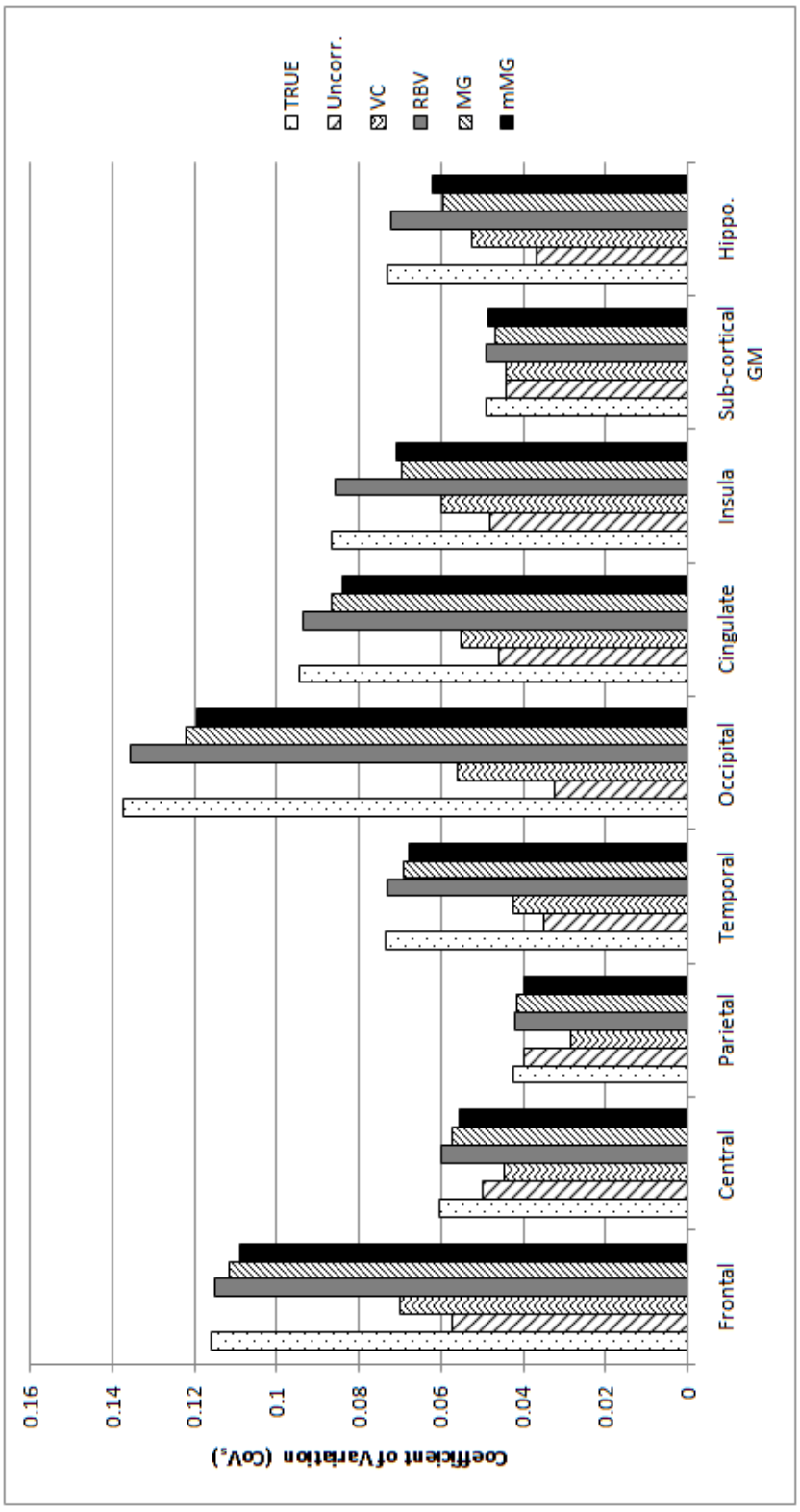


Figure 3.4: Inter-subject coefficient of variation in the grey matter regions of the healthy control phantom.

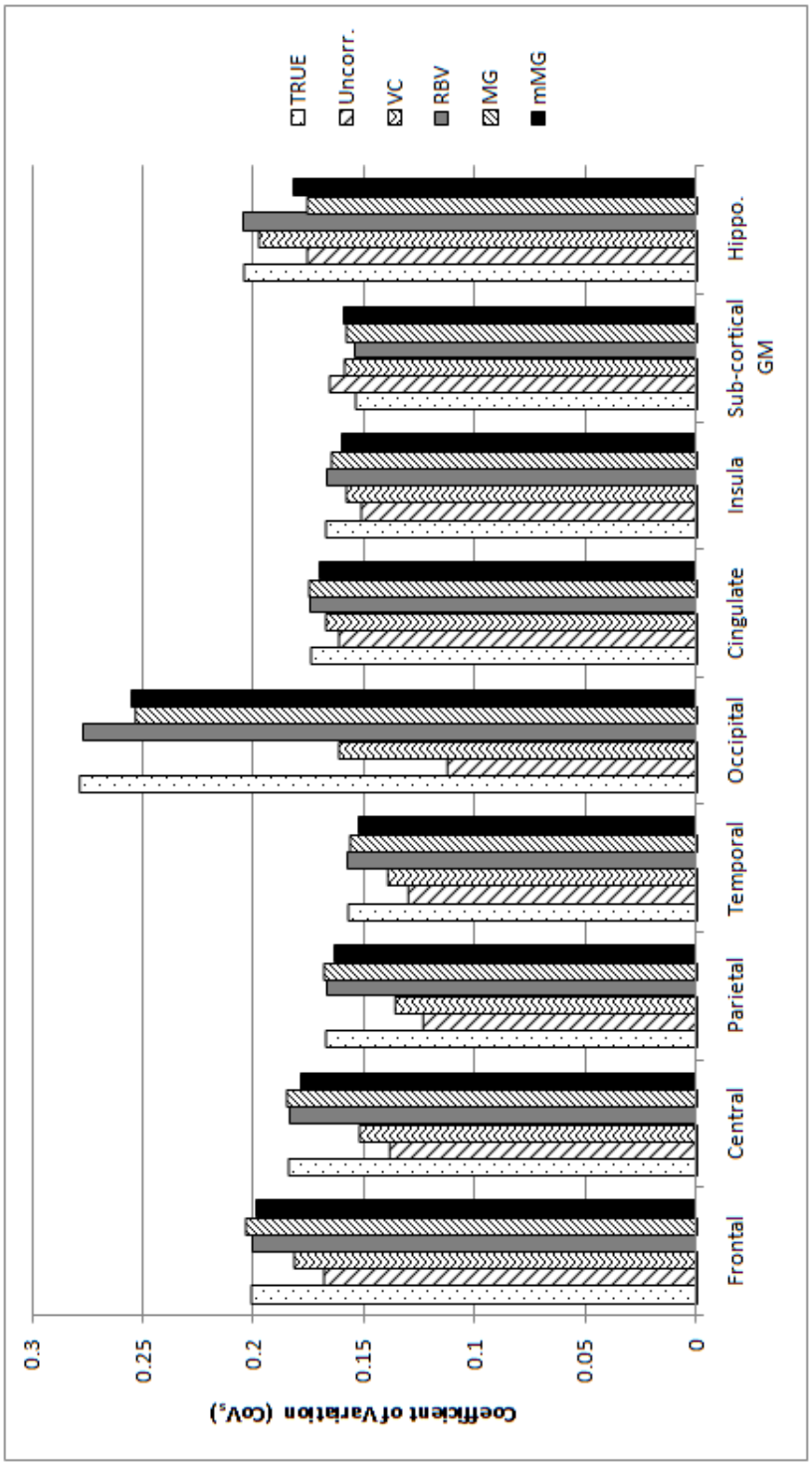


Figure 3.5: Inter-subject coefficient of variation in the grey matter regions of the Alzheimer's disease phantom.

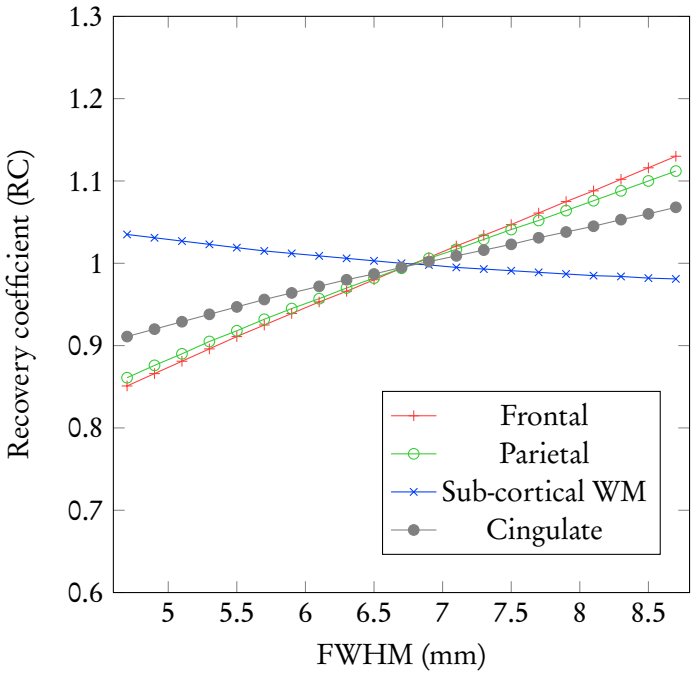


Figure 3.6: The effect of errors in the PSF measurement when performing RBV correction. The correct resolution is 6.7mm.

RBV correction probably results in a bias of no more than 5%. When compared to the -40% errors, without PVC, this assumption is acceptable.

The difference between the results of the MG and mMG methods are due to the estimated WM value used for the purposes of PVC. The mMG relies on GTM for the WM mean value estimation. The GTM in turn assumes regional uniformity. This assumption is valid for control subject and relatively small differences were observed between the two MG methods in the HC phantom. In AD subjects, cortical GM uptake can vary greatly throughout the cortex, with areas such as the occipital lobe and the sensorimotor cortex exhibiting visibly less uptake than regions such as the frontal lobe. This inhomogeneity causes a positive bias in the WM estimate produced by the GTM, leading to a larger negative bias in GM regions when applying the mMG rather than MG. The MG approach exhibits lower inter-subject variability which appears desirable. However, as this a phantom study, with a known ground truth, the lower variability was actually due to a smaller SD about an incorrect mean value ($RC \neq 1$).

The VC deconvolution approach improved recovery when compared to the uncorrected data, although the correction was not as accurate as the anatomy-based methods. This is likely to be due to the iterative process terminating prematurely. The early termination of the process resulted in incomplete recovery. The selection of a different stopping criterion

or alternative α value would produce difference results. The chosen parameters were based on the work of Tohka and Reilhac [2008]. However, as deconvolution-based approaches are known to greatly increase noise [Boussion et al., 2009], especially at high iteration numbers, realistic parameters were chosen for the VC approach.

The observed bias in the AD phantom hippocampus using the MG techniques can be attributed to intra-regional PVEs caused by spill-over from adjacent 'hot' GM regions, namely the amygdala and parahippocampal gyrus. The hippocampus tends to exhibit relatively low uptake even in AD subjects, whereas high uptake can be observed surrounding GM regions. The MG methods are unable to correct for spill-over between regions of the same tissue type, resulting in the observed hippocampal bias. In terms of recovery in the hippocampus, the mMG method appears to produce a lower error than the MG technique. The mMG actually over-corrects the surrounding WM, reducing the effect of the uncompensated GM-GM spill-over.

The digital phantoms were realistic in terms of both tracer distribution and anatomy. Uptake patterns observed in clinical data were used to generate the phantom datasets. The anatomical information was provided by the structural MRI of a healthy control and AD subject who were diagnosed by clinical evaluation. The phantoms were limited in the sense that other sources of error were not accounted for. No attenuation, scatter or randoms simulations were applied and no reconstruction or post-processing was included. In addition, no registration or segmentation errors were included. The aim of this phantom study was to evaluate the performance of PVC techniques in idealised conditions.

3.1.4 Conclusion

This phantom study has shown that PVC improves the quantitative accuracy of amyloid PET distributions. RBV correction achieved the most accurate recovery in all regions. The increase in inter-subject variability after PVC is genuine and should be expected when PVC is applied to clinical data. Both MG methods performed well for cortical GM regions. A positive bias was induced in the hippocampi of AD-like distributions when correcting with the MG and mMG. This was due to the methods' inability to account for spill-over from adjacent GM regions. The differences between the MG and mMG have been shown to be a result of the WM mean value estimates. The MG approach is more appropriate than mMG when a 3 compartment correction is applied to amyloid PET data. However, RBV correction out-performs both methods, although requiring a more complex segmentation to achieve this. The VC technique was inferior when compared to the anatomy-based methods.

3.2 Evaluating filters and priors for van Cittert deconvolution

This study describes investigations into the use of filters and priors as part of the VC deconvolution approach to PVC. Section 3.2.1 provides background information about deconvolution and the associated noise amplification issues. The motivation for evaluating these modifications is given in section 3.2.2. The approach used to generate appropriate phantom data is described in section 3.2.3. This is followed by a description of the methods applied (section 3.2.4). The remaining sections describe results (section 3.2.5) and provide a discussion (section 3.2.6) and conclusions (section 3.2.7) from the findings of this study.

3.2.1 Background

The VC technique is a data-driven image restoration technique proposed by van Cittert [1931]. It has been applied extensively in the field of astronomy and more recently has been proposed for PVC in PET imaging [Teo et al., 2007, Tohka and Reilhac, 2006, 2008]. The approach aims to iteratively deconvolve an image that has been degraded by a known PSF. However, deconvolution techniques are known to amplify noise as they iterate [Boussion et al., 2009]. The noise amplification occurs because deconvolution is an ill-posed problem [Tohka and Reilhac, 2008]. The remainder of this section describes why deconvolution problems are ill-posed:

3.2.1.1 The ill-posedness of deconvolution problems

During image formation, an object t undergoes a convolution operation (C_h), which is dependent on the PSF h and is a continuous operation:

$$t \xrightarrow{C_h} C_h t = t \otimes h, \tag{3.2}$$

The observed image i is then assumed to be:

$$i = t \otimes h. \tag{3.3}$$

t can be recovered from i by applying the inverse of the convolution operator:

$$t \xleftarrow{C_h^{-1}} i. \tag{3.4}$$

However, the measured image $\tilde{i}$ is actually i perturbed by noise. The similarity of $\tilde{i}$ to i does not imply that the inverse operators are similar:

$$\tilde{i} \simeq i \not\Rightarrow C_h^{-1}\tilde{i} \simeq C_h^{-1} i. \tag{3.5}$$

The inverse operators are discontinuous and this is why deconvolution is an ill-posed problem. The appropriate inverse operator for $\tilde{i}$ may be very different from that of i. This issue affects inverse problems in general [Natterer, 2001, pp. 85–101].

3.2.1.2 The reblurred van Cittert

As described in section 2.3.1.1 (page 29), the reblurred VC is a least squares approach to the problem of image restoration. The true object t is estimated by finding the minimiser of the least squares criterion:

$$\sum_x || \, i(x) - (t \otimes h)(x) \, ||^2 \tag{3.6}$$

A gradient decent approach to minimising 3.6 can be applied using the iterative rule:

$$\begin{aligned} t_{k+1}(x) &= t_k(x) + \alpha(h \otimes (i - t_k \otimes h))(x) \\ t_0(x) &= i(x). \end{aligned} \tag{3.7}$$

where $\alpha \in [0,2]$ is the step at each iteration, used to ensure convergence. Equation 3.7 can then be rewritten as:

$$\begin{aligned} t_{k+1}(x) &= t_k(x) + \alpha(\nabla\Psi(x)) \\ \Psi(x) &= (i - t_k \otimes h)(x) \\ t_0(x) &= i(x). \end{aligned} \tag{3.8}$$

The reblurred VC seeks the t_k that minimises the gradient, leading to the generalised inverse solution. In this case, the generalised inverse will always be discontinuous [Natterer, 2001, p.86]. Due to its iterative nature a termination criterion is usually employed. Tohka and Reilhac [2006] propose the following criterion:

$$\left\{ \frac{\sqrt{\sum_x (t_{k+1}(x) - t_k(x))^2}}{\sqrt{\sum_x i(x)^2}} \right\} < 0.01 \tag{3.9}$$

3.2.2 Motivation

A data-driven approach to PVC in PET is desirable from a clinical perspective as patient-specific structural information is not necessarily available outside of a research context. As described in the previous section, methods such as the VC deconvolution suffer from noise amplification issues. It has been suggested that the iterative process could be terminated prematurely, before the noise has degraded the image too much. However, premature termina-

tion tends to lead to incomplete recovery, as seen in the PVC performance study (section 3.1, page 54). The motivation behind this investigation is to evaluate whether the use of filters and priors can dampen the effects of noise amplification, allowing higher iteration numbers to be reached and therefore improved recovery achieved.

3.2.3 Digital phantom image generation

This section describes the generation of phantom PET images, used to assess the performance of PVC techniques throughout this work. Section 3.2.3.1 explains the motivation behind the image generator. How the image generator operates is described in section 3.2.3.2. A set of example images created with the image generator are shown in section 3.2.3.4. The image generator is summarised in section 3.2.3.5.

3.2.3.1 Motivation

PET images are degraded by number of factors as the image is formed. These include the effects and associated corrections for attenuation, scatter, the detection of random events, in addition to positron range, non-colinearity and the random nature of radioactive decay. The noise in the sinogram data generated by these factors propagates into the reconstructed image. Therefore, simply adding Poisson distributed noise to a digital phantom in the image domain is not realistic in terms of the observed noise in PET.

In order to evaluate the performance of PVC techniques with phantom data under realistic conditions, it is necessary to simulate noise. To achieve this it is necessary to generate data that are subjected to the same sources of noise which occur during PET image formation. The phantom image generator described in the remainder of this section includes the effects of attenuation, scatter and randoms, along with applying Poisson noise in the projection domain. Addition of simplified models and corrections for attenuation, scatter and randoms to simulate PET images results in a more realistic estimate of noise than merely convolving an image with a PSF in the image domain.

3.2.3.2 Methods and materials

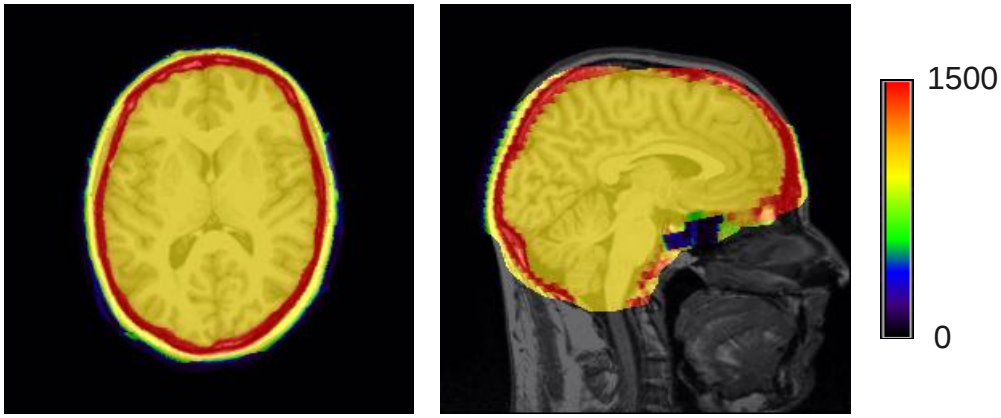
The generator requires a labeled anatomical image, a CT image of the same anatomy, a set of uptake values to be associated with the labeled image and a measure of the spatial resolution. Images are generated with a uniform resolution throughout the FOV, as the effects of the spatially varying PSF are relatively small across the cortex. Each of these inputs can be derived from real or phantom data. The phantom images used in this work (section 3.2 and chapter 5) were all created from actual MRI and CT data, with uptake distributions typical of that observed in clinical data. All images were stored in Neuroimaging Informatics Technology

Initiative (NIfTI) file format. The uptake distribution to be simulated, for a label image with N regions, is held in a comma separated variable (CSV) file with N rows; one row per region, with one value per row. The spatial resolution is measured in mm.

3.2.3.3 Data generation

The process for generating the phantom images from the given input data is shown in figure 3.19. Figure 3.20 focuses specifically on the noise process. Each stage has been assigned a label in the figures, subsequent references in the text relate directly to these labels.

1. **CT to MR registration:** The CT is registered to the T1-weighted MRI data using a rigid registration. The registration was performed using the `reg_aladin` application, which is part of the Nifty Reg package (see section 2.4.1.5, page 39). The result of the registration $T(CT)$ is then used for further processing. Figure 3.7 displays the CT image as an overlay on the MRI after registration has been performed.



(a) (b)

Figure 3.7: Axial (a) and sagittal (b) images of CT overlayed on T1-weighted MRI after rigid registration. Colour bar relates to the CT data.

2. **MR parcellation:** The MRI data are parcellated into regions that will then be used to represent the uptake distribution in the PET data. In the case of real MRI data, the FS package was used for parcellation. The detailed FS parcellation was then reduced to a smaller number of volumes using the region definitions provided in appendix D. An example of the parcellation and region reduction can be seen in figure 3.8

3. **Create PET distribution:** The activity distribution is created from the contents of the CSV file. The value at each row of the file is placed at every location defined by the mask of the associated region (figure 3.9). All regions have uniform activity.

4. **Convolve PET distribution:** The activity distribution from the previous step is then

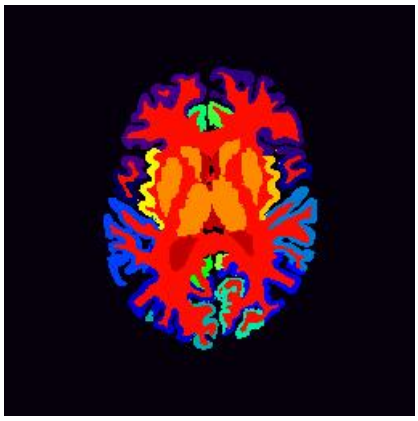


Figure 3.8: Parcellation of MR data using FreeSurfer, combined into a reduced number of regions.

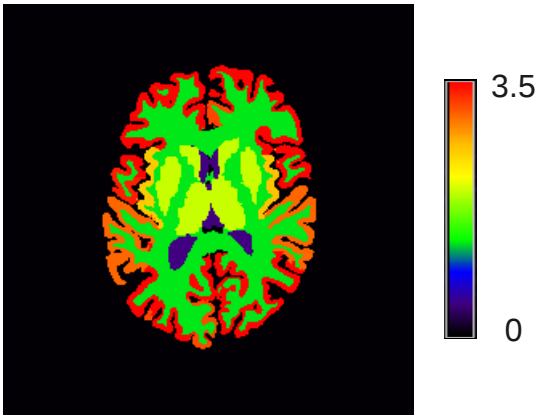


Figure 3.9: A simulated activity distribution before the noise model has been applied. Cerebellar GM normalised SUVR values are simulated and shown on a scale of 0 to 3.5.

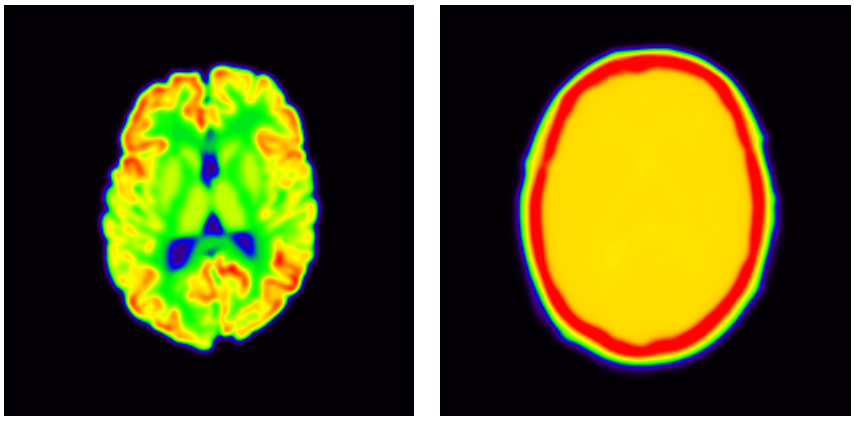
convolved by an isotropic Gaussian with a FWHM specified as the measure of the spatial resolution. The effect of the convolution on the activity distribution can be seen in figure 3.10a.

5. **Convolve CT image:** The registered CT image is also convolved by a Gaussian with the specified spatial resolution (figure 3.10b).

6. **Forward projection of the PET:** The convolved PET image is then forward projected. Each axial slice is projected over a set of angles using the radon transform. Figure 3.11 shows the result of forward projecting the PET slice shown in 3.10a.

7. **Forward projection of the CT:** The convolved CT image is also forward projected in the same manner as the PET.

8. **Application of noise:** Figure 3.20 depicts how the noise is applied to the sinogram data. The sinogram is first degraded by noise sources (8a - 8e). The effects of randoms (8f),



(a) (b)

Figure 3.10: The convolution of PET distribution (a) and CT (b) with an isotropic Gaussian (5mm FWHM).

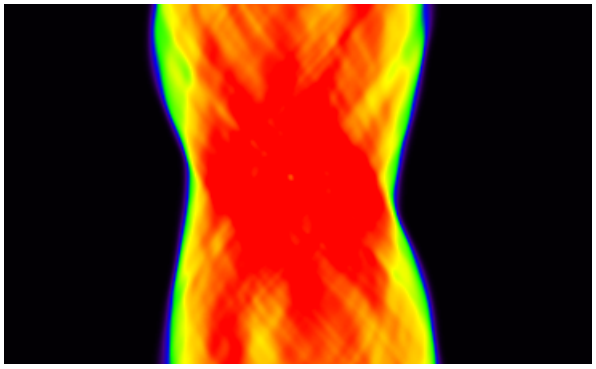


Figure 3.11: The forward projection of a convolved PET image slice.

scatter (8g) and attenuation (8h) are then removed. The process produces a sinogram that has in effect had attenuation, scatter and randoms correction applied.

(a) Adding attenuation:

A map of the attenuation is created from the CT data. To calculate the appropriate factors from the given CT, the following formula was used:

$$\rho(x)' = \exp\left[(\rho(x) - 1000) \times \frac{0.096}{1000} \times v\right] \tag{3.10}$$

where $\rho(x)$ is a voxel in the CT image and v is the voxel size in cm. The attenuation coefficient of water at 511 keV is $0.096\ \text{cm}^{-1}$ and this was the value assumed in the above equation. The subtraction of 1000 is applied to account for the CT data being stored as an unsigned integer. Equation 3.10 is used to create a map of attenuation. The PET projections are then divided on a voxel-wise basis by the attenuation map to apply the effect of attenuation (figure 3.12).

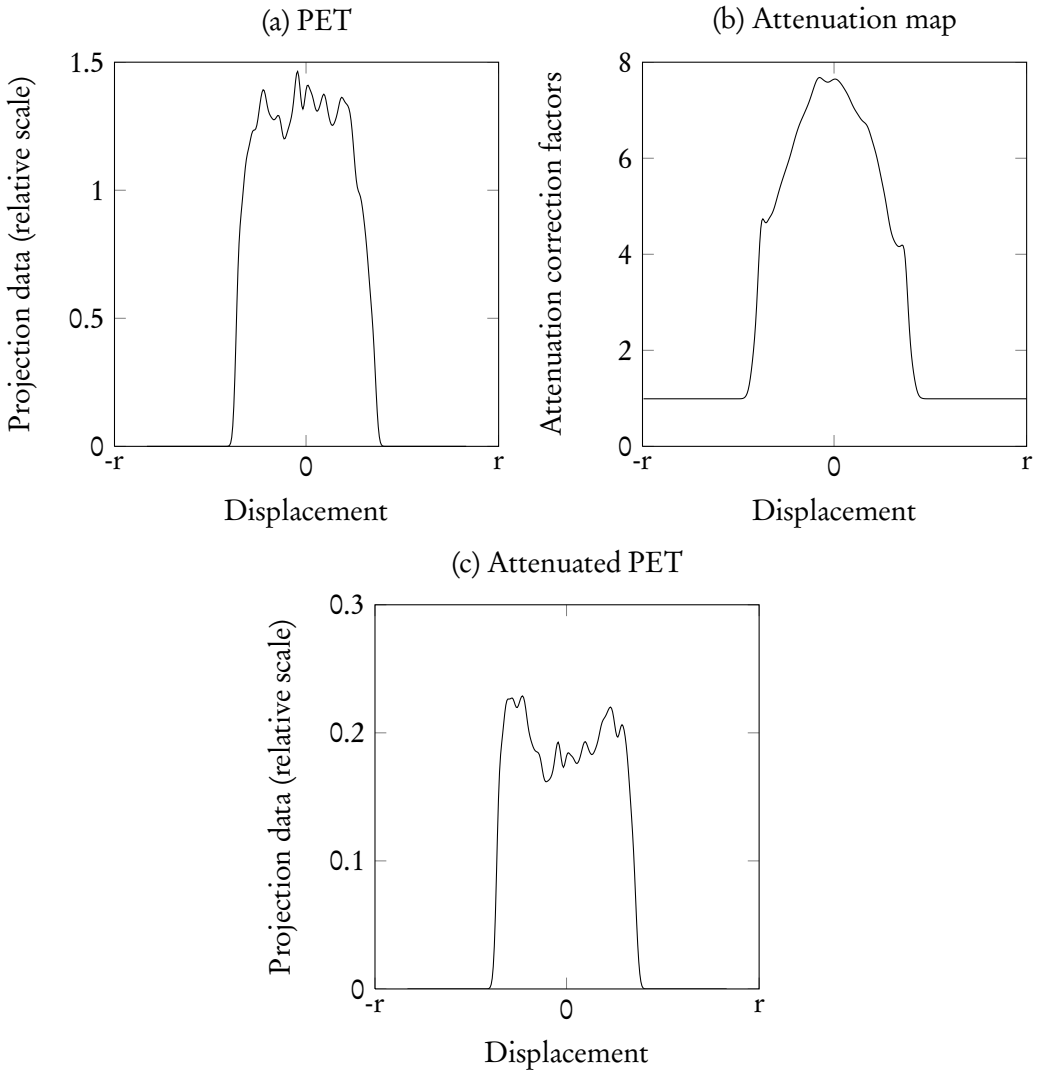


Figure 3.12: The effects of applying attenuation to the PET sinogram. A line profile through the radial axis of the PET sinogram in figure 3.11 is shown in (a). The associated line profile in the attenuation map can be seen in (b). The result of applying the attenuation map is shown in (c).

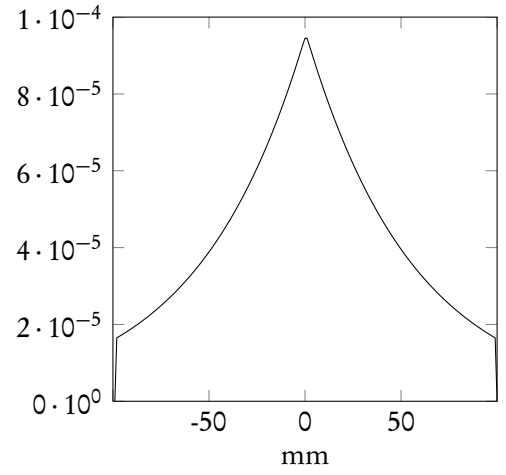


Figure 3.13: A line profile through the centre of the 2D scatter function.

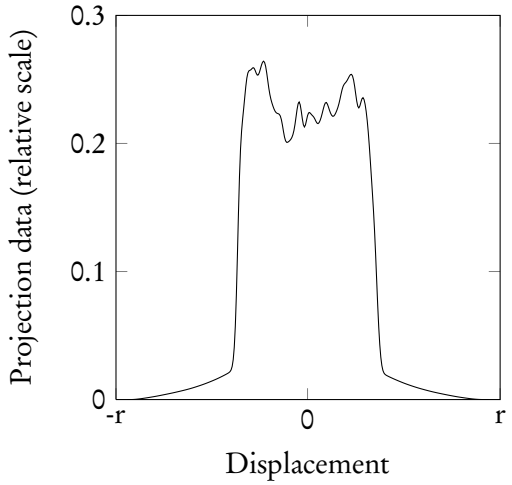


Figure 3.14: A line profile showing the effect of adding 25% scatter to the PET sinogram.

(b) Adding scatter:

A suitable scatter function was found experimentally, based on the measurements of Bergström et al. [1983]. The selected scatter function is monoexponential (figure 3.13). The function was defined as a 2D function. This was convolved with the 2D parallel projections. A fraction of the result of this convolution, in this case 25%, was added to the attenuated PET sinogram. The application of scatter in the same line profile can be seen in figure 3.14.

(c) Adding randoms:

To simulate the detection of random coincidences, a 5% random fraction was added to the PET sinogram. This was performed by calculating the sum of the sinogram and then dividing 5% of the sum by the total number of voxels, producing a 'randoms-per-voxel' value. This value was added to every location in the sinogram

(figure 3.15).

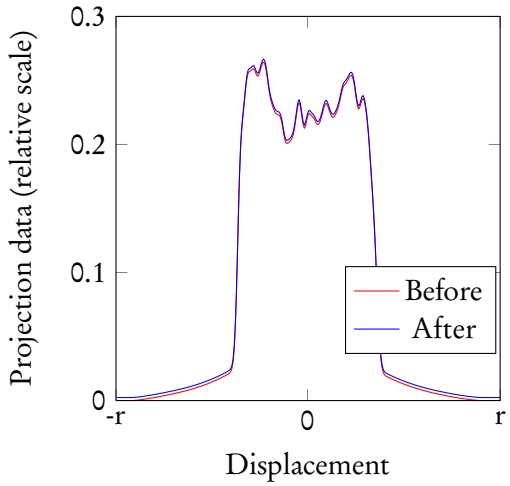


Figure 3.15: A line profile showing the effect of adding a 5% random fraction to the PET sinogram.

(d) Scale to count level

The PET sinogram is now re-scaled to the total count level. If the count level is not specified when the generator is called, a count level of $1e^8$ is assumed.

(e) Adding Poisson noise

After the image has been scaled, Poisson distributed noise is added to the sinogram data (figure 3.16). This was performed using the MATLAB function `poissrnd`, which is part of the Statistics Toolbox. It is important that the Poisson noise be added after scaling because, as discussed earlier, the amount of noise is proportional to the number of counts.

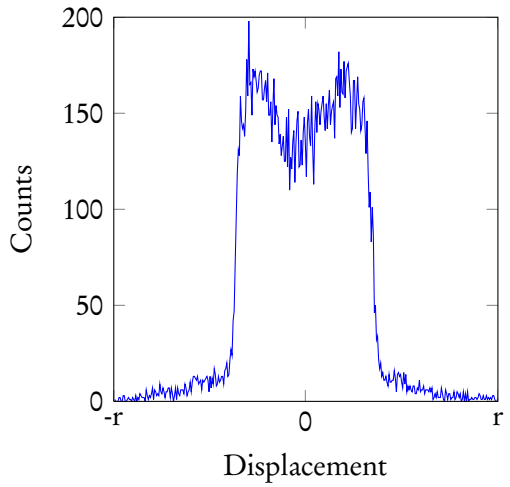


Figure 3.16: A line profile showing the effect of Poisson noise added to a PET sinogram. The image was scaled to $5.4e^8$ counts.

(f) Randoms correction

The random fraction is 'corrected' by subtracting the same 'randoms-per-voxel' value from the sinogram as found in stage 8c.

(g) Scatter correction

The scatter is also 'corrected' by removing the same scatter distribution that was applied to the sinogram in stage 8b.

(h) Attenuation correction

The sinogram is then multiplied by the attenuation map found in stage 8a. This has the effect of 'correcting' for the attenuation. Once all three corrections have been applied, the process of adding noise to the sinogram is complete (figure 3.17).

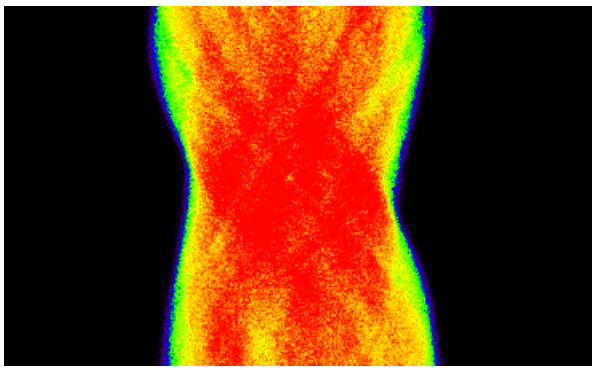


Figure 3.17: A single slice of the PET sinogram after the application of the model.

9. Applying a Hanning filter:

The Hanning filter (or window) is given by the equation:

$$w_n = \frac{1}{2}\left(1 - \cos\left[2\pi\frac{n}{N-1}\right]\right) \tag{3.11}$$

where n is the current point and N is the length of the filter. This is a low-pass filter, which removes high frequency noise. A 2D Hanning window was calculated, and this was applied to the parallel projections in the Fourier domain. The result applying the filter to the noisy PET data can be seen figure 3.18.

10. Back projection:

The filtered sinograms are finally reconstructed using FBP with a ramp filter. Figure 3.21 shows the back projection of the sinogram displayed in 3.18. All subsequent processing of the phantom data is then performed in the image domain.

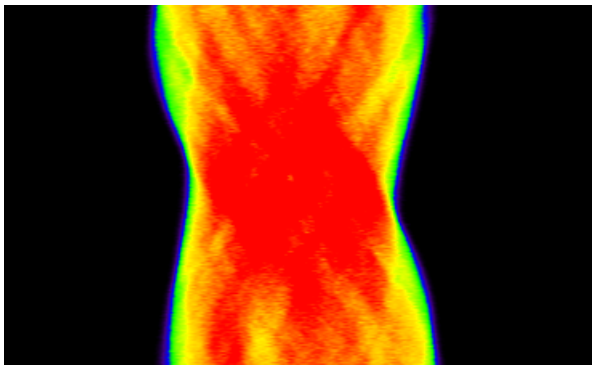


Figure 3.18: A single slice of the PET sinogram after the application of the Hanning filter.

3.2.3.4 Example output images

The following two figures show the same activity distribution having been generated at two separate count levels. Figure 3.21 shows the an axial slice of a reconstructed image at $5.4e^8$ counts. A reconstruction at $2.7e^8$ counts can also be seen in figure 3.22.

3.2.3.5 Summary

This section has described the process used to create phantom data containing realistic noise. The generated data are used throughout the remainder of this thesis to evaluate PVC techniques and as a test bed during algorithm development (section 5.2). The effects of attenuation, scatter, randoms and radioactive decay have been taken into account. The image generator does not directly simulate positron range or the non-colinearity of γ-rays, although this is implicit in the resolution parameter. The spatial resolution has been assumed to be uniform throughout the PET FOV. In reality, this assumption is incorrect, however the effects of space variance in brain PET are relatively small across the cortex, as is discussed in sections 3.1.2 and 3.1.3, and was therefore deemed acceptable for the purposes of phantom generation.

Monte-Carlo simulations using applications such as GATE [Schmidtlein et al., 2006] or PET-SORTEO [Reilhac et al., 2004] would be a more accurate approach for generating realistic phantom data, although they are restricted to the modeled scanner geometry. The aim of the image generator is to facilitate the creation of phantom data where the activity distribution and noise level can be easily changed. In addition, as FBP is used for reconstruction, the reconstructed resolution is known, making it straight-forward to evaluate PVC technique performance.

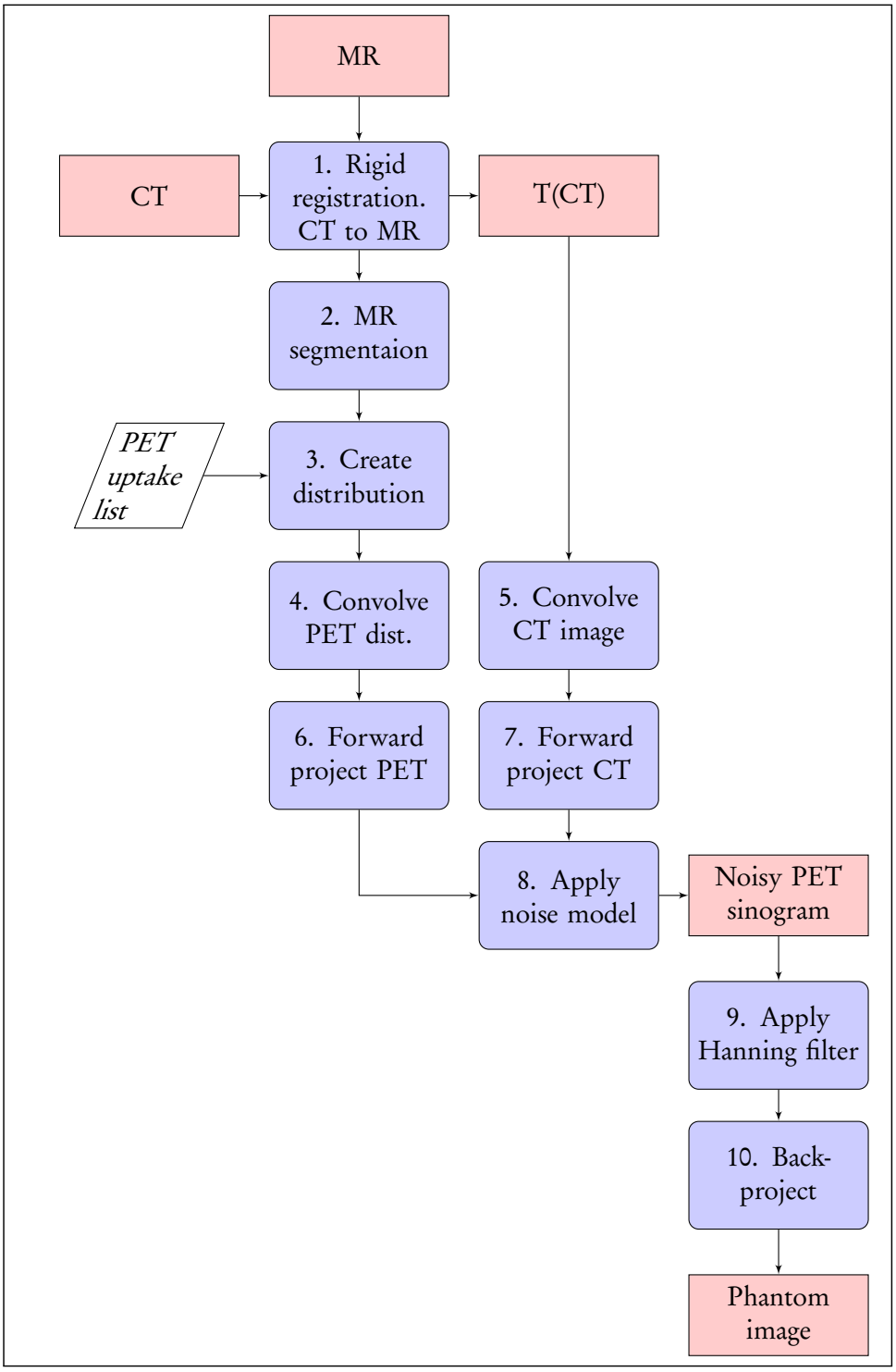


Figure 3.19: Overall system diagram of the process used to generate phantom images.

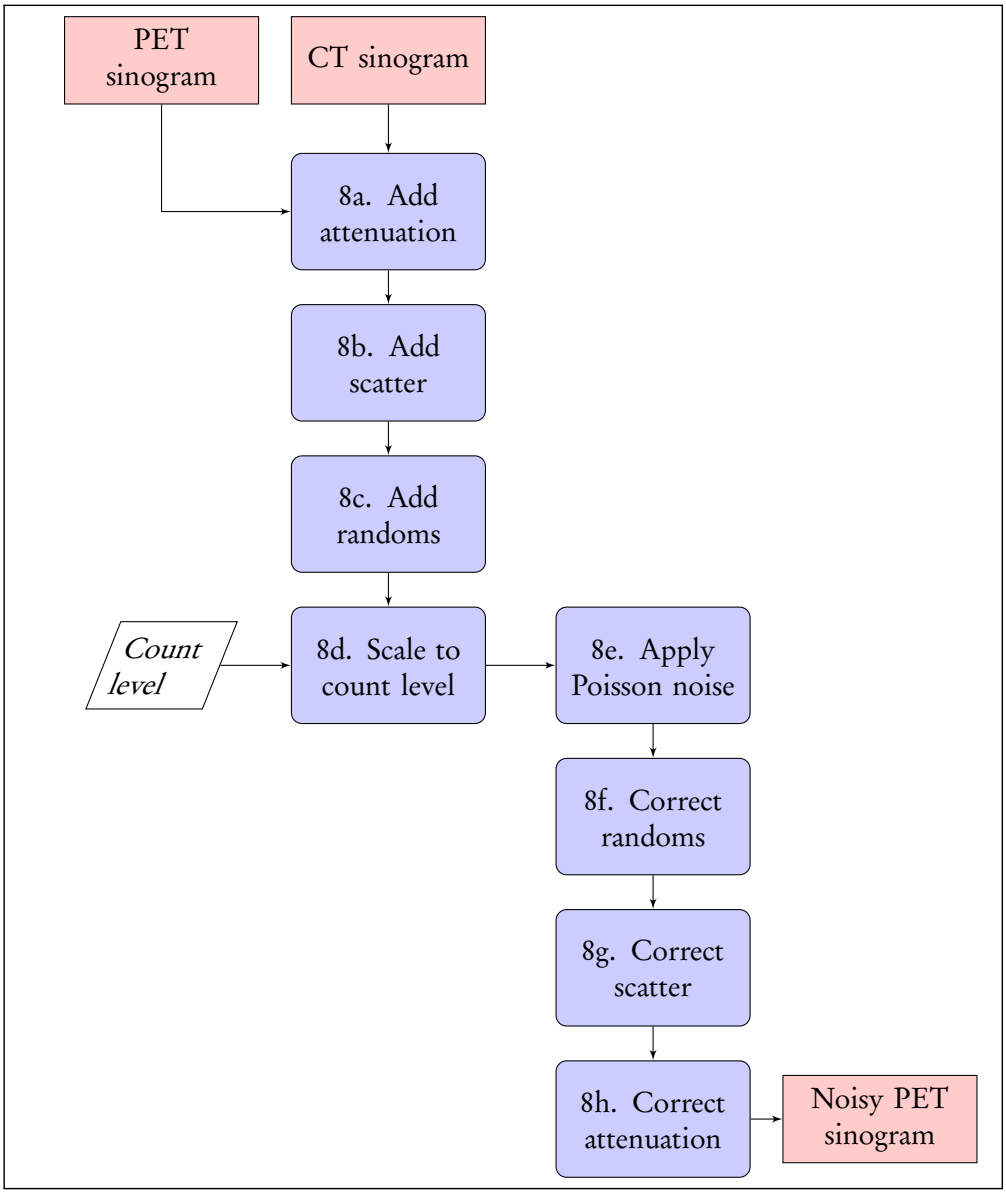


Figure 3.20: System diagram of the process used to generate noise in the phantom sinograms.

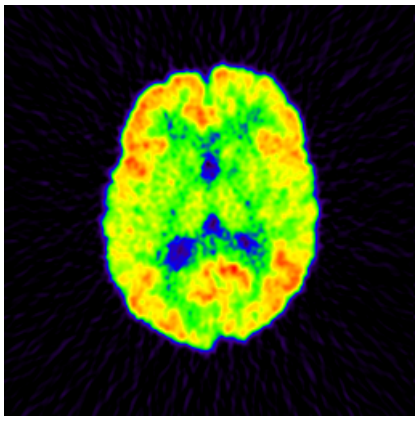


Figure 3.21: Example output of the phantom image generator at $5.4e^8$ counts.

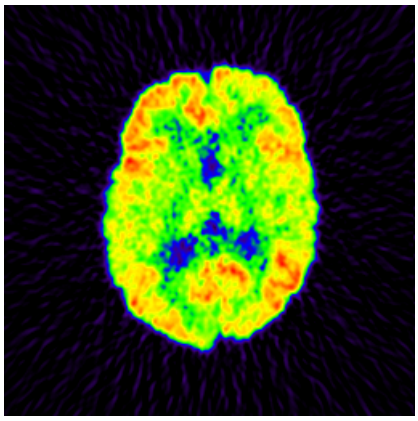


Figure 3.22: Example output of the phantom image generator at $2.7e^8$ counts.

3.2.4 Materials and methods

3.2.4.1 Dataset

A set of images were generated using the brain PET image generator (section 3.2.3, page 67). Ten noise realisation were created of an AD-like [^{11}C]PIB distribution. A count level of $5.4e^8$ was used to simulate a 30 minute [^{11}C]PIB acquisition [Vandenberghe et al., 2010b]. Figure 3.23 shows one of the ten images used during this study.

3.2.4.2 Methods

Two modifications to the VC method were evaluated in this study. The first was to apply additional Gaussian smoothing to the measured image before performing the standard reblurred VC. The additional smoothing is then accounted for in the PSF used for deconvolution. This extra smoothing is subsequently referred to as *VC-smooth*. The measured image i is smoothed by a further 2mm. As Gaussian FWHMs add in quadrature, to achieve a smoothing of 8mm total in an image with an initial resolution of 6mm resolution, a PSF (h') with a FWHM of

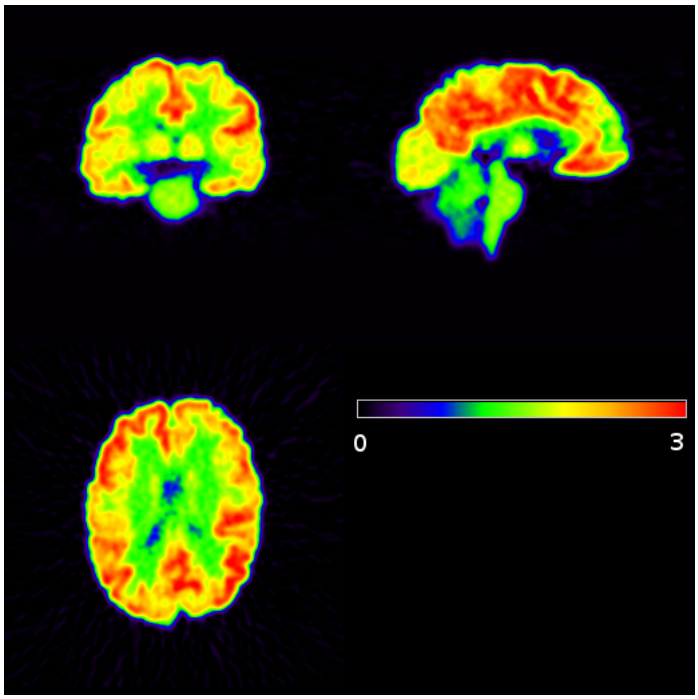


Figure 3.23: An example of an AD-like [^{11}C]PIB phantom SUVR image. Images were normalised using the cerebellar GM and are shown on a scale of 0 to 3.

5.29mm is required:

$$\sqrt{(6+2)^2-6^2}=5.29$$

Equation 3.7 then becomes:

$$\begin{aligned} t_{k+1}(x) &= t_k(x)+\alpha(\nabla\Psi(x)) \\ \Psi(x) &= (i'-t_k\otimes h)(x) \\ i'(x) &= (i\otimes h')(x). \end{aligned} \tag{3.12}$$

In this study all VC-smooth images have an additional smoothing of 2mm applied.

The second modification is the inclusion of the Huber prior in the deconvolution, referred to as *VC-prior*. The aim of including the prior is to reduce the noise as the deconvolution iterates, while not penalising edges. The prior penalises edges in the image less strongly than, for example, a Gaussian prior. The Huber function [Huber, 1981] is given by the equation:

$$\rho(z,\beta)=\begin{cases} z^2 & \text{if } |z|<\beta \\ 2\beta|z|-\beta^2 & \text{otherwise} \end{cases} \tag{3.13}$$

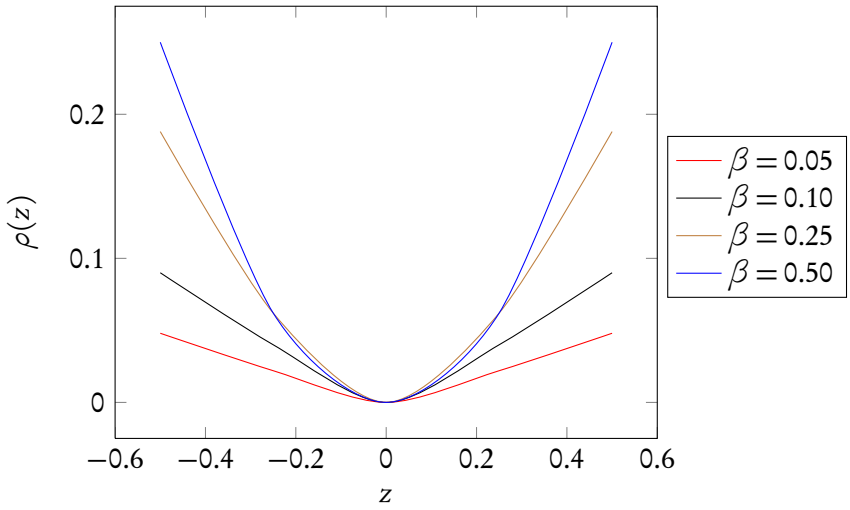


Figure 3.24: A plot of the Huber function for several β values.

It is quadratic about the centre and linear at the tails. The Huber function requires the parameter β. Large β values cause more smoothing at the edges, whereas small values smooth less. Figure 3.24 shows the Huber function for β values of 0.05, 0.1, 0.25 and 0.5. The value of the prior (H) is calculated over a $3 \times 3 \times 3$ neighbourhood:

$$H = \sum_{x} \sum_{j \in N(x)} \rho(|t_x - t_j|) \tag{3.14}$$

To include the Huber prior, into the VC deconvolution we first rewrite the least squares criterion (equation 3.6) to be:

$$\sum_{x} || \, i(x) - (t \otimes h)(x) \, ||^2 + \gamma H(x) \tag{3.15}$$

where H is the Huber prior and γ controls the influence of the prior. The iterative rule (equation 3.7) then becomes:

$$\begin{aligned} t_{k+1}(x) &= t_k(x) + \alpha(\nabla \Psi(x)) \\ \Psi(x) &= (i - t_k \otimes h)(x) + \gamma \times H(t_k)(x) \\ t_0(x) &= i(x). \end{aligned} \tag{3.16}$$

The parameters for β and γ were both set to 0.1 when evaluating the performance of the VC-prior deconvolution. The choice of parameters was found experimentally using preliminary data. The chosen β and γ values were selected in order to produce images which were visually

distinct from the standard VC method, while not introducing artefacts.

3.2.4.3 Analysis

Analysis was performed in terms of mean RC and mean regional coefficient of variation (CoV_r). The RC and CoV_r were both calculated using 10% trimmed mean values (see section 3.1.1.4, page 57). The reblurred VC, VC-smooth and VC-prior were evaluated for the 10 simulated [^{11}C]PIB images. The α parameter was set to 1.5 for all methods according the work of Tohka and Reilhac [2008]. The termination criterion (equation 3.9) was set to 0.001 and the maximum number of allowed iterations was 30.

3.2.5 Results

The images at 30 iterations of the VC deconvolution are shown in figure 3.25. All three VC deconvolution methods increased recovery in cortical GM regions compared to performing no PVC, with the largest increases observed in the the left frontal region. An increase of 21.5% in recovery was observed using VC, 17.7% with VC-prior and 10.7% using VC-smooth. Figure 3.26 shows the mean RC for the frontal, parietal and temporal lobes across noise realisations. At 30 iterations, no cortical regions had achieved complete recovery ($RC = 1$).

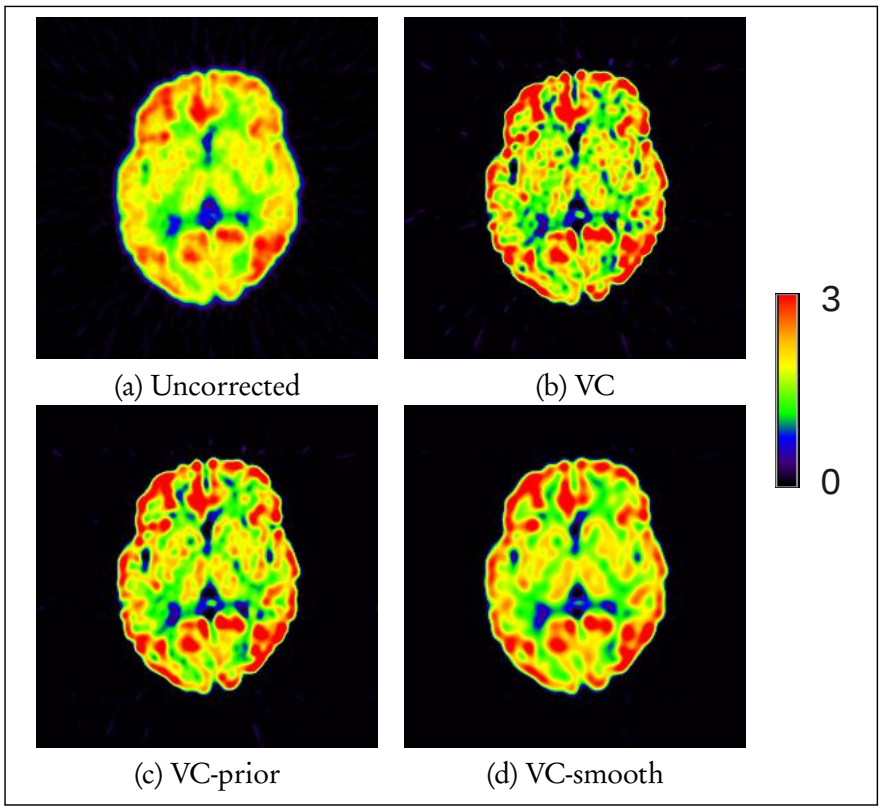
The CoV_r for the frontal, parietal and temporal lobes are reported in figure 3.26. Noise amplification was observed as each of the deconvolution methods iterated. Figure 3.27 shows the recovery versus CoV_r for 30 iterations of the left frontal and right parietal regions. The VC exhibited the highest CoV_r of the three, with VC-smooth being the lowest.

Figure 3.28 shows the recovery versus noise curve for the SWM. The RC in the SWM is greater than 1 due to spill-over from GM. The RC reduces as the deconvolution iterates. Reducing from 1.38 to 1.29 with VC, 1.38 to 1.31 with VC-prior and 1.41 to 1.35 with VC-smooth. After 30 iterations the SWM is still much greater than 1 and therefore positively biased. The increased RC in the initial image with VC-smooth is due to the pre-smoothing causing further PVEs.

VC-prior exhibited similar performance to VC but with a lower recovery and also lower noise. Quantitatively, the performance of the VC-prior was similar to that of VC in terms of both recovery and noise. Visually, the VC-prior images appeared smoother than VC, as can be seen in figure 3.25. The VC-prior deconvolution also appears to converge more quickly than VC and can be observed in the later iterations in both figures 3.27 and 3.28.

3.2.6 Discussion

The VC method, and deconvolution approaches in general, are known to suffer from noise amplification problems. In order to prevent noise destroying the recovered image, these ap-



(a) Uncorrected (b) VC

(c) VC-prior (d) VC-smooth

Figure 3.25: An axial slice from one realisation after 30 iterations. Images were normalised using the cerebellar GM and are shown on a scale of 0 to 3, without PVC (a), van Cittert (b), VC with a Huber prior (c) and VC with additional pre-smoothing (d).

proaches are often terminated prematurely, leading to incomplete recovery. This study has used realistic phantom data to evaluate methods for reducing the amplification of noise during the VC deconvolution. A [^{11}C]PIB distribution was simulated, with analysis focusing on improvement to cortical GM regions, as these areas are of clinical interest.

Two different noise suppression approaches were evaluated, VC-smooth and VC-prior. Lower noise was observed for both of the applied methods compared to the standard VC. However, this lower noise tended to be coupled with lower recovery. At 30 iterations, none of the deconvolved images had achieved complete recovery. The highest observed mean RC in cortical GM regions was found in right insula (VC: 0.83; VC-prior: 0.81; VC-smooth: 0.79). All other cortical regions had a RC of less than 0.8 irrespective of the applied technique. This suggests that in terms of recovery, 30 iterations is too early a point at which to stop. However, as can be seen in figure A.4 (appendix A, page 171), complete recovery is not acheived even at high numbers of iterations, suggesting that there is a limit to the recovery achievable with the VC technique. The VC-prior achieved similar, although lower, levels of recovery compared to the VC method. For example, recovery in the left frontal region VC-prior was 3.3% lower than VC. In the same region, VC-smooth was 9.2% lower than VC. The reduced recovery

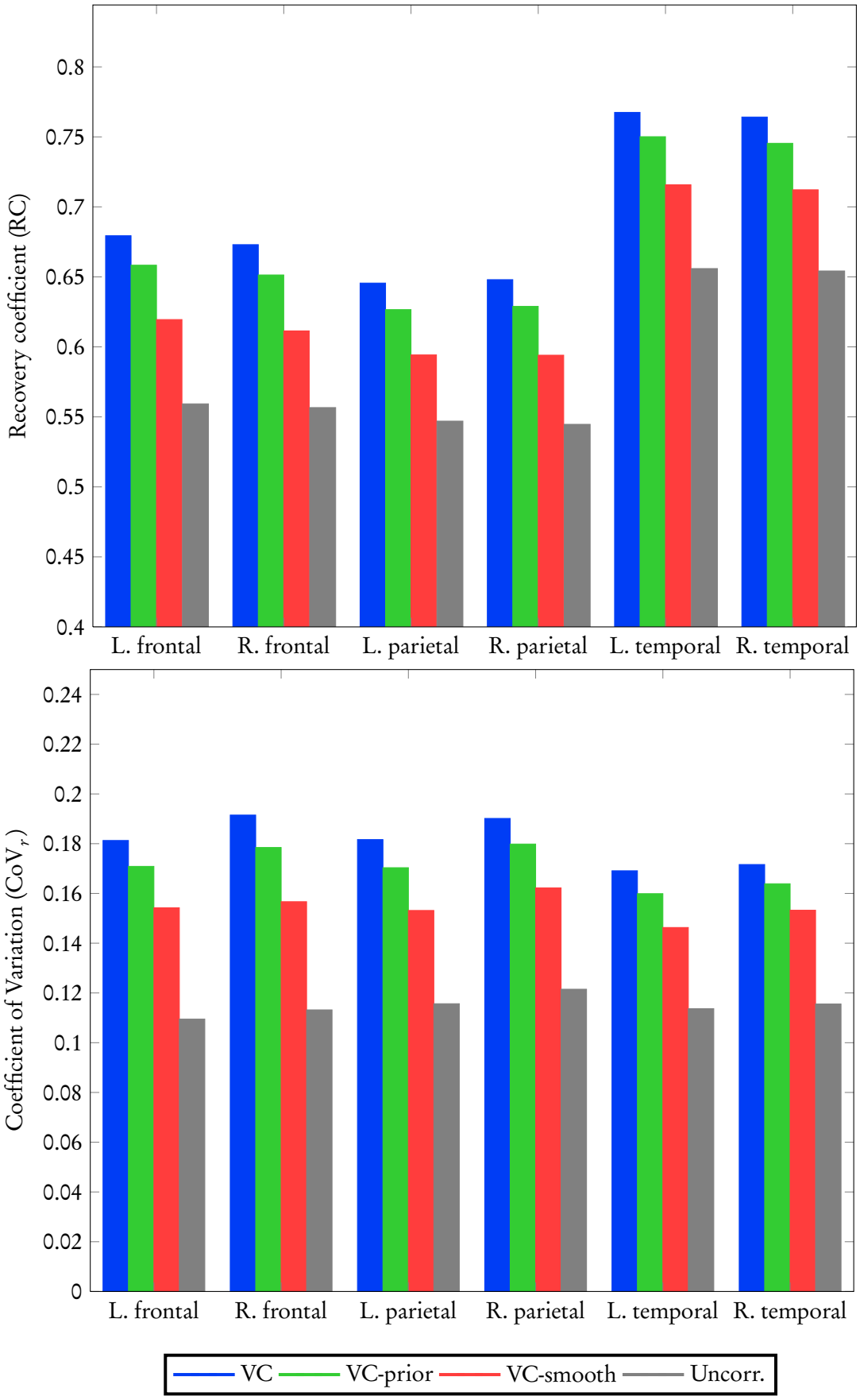


Figure 3.26: The performance in terms of recovery coefficient (top) and regional coefficient of variation (bottom) of the van Cittert (VC), VC with a Huber prior (VC-prior) and VC with additional pre-smoothing (VC-smooth). The RC without PVC (Uncorr.) is also shown.

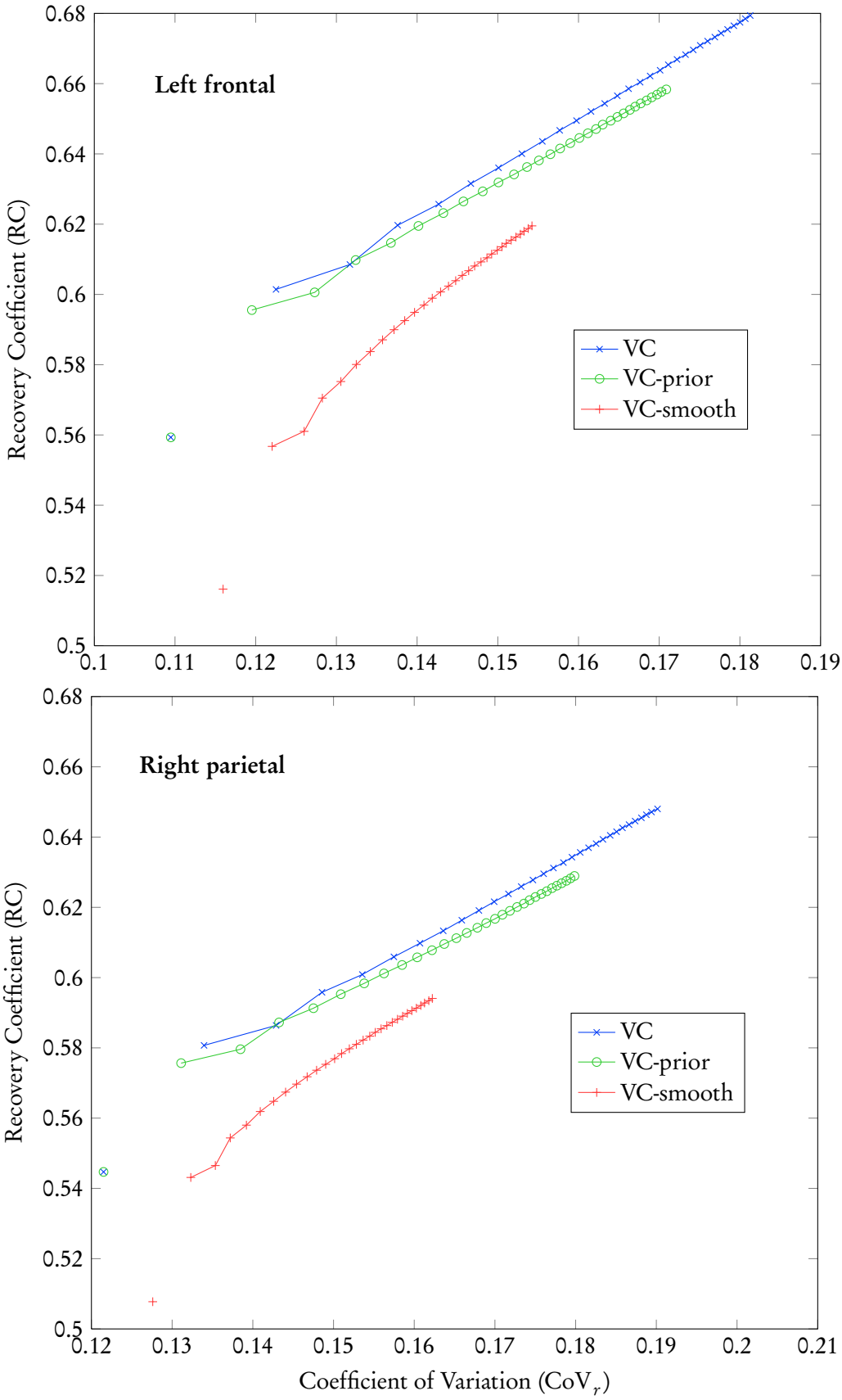


Figure 3.27: Recovery versus noise plots of the left frontal (top) and right parietal (bottom) regions for iterations 1 to 30. Single points indicate the initial start point before deconvolution.

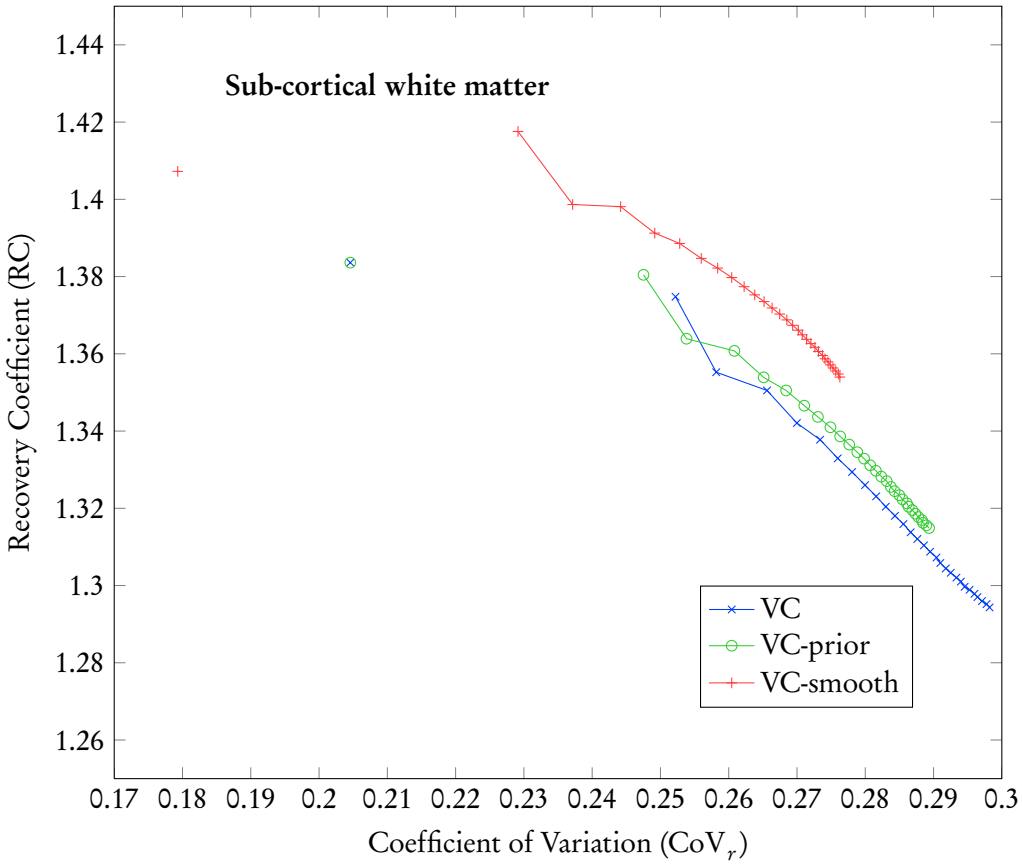


Figure 3.28: Recovery versus noise plot of the sub-cortical white matter for iterations 1 to 30. Single points indicate the initial start point before deconvolution.

observed with VC-smooth is due to the additional pre-smoothing causing further reductions in cortical GM regions and therefore requiring higher iteration numbers to reach similar levels of recovery.

Noise increases were observed over 30 iterations for each of the methods applied. These increases were highest with VC, with VC-smooth producing the lowest CoV_r of the three techniques. Again, assessing the left frontal region, the average CoV_r observed at 30 iterations was 0.18 with VC, 0.17 with VC-prior and 0.15 when VC-smooth was applied. As the region had an initial CoV_r of 0.11, this is an increase of 63.6%, 54.5% and 36.5% respectively. For VC (and perhaps VC-prior), the noise increase at 30 iterations is already at a point where the images begin to degrade qualitatively.

Both VC-smooth and VC-prior have the effect of smoothing the image, so while the voxel variance is reduced, the overall PVE has not been completely corrected. By its nature, VC-smooth increases the degree of PVE and this is most easily observed in the SWM (figure 3.28) where an increase of 2.2% was observed in the initial image (t_0) of VC-smooth. The large RC values seen in the SWM are caused by spill-over from the high cortical GM uptake.

As the images are deconvolved, the RC in the SWM reduces closer to 1. This is to be expected as GM spill-over is corrected for. However, although reduced, the SWM is still over-estimated ($RC > 1$) after deconvolution, again suggesting incomplete recovery due to early termination.

The performance of VC-prior was similar to that of VC. In cortical regions, lower recovery was observed with VC-prior, as was a lower CoV_r. It is the γ parameter that affects how much the estimated image is smoothed during the deconvolution. The chosen γ value of 0.1 was found experimentally with lower values resulting in negligible changes to the images. Higher γ values produced smoother images although ring-like artifacts were visible for values of more than 0.5. Severe artifacts were observed for $\gamma \geq 1$ and caused the deconvolution to diverge.

Similarly, the size of the PSF used to pre-smooth the images in VC-smooth was found by experimentation. The value of 2mm was selected because the images were visually less noisy and preliminary tests indicated that the pre-smoothing did not severely degrade the recovery in cortical GM. Once applied to the entire dataset, VC-smooth achieves lower levels of recovery compared to the other techniques, although this is coupled with markedly lower noise increases. However, iterating the VC-smooth technique beyond 30 iterations (data not shown) resulted in similar recovery and similar noise levels to VC and VC-prior, indicating that with more iterations the VC-smooth tends to a similar solution as the other methods.

The parameters chosen for the noise suppression and the deconvolution itself influence the observations of this study. The deconvolution parameters were based on that of Tohka and Reilhac [2008]. These may not be optimal for this application but were not investigated further in this study. The main issue with deconvolution techniques appears to be the noise amplification rather than the iterative scheme. As recently discussed, the parameters used for the Huber prior and the pre-smoothing also affected the results. Different priors may be more suited for deconvolving amyloid PET data although this is beyond the scope of this study. In summary, while deconvolution approaches are desirable from a clinical perspective, as there is no requirement for anatomical information, the attainable recovery in cortical GM is considerably lower than anatomy-based methods. The noise suppression techniques applied in this study did not notably improve the performance of VC.

3.2.7 Conclusion

The performance of the VC deconvolution was shown to be inferior to that of anatomy-based approaches in section 3.1. This study has evaluated the use of noise suppression with the VC deconvolution in an attempt to improve the performance of the method. Realistic simulations of [^{11}C]PIB PET data were used to investigate two approaches, VC-smooth and

VC-prior, and were compared against the standard reblurred VC. Reductions in noise were observed with both approaches, however this tended to be coupled with a reduction in recovery. Neither approach resulted in notable improvements compared to the standard VC. While different choices in parameters would have determined a different set of observations, the results suggests that there is a limit to the degree of recovery that can be achieved with VC.

Cortical GM regions in normal subjects are usually no more than 7mm thick. In diseased subjects with atrophied brains, the cortex may have a thickness of only a few mm. As the cortical thickness is equal to or less than the spatial resolution of the scanner, it is likely that deconvolution approaches may never reach full recovery even with the aid of noise suppression. The VC deconvolution may be more suited to applications where the target object is larger, such as striatal imaging or oncology.

Chapter 4

Comparing partial volume correction techniques using clinical data

This chapter describes the application of PVC to clinical amyloid PET data. Section 4.1 reports the effects of PVC on the quantification of a [^{18}F]flutemetamol cohort using two PVC techniques. The reproducibility of [^{18}F]flutemetamol is evaluated in section 4.2. In section 4.3 a comparison between two amyloid tracers, [^{11}C]PIB and [^{18}F]flutemetamol, is reported. The comparison evaluates alternative reference regions for image normalisation in combination with PVC.

4.1 Quantitative analysis

The quantitative accuracy of PET images is degraded by PVEs. The displacement of activity between neighbouring regions causes quantification errors. These errors may obscure the patterns of neurogenerative diseases, such as AD, that can be observed with amyloid PET ligands. Accurate quantitative assessment of these tracers is desirable for the purposes of tracking disease progression [Waragai et al., 2009] and also to evaluate potential new therapies [Chow et al., 2010].

AD is characterised by progressive atrophy and ventricular expansion. The volume losses in brain structures due to the destructive process of atrophy, in particular the cortical GM regions, cause these structures to be more severely affected by the PVE. In the presence of atrophy, apparent signal changes can at least be partially attributed to PVEs. Atrophy attenuates the true signal in the region. PVC aims to correct for this effect although it is not routinely applied during the quantification of amyloid PET images, despite diseased subjects often exhibiting widespread cortical atrophy.

This study investigates the effects of PVC on amyloid PET quantification in a clinical cohort, applying RBV correction and the mMG.

4.1.1 Materials and methods

4.1.1.1 Subjects

The cohort used in this study consisted of 72 subjects from a multicentre clinical trial (see Vandenberghe et al. [2010b] for details). Each subject underwent a [^{18}F]flutemetamol PET scan and T1-weighted MRI. The PET scans were acquired over 30 minutes (6 × 5-min frames), with acquisition commencing 85 minutes post-injection. Two subject were excluded from the analysis. The MRI cortical parcellation stage failed in the case of one AD subject. A HC subject was also excluded due to abnormally high levels of amyloid uptake.

The subjects were all classified by clinical evaluation as either AD ($n = 26$), MCI ($n = 20$) or HC ($n = 24$). The control group was further sub-divided into elderly healthy control (eHC) and young healthy control (yHC). The threshold used to divide the control group was 65 years of age. There were 10 participants in the eHC group and 14 yHC subjects. The median (range) ages of the AD, MCI, eHC and yHC groups were 71.3 (26.9), 70.9 (25.9), 73.7 (10.9) and 43.8 (33.9), respectively.

4.1.1.2 Image acquisition

The data were acquired at three sites. The first centre acquired PET data using a Siemens Biograph 16, the second centre used a Siemens ECAT Exact HR+, with the third centre using a GE Advance. Each centre employed different PET image reconstruction algorithms. OSEM was used at the first, with analytical reconstruction being performed at the other two sites.

4.1.1.3 Point-spread function estimation

The PVC techniques RBV correction and mMG require a measure of the PSF. The PSF was measured at each centre using a Radiology Support Devices (Radiology Support Devices Inc, Long Beach, California, USA) striatal brain phantom. 10MBq of [^{18}F]FDG (8kBq/ml) was used to fill the main chamber of the phantom. The right putamen was filled with 16 kBq/ml and the right caudate with 8 kBq/ml. Images were acquired for 30 minutes. Reconstruction was performed in 2-D with attenuation, scatter, randoms and dead-time corrections applied. Images were then post-smoothed with a Gaussian of between 3 and 5mm FWHM. The between-center variability in activity ratios between the putamen insert to the main chamber was estimated to be approximately 5% [Vandenberghe et al., 2010b]. The FWHM was estimated by finding the best fit between the reconstructed image of the striatum and the phantom convolved with a Gaussian PSF. Values were found to be 5.3mm for the first centre, 6.3mm for the second and 6.7mm at the third centre.

4.1.1.4 Image processing

Region definition: A parcellated subject MRI is required to perform the PVC. The parcellation of each scan was performed using the FreeSurfer (FS) package (see section 2.4.2.1, page 43). From the detailed FS parcellation, a mask image consisting of 14 volumes of interest (VOIs) was generated for each subject. The VOIs consisted of 10 GM regions and 4 other regions. The GM volumes consisted of the frontal, central, parietal, temporal, occipital, cingulate, insula, sub-cortical GM (caudate, putamen, thalamus, amygdala, pallidum), hippocampi and cerebellum. The 4 other VOIs consist of SWM, cerebellar WM, the brain stem and ventricular and CSF space.

The VOIs were defined to parcellate the brains into either clinically relevant regions such as the frontal, parietal and temporal lobes or areas that could suffer from large PVEs such as the occipital lobe. An additional central lobule, consisting of the pre-, para- and post-central gyrus was also included as this area is though to be relatively spared in AD until late stages of the disease [Braak and Braak, 1995]. The central lobule is also surrounded by 'hotter' GM regions (in AD-like patterns of uptake) and was thought to potentially suffer severe PVEs. For a complete list of the FS labels used to generate the mask images see appendix D.1. In addition, a three-class tissue segmentation mask (GM, WM and CSF) was generated from the FS labels on a per-subject basis for the purposes of the mMG correction.

Registration: Each of the 5-minute PET frames were rigidly registered together. The registered frames were then added together to create a 30-minute sum image. A rigid registration was performed using a block-matching based technique [Ourselin et al., 2001] between the subject PET and subject MRI. Subsequent image processing and analysis was performed in the subject MRI space.

PVC and normalisation: The PVC techniques, RBV correction and mMG were performed using the PET data and the region masks. The FWHM of the PSF used for PVC was applied according to the centre from which the scan was acquired. The scans (corrected and uncorrected) were normalised using the cerebellar GM as a reference region. Images were normalised using FSL v4.1 (FMRIB, Oxford University, UK). After normalisation, these images are referred to as SUVR images. All subsequent analysis was performed on these SUVR images.

4.1.1.5 Data analysis

Images were analysed in terms of regional mean SUVR values, using regions defined at the parcellation stage. Non-GM regions were not evaluated for the mMG as this correction method operates on GM voxels alone. To reduce the effects of outliers, a 10% trimmed mean value

was calculated for each region. The CoV_s was calculated across subjects within a group. The intra-regional variability (CoV_r) was calculated, averaged across groups. The image analysis was performed using FSL v4.1. Statistical analysis was undertaken in SPSS v17.0 (SPSS, Chicago IL.) and G*Power 3 (Institut für Psychologie, Christian-Albrechts-Universität Kiel, Kiel, Germany). Where statistical comparisons failed to reach significance, the Minimum Detectable Difference (MDD) was calculated, given by the effect size the test was powered to identify at 80% and a $p < 0.05$ (unless otherwise stated) multiplied by the pooled SD of the two groups.

4.1.2 Results

4.1.2.1 SUVR analysis

The application of PVC significantly increased SUVR values in cortical regions. The largest increases were observed in AD subjects. Figure 4.1 shows SUVR images of one AD subject with and without PVC applied. Visually, there is a noticeable improvement in the contrast between the GM \WM border when compared with the uncorrected data. SUVR values, with and without correction, for the frontal lobe, parietal lobe and hippocampi can be seen in tables 4.1, 4.2 and 4.3 respectively.

	Uncorrected	mMG	RBV
AD	2.07 ± 0.35 (0.17)	2.99 ± 0.63 (0.21)	3.24 ± 0.73 (0.22)
MCI	1.56 ± 0.41 (0.27)	2.03 ± 0.83 (0.41)	2.20 ± 0.96 (0.43)
eHC	1.21 ± 0.15 (0.12)	1.35 ± 0.24 (0.17)	1.41 ± 0.26 (0.18)
yHC	1.16 ± 0.08 (0.07)	1.26 ± 0.19 (0.15)	1.31 ± 0.22 (0.17)

Table 4.1: Effect of PV-correction on frontal lobe SUVR [mean ± SD (CoV_s)].

	Uncorrected	mMG	RBV
AD	2.02 ± 0.31 (0.15)	2.95 ± 0.62 (0.21)	3.26 ± 0.72 (0.22)
MCI	1.55 ± 0.38 (0.25)	2.02 ± 0.88 (0.44)	2.22 ± 1.03 (0.46)
eHC	1.20 ± 0.07 (0.06)	1.25 ± 0.08 (0.06)	1.31 ± 0.09 (0.07)
yHC	1.21 ± 0.07 (0.06)	1.30 ± 0.19 (0.14)	1.36 ± 0.22 (0.16)

Table 4.2: Effect of PV-correction on parietal lobe SUVR [mean ± SD (CoV_s)].

Of the two PVC methods, RBV correction resulted in the largest changes to cortical SUVR values. The parietal lobe exhibited the greatest changes of the assessed regions. In AD subjects, the parietal lobe increased 61.4% after RBV correction, with mMG the observed increase was 46.0%. The MCI group displayed increases of 43.2% using RBV correction and 30.3% after the mMG. In the control group, increases in SUVR with RBV correction was 9.2% in the eHC group and 12.4% in the yHC. When applying mMG, the control group

	Uncorrected	mMG	RBV
AD	$1.35^{\ddagger} \pm 0.16$ (0.12)	$1.49^{\star} \pm 0.22$ (0.15)	$1.36^{\dagger} \pm 0.21$ (0.16)
MCI	1.24 ± 0.14 (0.11)	1.31 ± 0.14 (0.11)	1.23 ± 0.13 (0.11)
eHC	$1.28^{\ddagger} \pm 0.07$ (0.05)	$1.26^{\star} \pm 0.08$ (0.06)	$1.27^{\dagger} \pm 0.10$ (0.08)
yHC	1.33 ± 0.07 (0.05)	1.21 ± 0.10 (0.08)	1.22 ± 0.11 (0.09)

Table 4.3: Effect of PV-correction on hippocampal SUVR [mean ± SD (CoV_s)]. ‡Two-tailed, unpaired Wilcoxon rank-sum test between AD and eHC not significant at 0.05 threshold ($p = 0.37$). *Two-tailed, unpaired Wilcoxon rank-sum test between mMG-corrected AD and eHC highly significant at 0.05 threshold ($p < 0.004$). †Two-tailed, unpaired Wilcoxon rank-sum test between RBV-corrected AD and eHC not significant at 0.05 threshold ($p = 0.27$).

showed increases of 4.2% and 7.4% for the eHC and yHC respectively.

The two PVC methods were also compared in terms of CoV_s. Overall the RBV correction exhibited higher values for CoV_s than the mMG. It should be noted that similar results were reported in the phantom experiments in section 3.1.3 (page 63), where this behaviour was thought to be caused by GM spill-over inducing a bias in the WM estimate. This could explain the apparent difference between the CoV_s for RBV and mMG in this study.

The CoV_r was also assessed for PV-uncorrected, mMG and RBV correction. A lower CoV_r was observed using RBV correction in GM regions than in the uncorrected data or when using mMG. The lower CoV_r suggests that regions corrected with RBV are more uniform than mMG or when no PVC has been applied. The only exception to this was in the hippocampus, where CoV_r increased after PVC. The mean CoV_r values for the frontal, parietal and hippocampal regions can be seen in tables 4.4, 4.5 and 4.6.

	Uncorrected	mMG	RBV
AD	0.13 ± 0.04	0.17 ± 0.02	0.11 ± 0.01
MCI	0.21 ± 0.07	0.19 ± 0.03	0.11 ± 0.01
eHC	0.24 ± 0.02	0.20 ± 0.06	0.11 ± 0.02
yHC	0.24 ± 0.03	0.23 ± 0.05	0.12 ± 0.03

Table 4.4: Regional coefficient of variation (CoV_r) for the frontal lobe [mean ± SD].

	Uncorrected	mMG	RBV
AD	0.14 ± 0.03	0.18 ± 0.02	0.11 ± 0.01
MCI	0.19 ± 0.06	0.19 ± 0.03	0.11 ± 0.01
eHC	0.22 ± 0.02	0.21 ± 0.04	0.10 ± 0.01
yHC	0.22 ± 0.02	0.23 ± 0.04	0.12 ± 0.03

Table 4.5: Regional coefficient of variation (CoV_r) for the parietal lobe [mean ± SD].

	Uncorrected	mMG	RBV
AD	0.13 ± 0.03	0.27 ± 0.04	0.18 ± 0.03
MCI	0.12 ± 0.02	0.25 ± 0.03	0.15 ± 0.02
eHC	0.11 ± 0.02	0.26 ± 0.02	0.15 ± 0.01
yHC	0.14 ± 0.03	0.29 ± 0.06	0.17 ± 0.04

Table 4.6: Regional coefficient of variation (CoV_r) for the hippocampus [mean ± SD].

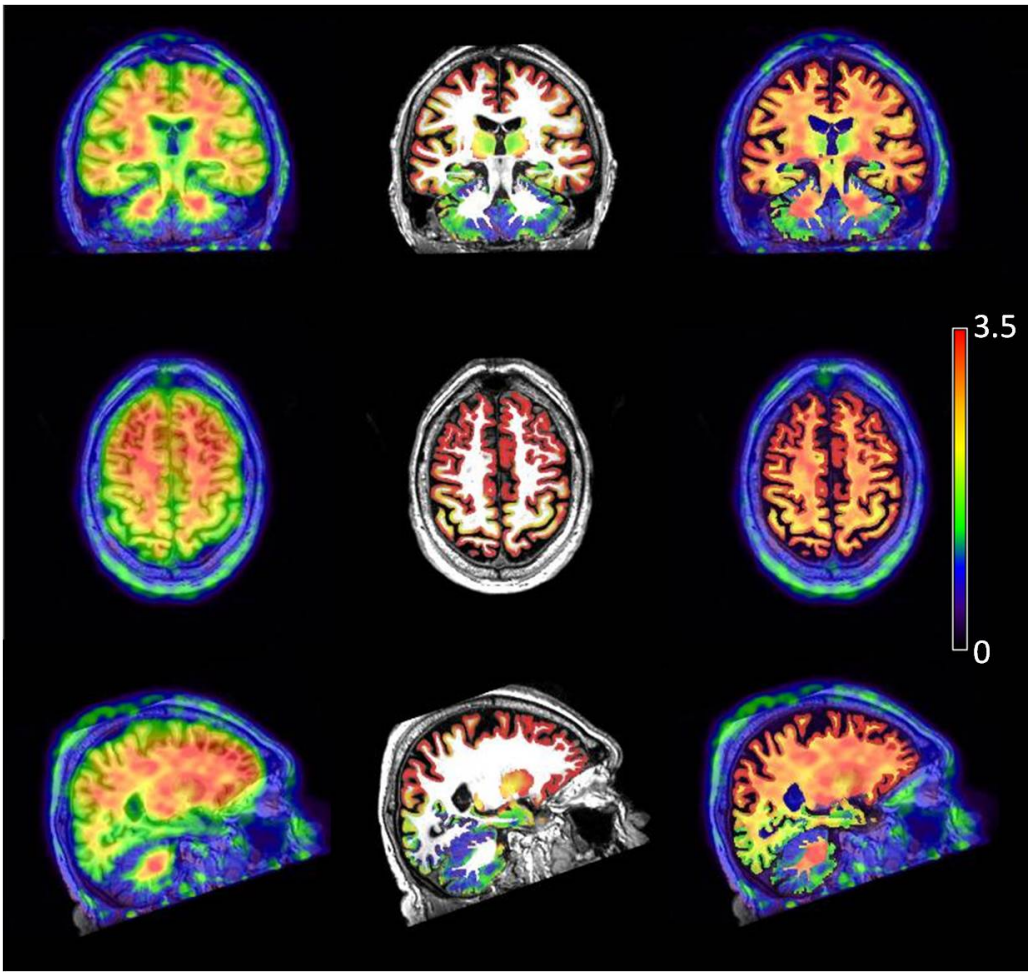


Figure 4.1: SUVR images of an AD subject with and without PVC, overlaid on T1-weighted MRI. The PV-uncorrected images (left), mMG-corrected images (centre) and RBV-corrected images (right) are shown, normalised to cerebellar GM.

4.1.2.2 MCI subjects

Higher variability was observed in the cortical regions of the MCI group than either the AD or control groups. The MCI group has a bimodal distribution in cortical brain regions. MCI subjects tend to present with either an AD-like or control-like distribution. The existence of bimodal distributions in MCI populations has been previously reported [Li et al., 2008]. MCI subjects are often classified as being either amyloid-positive or amyloid-negative. This classification is usually achieved by setting a SUVR threshold of between around 1.45 and 1.6

on a cortical region.

In order to derive a threshold for this cohort, a *composite cortical SUVR value* was found for each subject by averaging the mean values of the frontal, parietal, temporal and cingulate regions. The cortical SUVR values of the eHC and AD group were used to perform k-means cluster analysis. The threshold was then defined as being the average distance between the centres of the clusters. Without PVC, a threshold of 1.6 was found. When RBV correction was applied, this threshold increased to 2.3. The PV-corrected and -uncorrected threshold was applied to the MCI cohort. 11 subjects were classified as amyloid-negative and the remaining 9 as amyloid-positive. Both thresholds bisect the MCI group into a bimodal distribution when applied to their respective datasets. The distribution of subjects by their composite cortical values can be seen in figure 4.2.

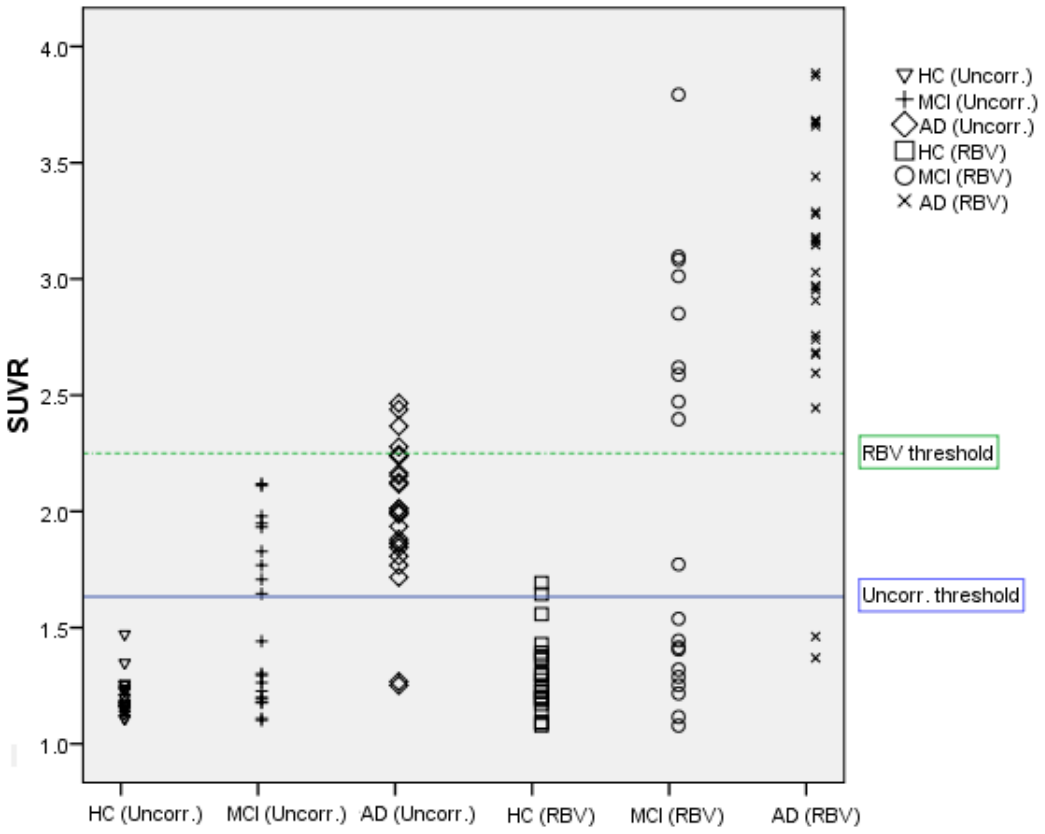


Figure 4.2: Cortical SUVR values (average of frontal, parietal, temporal and cingulate regions) of each subject with RBV correction and without PVC.

4.1.2.3 The hippocampus

Cortical GM values increased when both PVC approaches were applied. Although the two techniques produced different regional mean values, for cortical GM the two methods followed the same trend of large increases in AD subjects, some increases in the control group,

with MCI subjects exhibiting one or the other. The two methods did disagree with respect to the hippocampus. A small, less than 1% increase was observed in the mean SUVR of the AD group with RBV correction. However, the mMG method produced a 10.4% increase in the same group.

Paired Wilcoxon signed-rank tests were performed between the corrected and uncorrected SUVR values. The small increase with RBV correction was not found to be significant ($p = 0.85$) at the $p < 0.05$ threshold. The difference between the corrected and uncorrected hippocampal SUVR values with mMG was highly significant ($p < 0.001$). The difference between the two PVC methods for AD subjects was also tested and was found to be significant ($p < 0.001$). In the MCI group, mMG resulted in a 5.7% increase in the hippocampus, compared to a 0.8% decrease with RBV correction. Statistically significant differences between the AD and eHC groups were only observed when applying the mMG technique (see table 4.3).

4.1.3 Discussion

The effects of PVC were evaluated using a [^{18}F]flutemetamol dataset from a multicentre clinical trial. Degradation in the quantitative accuracy of brain PET by PVEs is influenced by scanner resolution and the degree of atrophy. Subjects with higher amyloid burden suffer from large quantification errors due to PVEs. The key finding of this study is that cortical GM regions are severely affected by PVEs. Spill-over from other tissue types and losses to the background are the main causes of PVEs in the cortical GM. However, this study also confirmed observations of *within*-tissue PVEs, made in the previous chapter during the phantom study (see section 3.1.3, page 60), that GM regions can suffer from PVEs due to spill-over from other GM regions. This condition was observed in the hippocampi. The RBV correction technique was earlier shown to correct for PVEs of this type, while the MG approaches induced biases. The same behaviour has been observed in this study using clinical data and suggests that RBV is superior to the mMG when faced with *within*-tissue variability.

Large increases in cortical SUVR values, particularly in AD subjects were observed after PVC. This trend was also observed in the phantom study (section 3.1) and suggests that the increases are genuine findings. Increases tended to be larger when applying RBV correction as opposed to the mMG correction. Again this was seen in the phantom study and it can be inferred that the RBV correction represents the true cortical uptake more accurately than the mMG approach.

The mMG relies on the GTM method to estimate the mean WM value for the purposes of the MG correction. The GTM relies on the assumption that each subregion is uniform.

When applying a 3-compartment mMG, as is performed here, the WM estimate is biased by spill-over from the (non-uniform) GM compartment. A positive bias in the WM causes a reduction in cortical GM values due to over-correction of the WM compartment. In addition, variability between WM regions has also been observed. This is covered in detail in section 4.3 (specifically in 4.3.2.1, page 109). WM variability also appears to affect the accuracy of the mMG.

Inter-subject variability (CoV_s) increased for all cortical GM regions after PVC. At first this may seem counter-intuitive as one might expect clinical groups to be more tightly distributed after correction. However, it appears that correcting for PVEs, and therefore more accurately representing the true distribution in each region of a subject, increases the inter-subject variability. This behaviour was observed in the previous phantom study and appears to be a genuine finding rather than a PVC-induced error.

The PV-corrected inter-voxel variability (or CoV_r) is affected by two factors. First, the CoV_r may decrease due to a reduction in regional bias. Second, an increase may be observed due to noise amplification. The measured CoV_r is a combination of these two factors. Reductions in CoV_r within cortical GM regions when applying RBV correction implies that the correction does not significantly increase noise and reduces variability in cortical regions. When applying the mMG approach, similar or increased variability was observed compared to PV-uncorrected data. This increase is likely to be due to biases induced by and inability to correct for *within*-tissue variability.

The MCI group were classified as amyloid-positive or amyloid-negative based on a cortical SUVR threshold. The threshold increases (+44%) after PVC has been applied. Given the large increases in cortical SUVR values, it is unsurprising that the threshold changes. The 'amyloid status' of all subjects remained constant irrespective of PVC. While the change in threshold is not important in the sense that it does not appear to improve discrimination between subject groups, it is important to state that the appropriate threshold should be applied depending on whether or not PVC has been performed. As can be seen in figure 4.2, applying the wrong threshold would result in the misclassification of subjects.

The key objective of this thesis was to evaluate whether PVC improves quantitative accuracy rather than improving discrimination. It is hoped that accurate PVC may reveal subtle changes in longitudinal studies. The application of PVC could be particularly important as temporal signal changes could be masked by PVEs. This masking of changes may occur if signal increases are coupled with atrophy. The effects of atrophy can attenuate the signal causing the activity to appear static. Section 5.5 addresses the issue of longitudinal tracking in greater

detail.

The earlier phantom study identified the hippocampus as an area that appeared to suffer from PVEs from neighbouring GM regions. As mentioned previously, the mMG is unable to correct for PVEs within a tissue class. In this study, a statistically significant difference in the AD group was found after mMG correction. This was not observed with RBV correction. In addition, a significant difference between the eHC group and the AD group was only observed after mMG was performed.

In AD, the hippocampi are surrounded by hot GM regions, in particular the parahippocampal gyrus. The observed increase when applying mMG is likely to be from uncompensated spill-over from this region. Previous studies based on pathology have shown that there are low levels of amyloid burden in the hippocampus [Braak and Braak, 1995], despite being one of the first areas to exhibit atrophy on MRI. This appears to be an erroneous bias induced by the mMG correction.

Comparisons between the AD and eHC groups did not reach significance when testing hippocampal SUVR values for the uncorrected data and the RBV-corrected data. The MDD was calculated for both tests and found to be 0.15 and 0.20, respectively. The difference in means between the groups for the uncorrected data was 0.07 and 0.09 in the case of the RBV-corrected. As the group difference is consistent across both datasets, this may suggest that the tests were underpowered in this instance. The comparison of the uncorrected and RBV-corrected AD group was also not significant (MDD =0.09) with a difference in means of 0.01.

4.1.4 Conclusion

This study has shown that large increases in cortical SUVR values are observed when PVC is applied to amyloid PET data. This is inline with the previous findings of the phantom study. RBV correction has been shown to reduce intra-regional variability, suggesting that regions are more uniform after RBV correction.

The mMG approach has been shown to suffer from *between*-compartment and *within*-compartment variability. This approach uses assumptions about regional uniformity which appear to be unsuitable for amyloid PET. GM uptake can vary greatly within a subject, particularly in diseased states. This study has shown that GM variability can induce biases when performing mMG correction.

Increases in inter-subject variability after PVC were observed for both RBV correction and mMG. The increase is believed to be a genuine result of correcting for PVEs, and more accurately reflects the true amyloid distribution in the subject scans. RBV correction was shown to outperform the mMG as it can account for variability within tissue compartments.

The improved quantification provided by RBV correction may improve discrimination and longitudinal tracking.

4.2 Assessment of reproducibility using clinical data

This section reports the evaluation of PVC in a clinical reproducibility dataset. The study investigates the quantitative reproducibility of [^{18}F]flutemetamol data. Quantification was assessed through regional analysis of uncorrected and RBV-corrected data.

4.2.1 Materials and methods

4.2.1.1 Subjects

A subset of the cohort that were part of the clinical trail [Vandenberghe et al., 2010b] described in section 4.1.1 underwent two [^{18}F]flutemetamol PET scans. The second scan was performed for five of the subjects within a 7-day interval (13 days in one subject). In addition, a T1-weighted MRI scan was acquired for all subjects at baseline only. The five subjects had been clinically diagnosed as meeting the criteria for AD. MCI or control data were unavailable in this instance.

4.2.1.2 PET acquisition

The clinical trial involved several imaging sites (see section 4.1.1.2, page 89). The scans belonging to this subset were performed at one site. The PET data were acquired using Siemens ECAT Exact HR+. All PET data reconstructed using FBP. A reconstructed resolution of 5.3mm FWHM was measured by the method described in section 4.1.1.3, page 89.

4.2.1.3 Image processing

Region definition: The T1 MRI data for each subject was processed using FreeSurfer (FS) (see section 2.4.2.1, page 43). The FS parcellation was used to create a region mask consisting on 34 VOIs. The region definitions can be found in appendix D.3. The GM volumes consisted of the frontal, central, parietal, temporal, occipital, anterior cingulate, posterior cingulate, insula, precuneus, sub-cortical GM (caudate, putamen, thalamus, amygdala, pallidum), hippocampi and cerebellum. All GM regions, with the exception of the sub-cortical GM and cerebellar GM, were split into separate regions for the left and right structures. The frontal, parietal and temporal regions were further sub-divided into inferior, superior, and in the case of the temporal lobe, medial regions. The four other non-GM VOIs consist of SWM, cerebellar WM, the brain stem and ventricular and CSF space.

Registration: As previously described in section 4.1.1.4, the PET frames were rigidly registered together to form a sum image. The block matching technique of Ourselin et al. [2001] was applied to perform a rigid registration between the sum image and the T1 MRI. The registration of the first and second PET scans to the MRI were performed independently. Image processing and data analysis was subsequently performed in the MRI space.

Partial volume correction: RBV correction was performed on the PET data using the parcellated brain masks. A space-invariant Gaussian PSF of 5.3mm FWHM was assumed for the purposes of PVC when performing RBV.

PET normalisation: The 10% trimmed mean of activity in the brain stem was used to normalise the PV-corrected and uncorrected PET data (see section 4.3). Image normalisation was performed in MATLAB. The brain stem normalised images are referred to as SUVR images. However, it should be noted that in amyloid PET imaging the term SUVR typically relates to cerebellar GM normalisation. All image analysis was performed on the SUVR images.

4.2.1.4 Analysis

The SUVR images for both the RBV-corrected and uncorrected data were evaluated through regional analysis. 10% trimmed mean SUVR values were calculated in each ROI. The CoV_r was also computed to demonstrate the effects of PVC in terms of regional variance. Results were reported for the precuneus, SWM and a composite cortical region comprising of frontal, central, parietal and temporal regions. The average SUVR across the group, for the two scans, was calculated for each ROI. The same calculation was also made for the CoV_r. The reproducibility of the two scans per subject was assessed through a repeated measures analysis of variance (ANOVA) on the SUVR values of the composite region. All analysis was performed in MATLAB.

4.2.2 Results

Large increases in cortical SUVR values were observed after RBV correction was applied. An average increase of 27.6% across the group, for the first scan (+28.3% for the second scan), was observed in the composite cortical region after RBV correction (figure 4.3). All cortical GM regions exhibited an increase in the average SUVR (table 4.7) after RBV correction. The SUVR increases observed in the precuneus can be seen in figure 4.4. Reductions in SWM SUVR values after RBV correction were also found (figure 4.5). The percentage difference (table 4.7) in SUVR values between scans for all regions was less than ±3% when uncorrected for PVEs and less than ±5% after RBV correction.

Reductions in the CoV_r after RBV correction, were observed in the majority of cortical GM regions (table 4.8). The largest average reduction (-25%) after PVC, when averaging CoV_r across the two scans, was observed in the left central region. As well as reductions after PVC in the subject CoV_r, the SD of the group CoV_r also tends to decrease and is consistent across scans. This behaviour can be observed in figures 4.3, 4.4 and 4.5 for the composite region, precuneus and SWM respectively.

A repeated measures ANOVA was performed between the SUVR values of the two scans in the compostite cortical region. The null hypothesis was that SUVR values did not differ significantly between scans at the $p < 0.05$ threshold. For the uncorrected data the result was not significant, $F(1,4) = 0.92, p = 0.39$. The RBV-corrected were also tested and did not reach significance: $F(1,4) = 0.81, p = 0.42$. Therefore, the null hypothesis could not be rejected, suggesting that the SUVR values are reproducible and that RBV correction does not degrade the reproducibility.

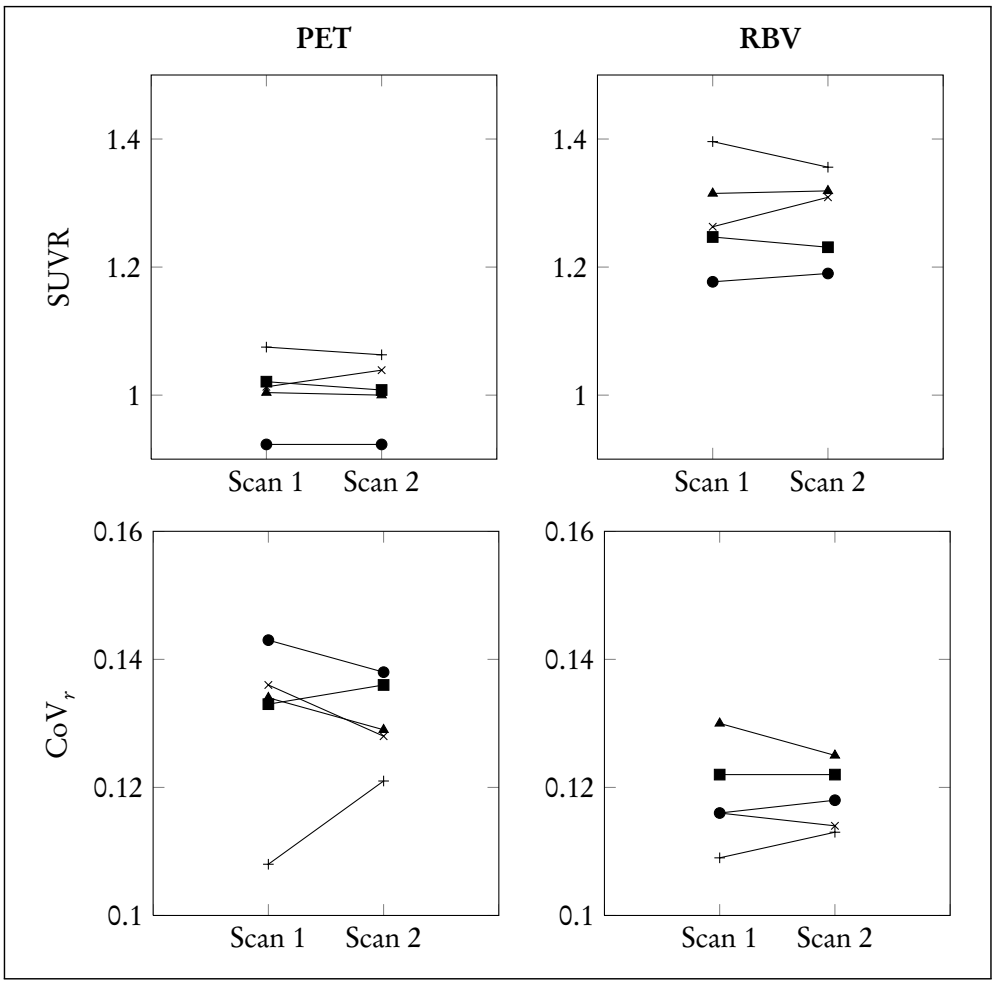


Figure 4.3: Graphs of SUVR and CoV$_r$ in composite cortical region, for both scans, without PVC (left column) and with after RBV correction (right column). SUVR is shown in the top row and CoV$_r$ in the bottom row.

4.2.3 Discussion

This study has investigated the reproducibility of quantitative measurements with [^{18}F]flutemetamol. The objective was to evaluate whether [^{18}F]flutemetamol quantification was reproducible without PVC and when applying RBV correction. The reproducibility of this cohort has already been assessed by Vandenberghe et al. [2010b] without PVC. How-

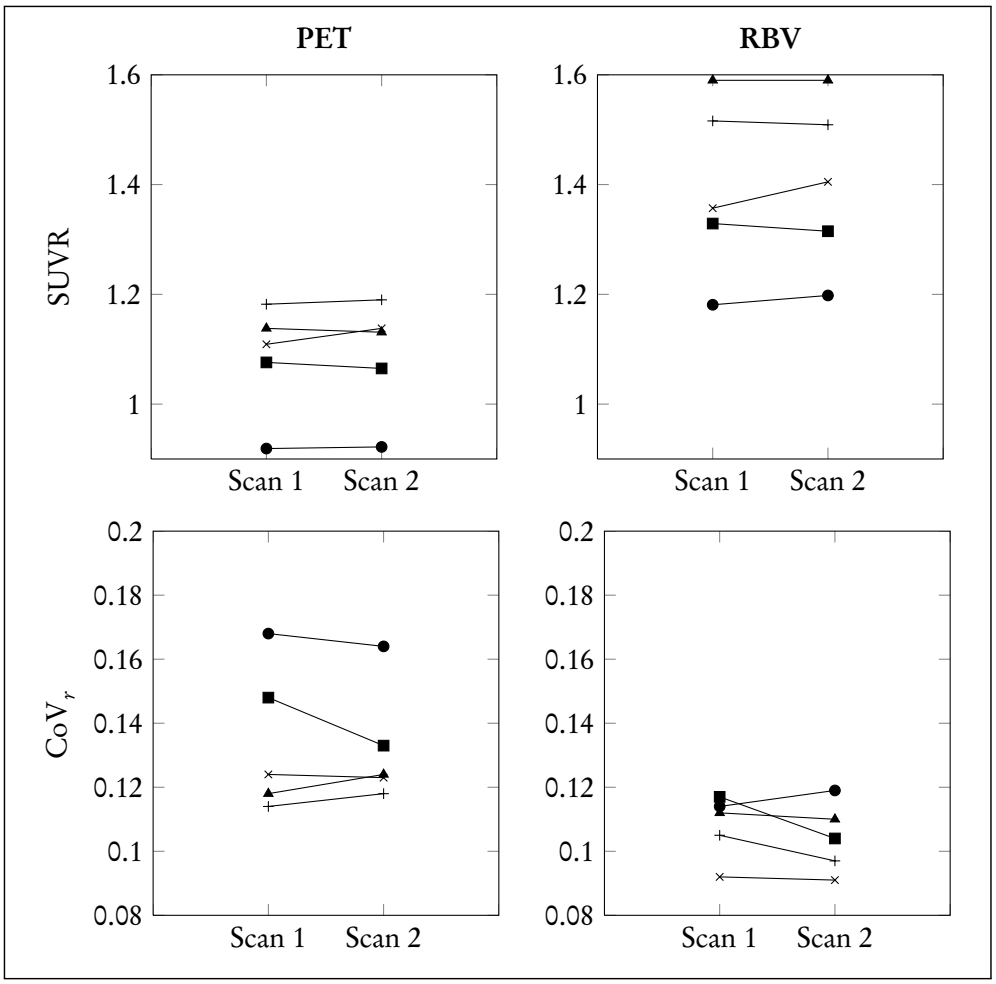


Figure 4.4: Graphs of SUVR and CoV$_r$ in the precuneus (left and right), for both scans, without PVC (left column) and with after RBV correction (right column). SUVR is shown in the top row and CoV$_r$ in the bottom row.

ever, the comparison reported here was specifically to evaluate the effects of PVC on the reproducibility of [^{18}F]flutemetamol.

The results of RBV correction were first compared against the uncorrected data. As expected, large increases in cortical SUVR values were observed when RBV correction was applied. The increases were consistent across subject scans, with SUVR values being within 5% of each other across scans, for all regions. This is in agreement with the results of Vandenberghe et al. [2010b] and suggests that the [^{18}F]flutemetamol signal is very reproducible, with or without PVC. SWM SUVR reductions were also observed after PVC and, as previously discussed, are likely due to correction for GM spill-over effects. Again, the reductions in SWM were consitent between scans, before and after PVC.

CoV$_r$ was also assessed in the comparison of uncorrected and RBV-corrected data. While CoV$_r$ is of less interest in terms of reproducibility, as external factors such as injected dose and

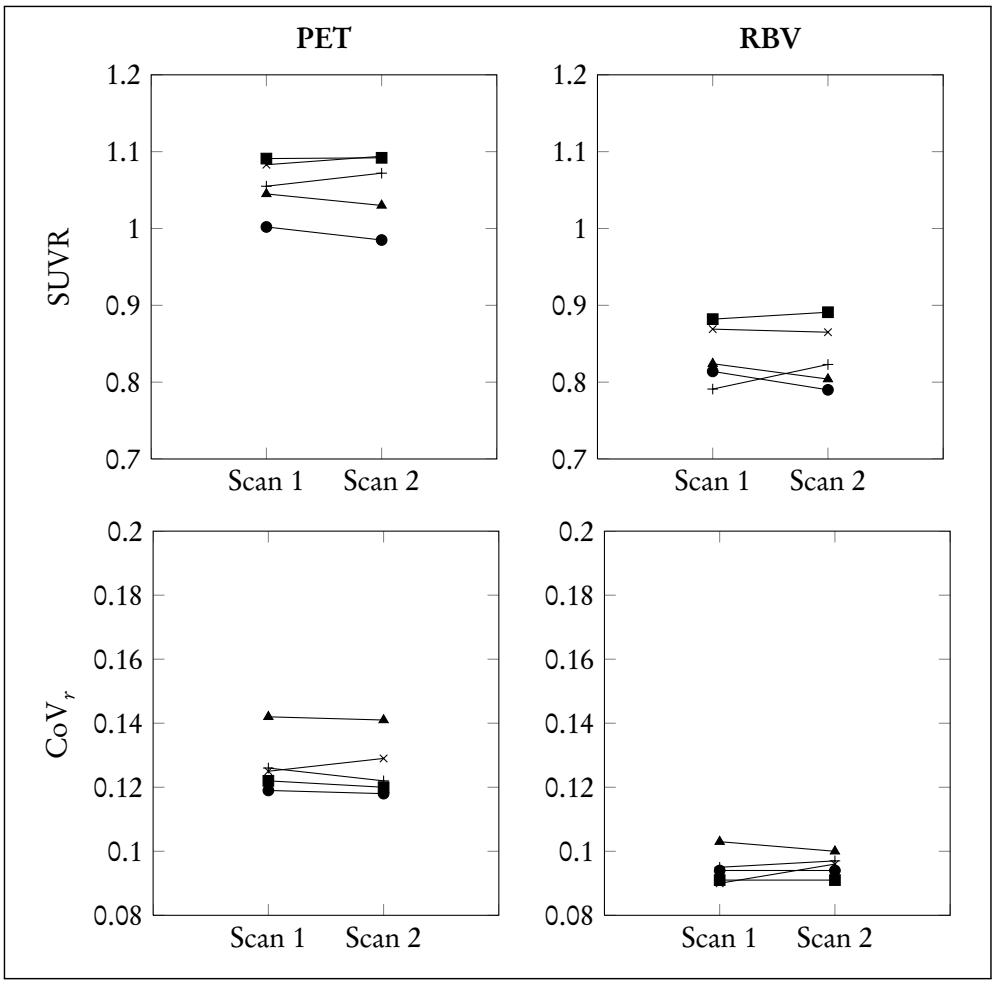


Figure 4.5: Graphs of SUVR and CoV$_r$ in the sub-cortical white matter (left and right), for both scans, without PVC (left column) and after RBV correction (right column). SUVR is shown in the top row and CoV$_r$ in the bottom row.

positioning, will affect the measurement, the values are reported in this study as they show that RBV correction results in consistent reductions in CoV$_r$. Most cortical GM regions exhibited marked reductions in CoV$_r$. However, for example, large increases (> 40%) in the insula were observed after PVC. This may be a genuine finding, or an indication that the parcellation is inappropriate at that point. The CoV$_r$ also increased in the hippocampus. This behaviour has been previously discussed and can be seen in table 4.6, page 93.

Repeated measures ANOVAs were performed between SUVR values of the composite cortical region for the uncorrected scan pairs and the RBV-corrected scan pairs. These were calculated to assess whether there was a significant difference between the scans of a subject. Neither the uncorrected nor RBV-corrected data exhibited a significant difference across scans. This suggests both that [^{18}F]flutemetamol SUVR values are reproducible and that RBV correction does not degrade the reproducibility of cortical SUVR values.

In summary, this study has shown that [^{18}F]flutemetamol uptake is reproducible across scans and agrees with the results of Vandenberghe et al. [2010b] who analysed the same cohort, albeit it using a different methodology. RBV has also been shown to reduce regional variability, while maintaining the reproducibility of the SUVR values. This suggests that PVC can and should be applied to longitudinal amyloid PET studies as it improves quantification.

Region	PET 1	PET 2	% diff.	RBV 1	RBV 2	% diff.
Left sup. frontal	1.05	1.06	1.02%	1.34	1.36	1.53%
Left inf. frontal	1.08	1.07	-0.75%	1.37	1.36	-0.78%
Right sup. frontal	1.08	1.07	-0.72%	1.39	1.38	-0.76%
Right inf. frontal	1.10	1.08	-1.71%	1.41	1.38	-2.06%
Left central	0.88	0.89	1.72%	1.09	1.12	2.87%
Right central	0.92	0.91	-1.09%	1.15	1.13	-1.29%
Left precuneus	1.09	1.10	0.25%	1.42	1.42	0.22%
Right precuneus	1.08	1.08	0.52%	1.37	1.38	1.06%
Left inf. parietal	1.00	1.01	1.63%	1.29	1.32	2.26%
Right inf. parietal	1.04	1.03	-0.87%	1.34	1.33	-1.03%
Left sup. parietal	0.97	0.98	1.32%	1.28	1.31	1.76%
Right sup. parietal	0.99	0.98	-0.41%	1.30	1.29	-0.42%
Left sup. temporal	1.01	1.01	0.97%	1.30	1.32	1.53%
Right sup. temporal	1.01	1.00	-1.89%	1.27	1.24	-2.29%
Left mid. temporal	1.00	1.03	2.71%	1.28	1.33	4.04%
Right mid. temporal	1.05	1.03	-1.73%	1.34	1.31	-2.12%
Left inf. temporal	0.92	0.91	-0.52%	1.09	1.08	-0.76%
Right inf. temporal	0.88	0.86	-1.49%	1.02	1.00	-1.97%
Left occipital	0.76	0.77	1.28%	0.83	0.85	2.06%
Right occipital	0.77	0.77	0.71%	0.84	0.85	1.29%
Left ant. cingulate	1.07	1.07	0.20%	1.29	1.30	0.75%
Right ant. cingulate	1.04	1.04	0.12%	1.25	1.25	0.32%
Left post. cingulate	1.06	1.03	-2.33%	1.28	1.24	-2.94%
Right post. cingulate	1.03	1.03	-0.65%	1.24	1.23	-0.60%
Left insula	1.00	1.01	0.39%	1.06	1.07	0.66%
Right insula	1.04	1.01	-2.11%	1.11	1.08	-2.56%
Deep GM	0.96	0.96	0.56%	0.90	0.90	1.05%
Left hippocampus	0.69	0.71	2.19%	0.63	0.66	4.73%
Right hippocampus	0.68	0.68	-0.24%	0.62	0.62	0.07%
Sub-cortical WM	1.06	1.05	-0.07%	0.84	0.83	-0.17%
Cerebellar WM	0.96	0.96	-0.24%	0.95	0.95	-0.35%
Cerebellar GM	0.48	0.48	0.46%	0.42	0.43	0.81%
CSF	0.40	0.40	-0.58%	0.16	0.15	-2.21%
Brain stem	1.00	1.00	0.00%	1.00	1.00	0.00%

Table 4.7: SUVR values for each ROI, without PVC (second and third columns) and after RBV correction (fourth and fifth column). The percentage difference (% diff.) is given between the scans in the third column for uncorrected data and the sixth column after RBV correction. The average values were calculated across across the group for each scan.

Region	PET 1	PET 2	% diff.	RBV 1	RBV 2	% diff.
Left sup. frontal	0.14	0.13	-2.13%	0.11	0.11	1.79%
Left inf. frontal	0.12	0.11	-4.13%	0.11	0.11	-0.66%
Right sup. frontal	0.14	0.13	-2.60%	0.12	0.12	-0.43%
Right inf. frontal	0.11	0.11	0.99%	0.12	0.11	-1.95%
Left central	0.14	0.14	0.38%	0.11	0.10	-0.84%
Right central	0.14	0.13	-2.18%	0.11	0.11	-2.59%
Left precuneus	0.12	0.12	3.04%	0.10	0.10	0.39%
Right precuneus	0.15	0.14	-5.22%	0.12	0.11	-7.07%
Left inf. parietal	0.14	0.14	4.02%	0.11	0.12	3.63%
Right inf. parietal	0.12	0.12	0.90%	0.11	0.12	1.78%
Left sup. parietal	0.14	0.15	3.41%	0.11	0.11	2.19%
Right sup. parietal	0.14	0.14	0.03%	0.12	0.12	1.94%
Left sup. temporal	0.12	0.11	-2.06%	0.11	0.10	-1.23%
Right sup. temporal	0.11	0.11	-5.42%	0.11	0.11	-2.96%
Left mid. temporal	0.11	0.11	-2.32%	0.11	0.11	-0.78%
Right mid. temporal	0.09	0.09	-0.23%	0.12	0.12	-2.41%
Left inf. temporal	0.16	0.17	5.53%	0.16	0.17	5.26%
Right inf. temporal	0.18	0.18	1.19%	0.18	0.18	-0.74%
Left occipital	0.20	0.20	-0.25%	0.17	0.17	-0.72%
Right occipital	0.19	0.19	-1.19%	0.18	0.17	-2.33%
Left ant. cingulate	0.10	0.09	-5.49%	0.11	0.11	2.50%
Right ant. cingulate	0.12	0.12	-0.46%	0.11	0.11	-3.22%
Left post. cingulate	0.09	0.09	-0.61%	0.10	0.11	10.67%
Right post. cingulate	0.12	0.11	-5.18%	0.11	0.11	2.79%
Left insula	0.09	0.09	-0.01%	0.13	0.13	-0.75%
Right insula	0.08	0.08	8.15%	0.13	0.13	1.02%
Deep GM	0.18	0.18	-1.22%	0.14	0.13	-3.07%
Left hippocampus	0.12	0.12	-3.45%	0.16	0.15	-6.86%
Right hippocampus	0.13	0.13	-3.76%	0.18	0.17	-5.32%
Sub-cortical WM	0.13	0.13	-0.53%	0.09	0.10	0.96%
Cerebellar WM	0.20	0.19	-2.69%	0.13	0.12	-5.83%
Cerebellar GM	0.27	0.26	-2.24%	0.18	0.18	-2.73%
CSF	0.38	0.38	1.59%	0.23	0.23	3.71%
Brain stem	0.16	0.16	-1.38%	0.12	0.12	-3.80%

Table 4.8: CoV_r values for each ROI, without PVC (second and third columns) and after RBV correction (fourth and fifth column). The percentage difference (% diff.) is given between the scans in the third column for uncorrected data and the sixth column after RBV correction. The average values were calculated across across the group for each scan.

4.3 Amyloid PET normalisation

Amyloid PET scans are typically normalised using the mean value found in the cerebellar GM. The region is thought to be relatively free of fibrillar amyloid plaques, even during the late stages of AD [Klunk et al., 2004, Lopresti et al., 2005]. Amyloid plaques have however been observed in the cerebellar GM for certain variant forms of the disease [Verkkoniemi et al., 2001], leading Koivunen et al. [2008] to report SUVR values using the cerebellar GM and the pons as reference regions. Wong et al. [2010] chose WM reference regions to perform the normalisation of [^{18}F]FDDNP images.

A comparison between [^{11}C]PIB and [^{18}F]flutemetamol has been previously reported by Vandenberghe et al. [2010b]. This section reports an investigation into alternative reference regions for image normalisation using the [^{11}C]PIB and [^{18}F]flutemetamol data collected during the same clinical trial. The interaction of PVC with the choice of reference region is also assessed. In addition, WM regions are compared with and without PVC for both tracers. The regional analysis was undertaken to evaluate as to whether WM uptake is homogeneous or not as this is an important assumption of some PVC techniques.

The aims of this study are to compare [^{11}C]PIB and [^{18}F]flutemetamol, evaluate WM regions and assess the performance of alternative reference regions, while demonstrating these relationships with appropriate PVC.

4.3.1 Materials and methods

4.3.1.1 Subjects

This study uses a subset of the subjects belonging to the clinical cohort described in section 4.1.1.1 (page 89). The WM analysis reported in section 4.3.2.1 (page 109) is performed on [^{18}F]flutemetamol data alone. All subjects in the AD ($n = 26$) and HC ($n = 24$) group are included in the WM analysis.

As part of the same clinical trial, all MCI subjects ($n = 20$) and the majority of AD subjects ($n = 20$) underwent an additional [^{11}C]PIB PET scan [Vandenberghe et al., 2010b]. Therefore, the comparison between the two tracers and the evaluation of alternative reference regions is restricted to the MCI and AD subject group.

4.3.1.2 Image processing

All images were processed in the same manner as described in sections 4.1.1.2 to 4.1.1.4. A greater number of regions were included in the brain mask to reduce potential errors due to lateralisation of uptake. The number of GM regions was increased from 10 to 20, bringing the total number of regions to 24. The 20 regions represent the same anatomical structures as

the set described in section 4.1.1.4, but sub-divided into the left and right hemispheres. The full list of FS labels are provided in appendix D.2.

4.3.1.3 Partial volume correction

The aim of this study is to investigate WM regions, alternative reference regions and evaluate the interaction of PVC. It is necessary to apply a PVC technique that can accurately correct multiple regions. The study in the previous section (4.1) showed that the performance of RBV was superior to that of the mMG. The mMG is also restricted to correcting the GM compartment only. Therefore, RBV was the only PVC applied during this study.

4.3.1.4 Data analysis

SUVR values were calculated for all segmented regions using a 10% trimmed mean value. Left and right regions were combined for the purposes of analysis in order to compare the results with that of Vandenberghe et al. [2010a]. The SD across subjects in a group was also calculated. The [^{18}F]flutemetamol WM analysis was carried out using cerebellar GM normalisation. The SWM and cerebellar white matter (CWM) were evaluated through SUVR analysis and also by linear regression.

Both SUVR and linear regression analysis were performed for the [^{11}C]PIB \ [^{18}F]flutemetamol comparison. The MCI group was sub-divided into amyloid-positive ($n = 9$) and amyloid-negative ($n = 11$) groups. The division was carried out using the SUVR threshold as described in section 4.1.2.2 (page 93). All images were normalised and analysed using MATLAB R2010b (Mathworks, Natick, M.A., USA). Statistical analysis was performed in MATLAB and G*Power. Where the MDD was calculated, a power of 90% and $p < 0.01$ was assumed. The linear regression analysis was performed between the [^{11}C]PIB and [^{18}F]flutemetamol SUVR for all 40 subjects, for each of the three reference regions under evaluation. The assessed regions were the cerebellar GM, SWM and brain stem.

In the tracer comparison, the differences in CWM and SWM were reported in terms of a cerebellar WM to sub-cortical WM ratio (WMR), defined as:

$$\text{WMR} = \frac{\text{cerebellar WM}}{\text{sub-cortical WM}} \tag{4.1}$$

The WMR was calculated using the non-normalised images of each subject, with and without PVC, for both tracers. The mean WMR and SD across subjects within a group was calculated. The results of the linear regression were reported in terms of the slope of the line (m), correlation coefficient (r) and the normalised standard error of the estimate (NSEE) was calculated

using the sum of square errors normalised to the mean of the data. The $NSEE$ is defined as:

$$NSEE = \frac{1}{x}\left(\frac{\sqrt{\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}}{N-2}\right) \tag{4.2}$$

where y_i is the [^{11}C]PIB value, $\hat{y}_i$ is the predicted value, N is the number of points ($N = 40$) and x is the mean of the [^{18}F]flutemetamol values.

4.3.2 Results

4.3.2.1 [^{18}F]flutemetamol white matter analysis

SUVR analysis was performed on the WM regions of the SWM and CWM. As previously described in section 4.1.1.1 (page 89), the HC group was originally sub-divided into elderly and young subjects. Wilcoxon rank-sum tests were performed between the eHC and yHC groups to establish if there was any WM differences within the control group. Tests were performed between the groups for both the SWM and CWM. No significant difference at the $p < 0.05$ threshold was found for either region (SWM: MDD =0.11; CWM: MDD =0.07). As a result, the controls were considered as a single group for the WM analysis.

Trimmed mean SUVR values were calculated in both WM regions for the AD and HC groups. These values can be seen in table 4.9. The CWM values were consistent across the groups before and after correction. Wilcoxon signed-rank tests were highly significant ($p < 0.001$) between the regions without PVC. SWM exhibited higher SUVR values than CWM in the AD subjects. This pattern was reversed in control subjects.

When RBV correction was applied, both subject groups had a higher CWM than SWM SUVR. No significant ($p > 0.05$) difference was found between the AD and HC group, for either region, after PVC. However, Wilcoxon signed-rank tests between the regions of both subject groups exhibited highly significant differences ($p < 0.001$).

		Uncorrected	RBV
Cerebellar WM	AD	1.96 ± 0.16	2.30 ± 0.30
	HC	1.98 ± 0.13	2.35 ± 0.23
Sub-cortical WM	AD	2.17 ± 0.22	2.01 ± 0.32
	HC	1.78 ± 0.09	2.01 ± 0.22

Table 4.9: White matter SUVR values [mean ± SD] in AD and control groups, with and without RBV PV-correction.

In addition to an SUVR analysis of the WM regions, a linear regression analysis was performed for RBV-corrected and the uncorrected data. The results can be seen in figure 4.6.

A much stronger relationship between SWM and CWM was observed after RBV correction. The correlation coefficient increased from 0.51 to 0.82 for the HC group and from 0.37 to 0.89 for the AD group. The changes in correlation coefficient were compared using the $\bar{Z}_2^*$ test for differences between dependent correlations proposed by Steiger [1980]. The increases in correlation coefficient after PVC were highly significant ($p < 0.001$) for both groups. Before PVC, the HC and AD groups appear distinct. After RBV correction, this distinction disappears, and suggests that disease has no influence on the true signal in WM. The higher SWM values in the AD group are due to PVEs as they are not observed in the PV-corrected data.

4.3.2.2 [^{11}C]PIB and [^{18}F]flutemetamol comparison

All cortical GM regions exhibited strong agreement during linear regression analysis between [^{11}C]PIB and [^{18}F]flutemetamol SUVR values. Figure 4.7 shows the linear regression analysis of the composite cortical region, for each reference region, with and without PVC. The agreement was observed for each of the three reference regions, irrespective of whether PVC had been applied. In WM regions, the weak agreement between the two tracers was further reduced after PVC (see figure A.1, page 167). The results of the linear regression analysis are shown in table 4.10. The RBV-corrected and uncorrected SUVR images of both tracers from an AD subject are shown in figure 4.8.

In terms of reference regions, the brain stem produced the lowest $NSEE$ values (table 4.11) after PVC, suggesting that of the three regions, brain stem normalisation resulted in the strongest agreement between the tracers. Assessing the slopes of the lines, both cerebellar GM and brain stem normalisation were consistent before and after RBV correction. SWM normalisation exhibited greater changes in slope when comparing the uncorrected with the RBV-corrected data.

Further F-tests were performed on the $NSEE$ values. The $NSEE^2$ values were evaluated for each pair of reference regions (cerebellar GM vs. SWM, cerebellar GM vs. brain stem and SWM vs. brain stem), with and without PVC. For the uncorrected data, comparison of the cerebellar WM to both the brain stem ($F_{(1,39)} = 6.61$) and SWM ($F_{(1,39)} = 12.96$) achieved significance at the $p < 0.05$ threshold. Given the lower $NSEE$ values of the brain stem and SWM, this suggests that these reference regions result in stronger agreement between the two tracers. The comparison of brain stem to SWM did not reach significance. When assessing the RBV-corrected data, only the comparison between cerebellar GM and the brain stem reached significance ($F_{(1,39)} = 5.76$). This suggests that PVC makes the three reference regions become more similar to each other, at least in terms of their spread about the regression line. The

'dissimilarity' in the PV-uncorrected reference regions may therefore be logically attributed to PVEs.

4.3.2.3 Effects of PVC on the white matter ratio

Further to the initial investigation of WM regions using [^{18}F]flutemetamol (section 4.3.2.1). The CWM and SWM regions were analysed for the MCI and AD subjects that had both [^{11}C]PIB and [^{18}F]flutemetamol scans. The WM regions were analysed in terms of WMR values. The results for both tracers, with and without PVC, can be seen in figure 4.9 for the AD, amyloid-positive and amyloid-negative groups.

For there to be no difference between the CWM and SWM across the cohort, a mean WMR = 1 would be expected. This null hypothesis was evaluated using Wilcoxon signed-rank tests for each clinical group, for each tracer, with and without PVC. The hypothesis was rejected in all tests and was highly significant ($p < 0.01$). The AD and amyloid-positive group exhibited WMR < 1, with all amyloid-negative subjects displaying a WMR > 1. When RBV-correction was applied, all groups had an average WMR > 1.

Wilcoxon rank-sum tests were used to compare the clinical groups to each other. A significant ($p < 0.01$) difference between the amyloid-negative group and the amyloid-positive and AD groups was found in both tracers when PVC had not been applied. When RBV correction was applied, this difference remained significant. This was further investigated using the PV-corrected SUVR values normalised to the brain stem, rather than the WMR. No significant differences were found in the SWM for either tracer. Uptake in the CWM was found to be higher in the amyloid-negative group. The brain stem normalised CWM values can be seen in figure 4.10. Wilcoxon rank-sum tests were performed using the cerebellar WM SUVR values. Tests between amyloid-positive and amyloid-negative subjects were not significant at the $p < 0.01$ threshold ([^{11}C]PIB: $p = 0.095$, MDD=0.09; [^{18}F]flutemetamol: $p = 0.015$, MDD=0.09). Significance was also not found between the amyloid-negative and AD subjects ([^{11}C]PIB: $p = 0.016$, MDD=0.07; [^{18}F]flutemetamol: $p = 0.069$, MDD=0.09).

The WMR was also compared between tracers for each group, with and without PVC. Wilcoxon signed-rank tests were performed, assuming a threshold of ($p < 0.01$). Without PVC the amyloid-negative group showed no significant difference ($p = 0.15$, MDD=0.08). However, significant differences were observed for both the amyloid-positive and AD group (amyloid-positive: $p = 0.0007$; AD: $p = 0.002$). When RBV correction was applied, no significant difference between tracer was observed for any subject group (amyloid-negative: $p = 0.47$, MDD=0.09; amyloid-positive: $p = 0.26$, MDD=0.10; AD: $p = 0.024$, MDD=0.06).

	Reference region											
	Cerebellar GM				Sub-cortical WM				Brain stem			
	Uncorr.		RBV		Uncorr.		RBV		Uncorr.		RBV	
	m	r	m	r	m	r	m	r	m	r	m	r
Composite	1.00	0.90	1.01	0.93	1.13	0.98	1.35	0.98	1.49	0.99	1.50	0.99
Frontal	1.02	0.92	1.01	0.93	1.14	0.98	1.35	0.98	1.48	0.99	1.49	0.99
Parietal	1.06	0.92	1.08	0.95	1.09	0.98	1.32	0.98	1.48	0.99	1.49	0.99
Temporal	0.99	0.92	0.98	0.93	1.01	0.96	1.28	0.97	1.46	0.99	1.46	0.99
Post. cingulate	0.98	0.88	0.98	0.90	1.15	0.98	1.37	0.98	1.49	0.99	1.51	0.99
Ant. cingulate	0.93	0.88	0.91	0.90	1.20	0.99	1.40	0.99	1.49	0.99	1.49	0.99
Occipital	1.13	0.94	1.13	0.98	0.97	0.90	1.24	0.95	1.52	0.97	1.46	0.97
Deep GM	0.82	0.78	0.86	0.82	1.00	0.87	1.25	0.92	1.53	0.98	1.54	0.97
Hippocampus	0.64	0.78	0.66	0.83	1.06	0.81	0.83	0.63	1.35	0.90	1.16	0.76
Cerebellar WM	0.33	0.47	0.17	0.29	1.13	0.97	0.79	0.77	0.77	0.80	0.54	0.71
Brain stem	0.41	0.49	0.29	0.41	1.16	0.97	0.55	0.65	-	-	-	-
Sub-cortical WM	0.53	0.56	0.20	0.36	-	-	-	-	1.45	0.96	0.53	0.64
Cerebellar GM	-	-	-	-	0.87	0.62	0.37	0.35	0.68	0.45	0.51	0.35

Table 4.10: Slope (m) of linear fit and correlation coefficient (r) of [^{11}C]PIB and [^{18}F]flutemetamol SUVR comparison using three reference regions, with (RBV) and without (Uncorr.) partial volume correction.

	Reference region					
	Cerebellar GM		Sub-cortical WM		Brain stem	
	Uncorr.	RBV	Uncorr.	RBV	Uncorr.	RBV
Composite	0.018	0.024	0.005	0.017	0.007	0.010
Frontal	0.019	0.025	0.007	0.021	0.008	0.013
Parietal	0.018	0.023	0.007	0.020	0.008	0.014
Temporal	0.016	0.021	0.007	0.020	0.008	0.013
Post. cingulate	0.020	0.027	0.006	0.015	0.008	0.011
Ant. cingulate	0.021	0.028	0.005	0.015	0.009	0.011
Occipital	0.012	0.015	0.011	0.033	0.013	0.029
Deep GM	0.015	0.020	0.007	0.018	0.007	0.011
Hippocampi	0.011	0.012	0.011	0.023	0.012	0.020
Cerebellar WM	0.009	0.012	0.006	0.007	0.004	0.005
Brain stem	0.010	0.012	0.006	0.008	-	-
Sub-cortical WM	0.016	0.014	-	-	0.009	0.008
Cerebellar GM	-	-	0.021	0.027	0.019	0.025

Table 4.11: The normalised standard error of the estimate ($NSEE$) for the [^{11}C]PIB and [^{18}F]flutemetamol SUVR comparison using three reference regions, with (RBV) and without (Uncorr.) partial volume correction.

4.3.3 Discussion

4.3.3.1 [^{18}F]flutemetamol white matter analysis

The WM signal observed in amyloid PET images is thought to be non-specific uptake, with no disease-related component [Vandenberghe et al., 2010b]. This study has investigated whether CWM uptake is the same as SWM. The comparison was carried out using the AD and HC groups of the [^{18}F]flutemetamol dataset. CWM activity was consistent across the subject groups before and after PVC. The increase in CWM after RBV correction is likely to be a result of correction for spill-over into the background. When PVC has not been applied, the SWM uptake in AD subjects (see table 4.9) is higher than in the HC group. After correction this difference is no longer observed and the SWM of AD subjects actually reduces. Given that the difference between groups ceases to exist after PVC, it is likely to have been caused by PVEs due to spill-over from 'hot' GM regions.

Comparisons between the AD and HC groups for both WM regions failed to reach significance after PVC. The MDD was calculated as 0.22 for both comparisons, with a difference in means between the groups of 0.05 for the CWM and 0.00 for the SWM. Given that a larger difference (0.39) was observed in the uncorrected SWM, these negative results add further weight to the belief that WM uptake has no disease component associated with it.

The difference between the CWM and SWM is of particular interest in terms of PVC.

Techniques such as the MG approaches, operate on the assumption that WM can be considered to be a uniform region. The MG method would therefore estimate activity based on the average of the CWM and SWM values. As previously discussed, in subjects with high amyloid burden, cortical GM spill-over appears to increase SWM activity. This over-estimation of SWM would lead to lower cortical GM values. As a consequence of this, CWM is also over-estimated, resulting in a negatively-biased cerebellar GM region. When used as a reference region, the cerebellar GM would have the effect of increasing cortical GM values. Therefore, the differences between WM regions may cause both increases and decreases in cortical GM values of subjects with high amyloid burden. The over-estimation in SWM depends on the amount of spill-over from cortical GM. This means the degree of over-estimation is dependent on the ratio of cortical GM to WM rather that the magnitude of the cortical GM. The greater the signal ratio between the two, the larger the over-estimation will be.

In subjects with low amyloid burden, cortical GM spill-over does not affect SWM. However, when applying MG in this situation, the CWM results in an average of the two regions, which causes an over-estimation of the SWM. The CWM then becomes under-estimated, creating a positively-biased cerbellar GM. As a reference region, the use of the cerebellar GM would result in reduction in cortical GM values. Further phantom experiments (as described in section 3.1.1.1) were carried out to test this hypothesis. Variability in the WM was applied according to the regional differences observed in this study. The results can be seen in appendix A (figures A.2 and A.3). The mMG method resulted in further under-estimations in cortical GM regions of between 3% and 5%. A positive bias (+6%) in cerebellar GM can also be observed.

4.3.3.2 [^{11}C]PIB and [^{18}F]flutemetamol comparison

The findings of WM variability prompted the further investigations in this study. As the cerebellum is susceptible to biases when inappropriate PVC techniques are applied, we evaluated alternative reference regions. The comparison between [^{18}F]flutemetamol and [^{11}C]PIB was carried out to reproduce the WM findings in another tracer, evaluate potential reference regions and also investigate the effects of PVC. The cerebellar GM, SWM and brain stem were evaluated as reference regions. The brain stem was selected due to the ease of delineation compared to the pons. The results of using the pons as a reference region from a previously published study [Vandenberghe et al., 2010a] were highly comparable to the results of this study, suggesting that both regions are similar, with non-specific uptake.

Linear regression analysis was performed between [^{11}C]PIB and [^{18}F]flutemetamol. These tracers showed strong agreement in the cortical GM regions for all three reference re-

gions. In the PV-corrected data, brain stem normalisation produced the lowest $NSEE$ values. Stronger agreement between tracers was observed when using SWM compared to cerebellar GM. However, the slopes change after PVC, when using SWM as the reference region. This is to be expected, as SWM is severely affected by PVEs, especially in subjects with high amyloid burden.

In order for SWM to be applied as the reference region, either PVC would have to be applied or a reference volume would have to be defined such that it was not affected by PVEs. For example, a spherical or ellipsoidal volume in the centrum semiovale could be defined for this purpose, assuming the boundaries of the volume are at a distance of at least three times the FWHM of the scanner PSF from the cortex and ventricles. When applying either brain stem or SWM normalisation, the slopes of the line deviate from unity ($m \neq 1$). It would therefore be necessary to find tracer-specific correction factors to ensure that data from multi-centre, multi-tracer trials were comparable.

Weaker agreement between the two tracers was observed in WM regions and reduced further after PVC was applied. This observation is understandable if WM uptake is really non-specific. The higher WM signal seen with [^{18}F]flutemetamol compared to [^{11}C]PIB is believed to be a result of differing washout characteristics [Vandenberghe et al., 2010b]. The initial correlations in WM between the two tracers could be a GM spill-over PVE.

The previously discussed differences between CWM and the SWM were further evaluated in the subjects that had both [^{18}F]flutemetamol and [^{11}C]PIB. The finding that CWM is hotter than SWM was confirmed for both tracers. [^{18}F]flutemetamol is known to have a higher WM signal than [^{11}C]PIB, but the differences between the regions were observed irrespective of overall WM differences between the tracers. The ratio of the CWM to the SWM does not appear to have any clinical relevance in itself. The inter-regional differences become important when PVC is to be applied. Assuming that WM is uniform when performing PVC could introduce biases in the quantitative analysis of regions that are clinically relevant, such as cortical GM.

4.3.3.3 Effects of PVC on the white matter ratio

The inter-regional differences were analysed using the WMR. Wilcoxon signed rank tests on the WMR were highly significant, irrespective of clinical group or tracer type. This result suggests that the uptake in the CWM is different to that of the SWM. As can be seen in figure 4.9, before PVC the AD and amyloid-positive subjects have higher SWM uptake than CWM uptake (WMR < 1) for both [^{11}C]PIB and [^{18}F]flutemetamol, with amyloid-negative subjects exhibiting the opposite characteristics (WMR > 1). When RBV correction was applied, all

subject groups had a WMR > 1. The same observation was made between AD and HC subjects in [^{18}F]flutemetamol alone and has now also been observed in the [^{11}C]PIB dataset.

The subject groups were compared using unpaired Wilcoxon rank-sum tests to evaluate whether there existed a between-group difference in WMR values. A significant difference was observed between the amyloid-negative group and the other subject groups. This remained significant at the $p < 0.01$ threshold after PVC. Further investigation was undertaken using the brain stem normalised SUVR values instead of the WMR (figure 4.10). When RBV correction was applied, no significant differences were found between groups for either the CWM or SWM SUVR values. However, the CWM SUVR was higher in the amyloid-negative group for both [^{18}F]flutemetamol and [^{11}C]PIB. While the CWM SUVR did not reach significance at the $p < 0.01$ threshold, three of the four comparisons were less than 0.05. It is this higher CWM value in amyloid-negative subjects that gave rise to the group difference in WMR. As the SWM differences do not remain significant after PVC has been applied, it suggests that the (uncorrected) SWM differences are due to PVEs from GM. As SWM appears consistent across subject groups it suggests that SWM uptake is non-specific and does not reveal any information about disease state.

All of the statistical comparisons of the CWM and SWM SUVR values failed to reach significance. The MDD was calculated for both regions and was found to be between 0.07 and 0.10 for all tests. The difference between group means was less than the MDD for each test. However, some group mean differences were close to the MDD. For example, when comparing the CWM between the [^{11}C]PIB SUVR values of the amyloid-negative and AD groups, the mean difference was 0.06, with a MDD of 0.07. There could be subtle differences between clinical groups in WM regions, although the small sample sizes of the amyloid-negative and -positive group have prevented this from being detected.

A comparison of the WMR between [^{18}F]flutemetamol and [^{11}C]PIB, across groups, was carried out. Paired Wilcoxon signed-rank tests showed significant differences between [^{18}F]flutemetamol and [^{11}C]PIB in amyloid-positive and AD subjects when PVC was not applied. The [^{11}C]PIB WMR tended to be lower than [^{18}F]flutemetamol. No significant difference in WMR was observed between tracers for any subject group once RBV correction was performed. The WMR increase after RBV correction is a result of PVC for cortical GM spill-over.

The MDD was calculated for the negative statistical comparisons between the two tracers. The amyloid-negative group exhibited a mean difference across tracer of 0.02, with a MDD of 0.09. The amyloid-positive also failed to reach significance, with a mean difference

of 0.04 and a MDD of 0.10. However, in the AD group, a mean difference between tracer of 0.05 was observed (MDD=0.06). While this negative result may be due to sample size, it is unlikely that there is a difference between [^{11}C]PIB and [^{18}F]flutemetamol for the AD subjects alone.

4.3.4 Conclusion

This study has shown that differences in activity exist between the CWM and SWM. The difference was first identified in AD and HC subjects using [^{18}F]flutemetamol. The finding was subsequently confirmed in a [^{11}C]PIB dataset. These differences have implications for the application of PVC in amyloid PET imaging. RBV correction was used to show that SWM uptake has no disease-specific component in either [^{11}C]PIB or [^{18}F]flutemetamol. After correction, CWM uptake was shown to be consistently higher than the SWM, irrespective of clinical group, or tracer.

PVC techniques which assume uniformity in WM, such as the MG and mMG, may induce biases due to erroneous estimation of WM used during the PVC. A method that corrects for spill-over from multiple regions is necessary to accurately correct amyloid PET images as they contain both regional GM and WM variability. In particular, the effects of inaccurate correction in the CWM are likely to affect the accuracy of cerebellar GM as a reference region.

An investigation of alternative reference regions has been carried out in this study, evaluated through a comparison of [^{18}F]flutemetamol and [^{11}C]PIB. Brain stem normalisation in combination with PVC was found to produce the strongest agreement between the tracers for cortical GM regions. In terms of tracer agreement, SWM was also superior to the cerebellar GM, although this can be partially attributed to PVEs from cortical GM spill-over.

In summary, WM variability in amyloid PET images mandates that a PVC technique, such as RBV, should be applied. This is particularly important if a target reference region could be affected by PVC-induced biases. Brain stem appears preferable to the cerebellar GM for the purposes of normalisation as it has been shown to improve agreement between [^{11}C]PIB and [^{18}F]flutemetamol, maintains clinical group separation and is arguably easier to segment than the cerebellar GM.

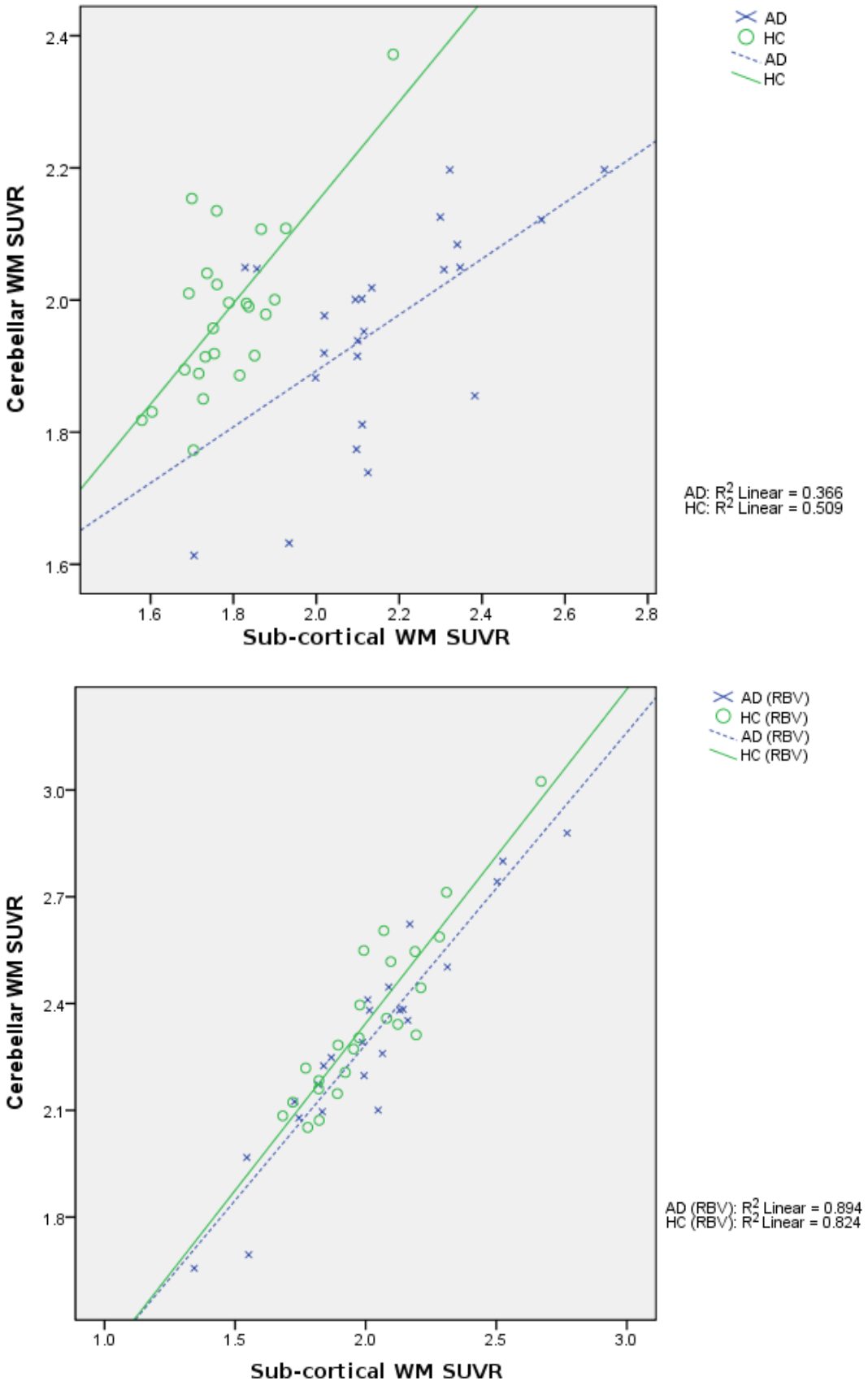


Figure 4.6: Scatter plots of sub-cortical WM against cerebellar WM without PV-correction (top) and after RBV correction (bottom). Images were normalised to the cerebellar GM. Controls are marked as green circles and AD subjects as blue crosses. The solid green lines represent the line of best fit in the control group, with the dashed blue line representing the AD group.

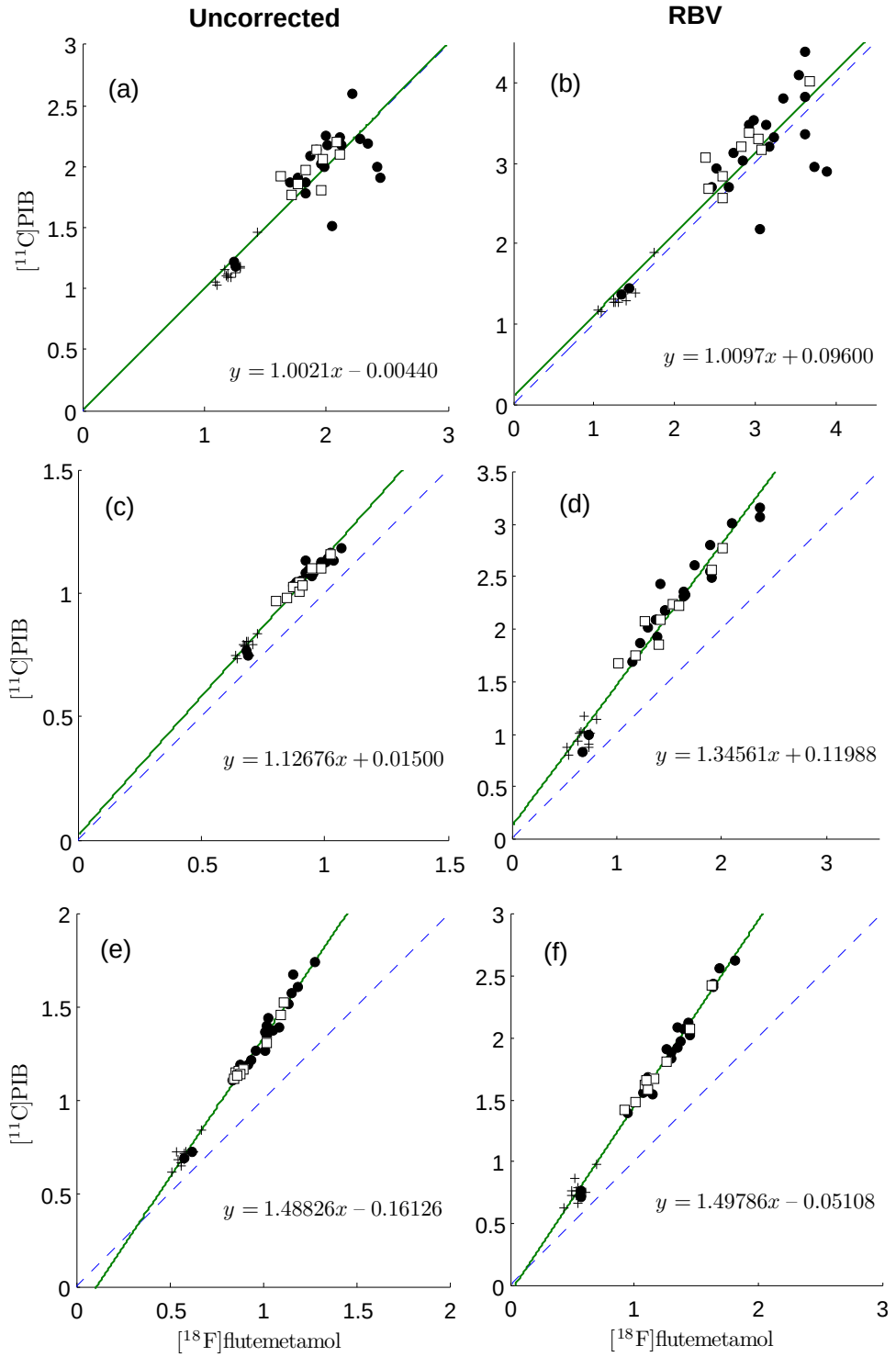


Figure 4.7: Scatter plots of mean [^{18}F]flutemetamol SUVR and mean [^{11}C]PIB SUVR for the composite cortical region of AD (black circles), amyloid-positive (white squares) and amyloid-negative (black crosses). Plots *a* and *b* show cerebellar GM normalisation, *c* and *d* are normalised to sub-cortical WM and plots *e* and *f* are normalised using the brain stem. *a*, *c* and *e* are uncorrected for PVEs. RBV correction is shown in *b*, *d* and *f*. The solid line represents the line of best fit. The dashed line is the line of identity.

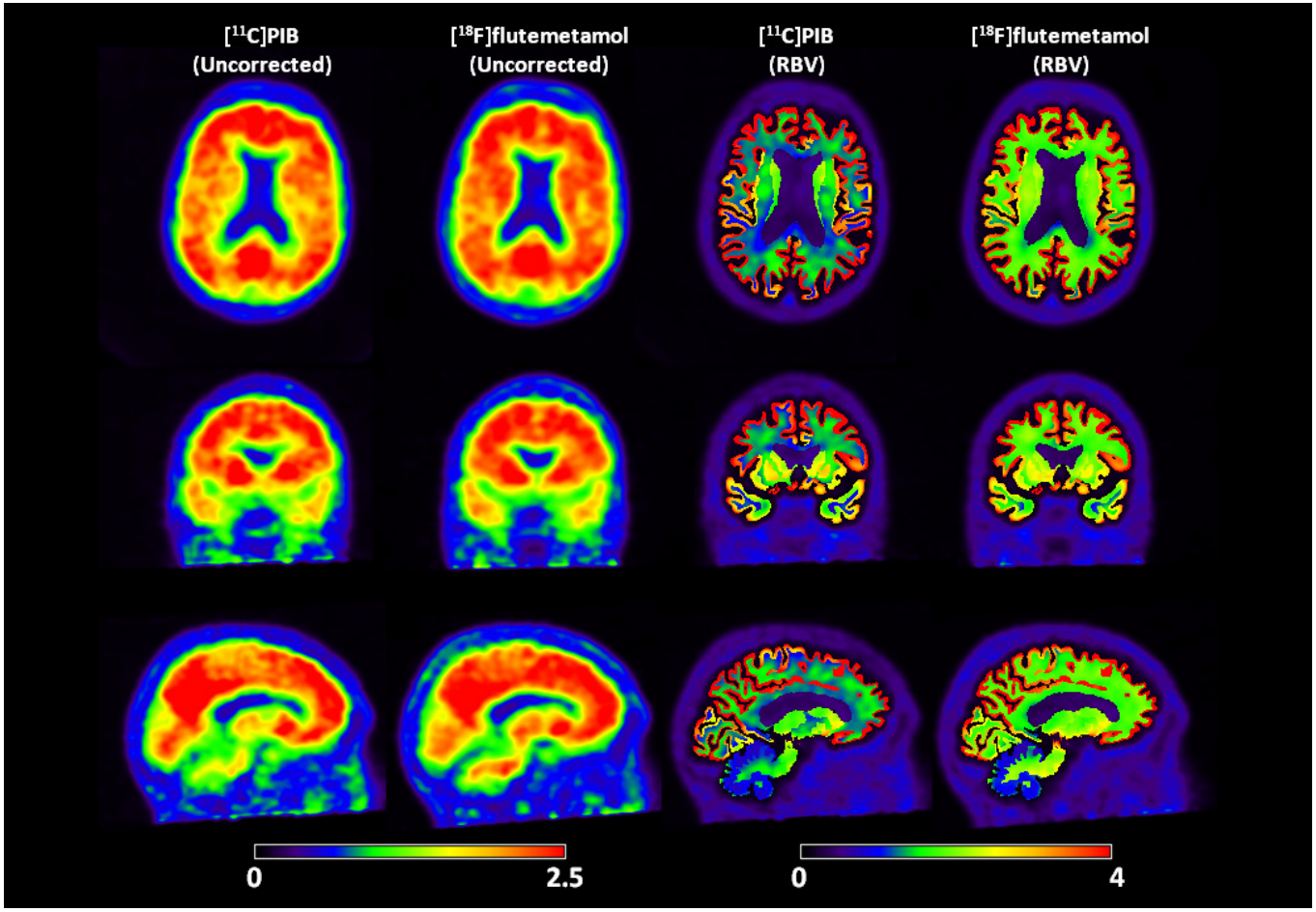


Figure 4.8: SUVR images of an AD subject normalised to cerebellar GM. First column shows the [^{11}C]PIB scan without RBV. The second column depicts the [^{18}F]flutemetamol images without correction. The third and fourth columns show the RBV-corrected images for the [^{11}C]PIB and [^{18}F]flutemetamol respectively.

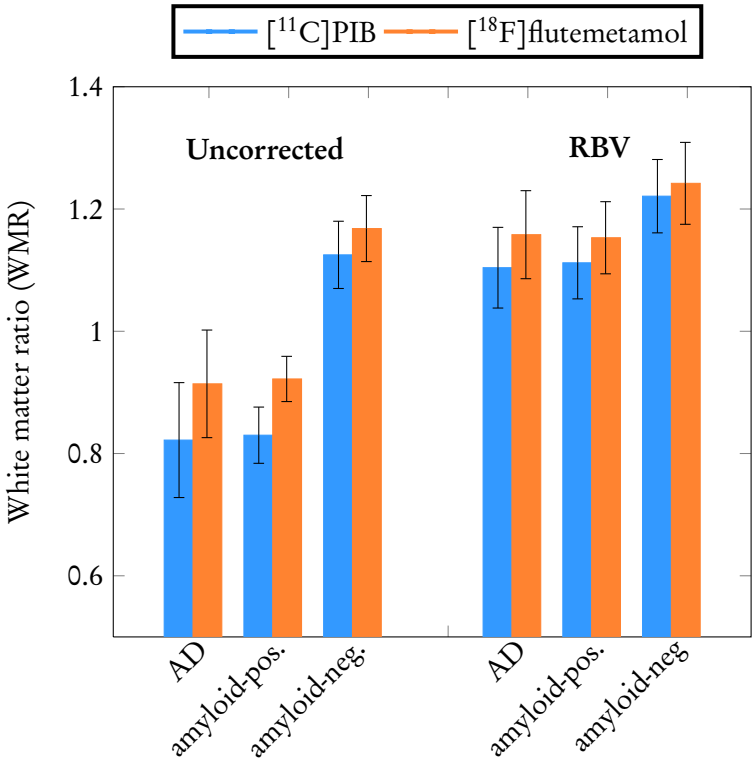


Figure 4.9: Mean cerebellar WM to sub-cortical WM ratio (± SD), with (RBV) and without (Uncorrected) partial volume correction.

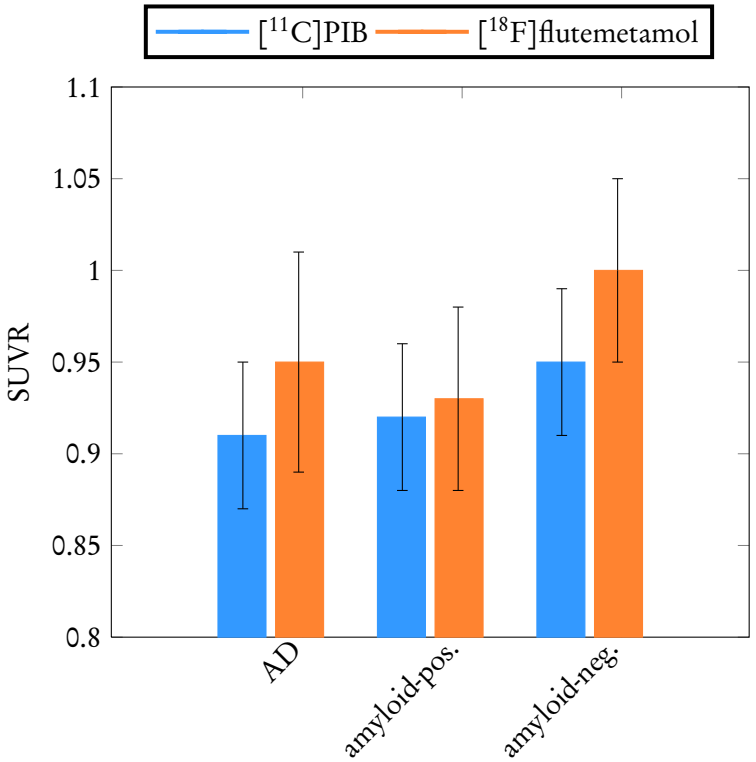


Figure 4.10: The mean cerebellar WM SUVR (± SD) after RBV correction. The brain stem was used as a reference region.

Chapter 5

Segmentation-modifying partial volume correction

This chapter describes the development and application of a novel anatomy-based PVC technique that iteratively modifies the MRI parcellation to fit the PET data. The technique is called segmentation-modifying partial volume correction (SMPVC). The motivation behind the technique is given in section 5.1. The algorithm development is described in section 5.2. Performance is then evaluated using phantom data (section 5.3), [^{18}F]flutemetamol reproducibility data (section 5.4) and [^{11}C]PIB longitudinal data (section 5.5). Conclusions are then given in section 5.6.

5.1 Motivation

Most anatomy-based PVC techniques make an assumption of regional uniformity. Region definitions are typically based on parcellations of structural imaging data. These regions do not necessarily represent regions of uniform uptake in the PET data. Using a set of region definitions that poorly describe the PET distribution will induce bias, particularly at region borders. This has previously been demonstrated when applying the MG to AD subject scans (section 3.1).

For example, amyloid PET imaging is an area where there is likely to be some discordance between the PET and parcellated MRI data. Braak and Braak [1991] state in their pathological AD study that during the early stages of the disease, considerable inter-subject variability was observed in terms of the distribution of amyloid plaque. Whereas, amyloid deposition in the late stages of the disease followed fairly constant pattern. As a result, the authors were only able to define three stages (A, B and C) of amyloid deposition, as opposed to the six stages that could be defined for the accumulation of NFTs. The first stage of amyloid deposition, stage A, consists of low density amyloid deposition, predominantly in the

inferior frontal, parietal and occipital lobes. By stage B, all cortical areas, with the exception of the sensorimotor cortex exhibit amyloid deposition. By the final stage, amyloid plaques are observed in almost the entire cortex as well as sub-cortical GM structures. In effect, the cortex appears to *fill-up* with amyloid. The "scattered amyloid deposits" [Braak and Braak, 1991], of the earlier stages of AD are unlikely to lie on distinct anatomical boundaries. As this example shows, the ability to define regions in a way that more accurately represents the PET distribution is desirable from a PVC perspective.

RBV correction

RBV correction is able to account for variability within tissue types, although errors may still exist in a sub-region. While it is clear that different tissue compartments should be defined as separate regions, the appropriate approach to sub-dividing these compartments is less obvious. This is especially the case for GM in AD subjects undergoing amyloid PET imaging. RBV correction relies on the assumption that the PET distribution within a parcellated region is uniform. As discussed above, it is unlikely that areas of amyloid PET uptake (or hypometabolism in [^{18}F]FDG) will correspond exactly with anatomically derived regions. Applying RBV correction with region definitions that do not satisfy this assumption will cause errors in the region, especially at boundaries and also in neighbouring regions such as sub-cortical WM. Conversely, if the regions can be defined in order to satisfy the uniformity assumption, RBV correction will accurately perform PVC.

Hyper-plane PVC

The anatomy-based PVC methods described so far in this thesis make an assumption of uniformity in one or more of the regions. Hyper-plane partial volume correction (h-PVC) [Erlandsson and Hutton, 2011] allows every region to contain a gradient in one direction. The technique uses the same correction step as given by equation 3.1. Each region is represented by its mean value and a 3D vector which represents the gradient measured from the observed PET image. During the formation of the image used to calculate the correction factors, the value of each voxel is calculated as the product of the regional mean value and the gradient at that point. The voxel-wise correction is then performed in the same manner as RBV.

Preliminary results using SPECT phantom data [Erlandsson and Hutton, 2011] indicate that h-PVC would more accurately correct for PVEs if a gradient were to exist in the uptake of a region. While the assumption of a gradient may not be valid, the notion of fitting some function to a region may prove to be better than assuming uniformity. h-PVC was applied to striatal SPECT data, although the concept may also be appropriate for correction in the

cortex. Like RBV, h-PVC still relies on a suitable region definition based on MRI data. While h-PVC does not assume uniformity, it does assume that the region defined by the MRI contains an activity distribution that can be described by a gradient and therefore its performance is still reliant on the anatomical definition.

Changing the region definitions to better fit the observed PET data will result in a more accurate PVC, reducing intra-regional bias. In this chapter, a novel algorithm is proposed which iteratively modifies the initial anatomical parcellation to find the optimal set of regions. This optimisation is driven by the PV-corrected PET data. The current implementation operates on the assumption of regional uniformity, although could also be applied to h-PVC. This new technique has the additional benefit of providing a parameter of *functional volume change* which may be useful for longitudinal studies.

5.2 Algorithm development

This section describes the process undertaken during development of SMPVC. Two algorithms are reported, the first was designed as a voxel-based method, the second is a surface-based approach. The aim of both is to alter the underlying segmentation that is used to perform PVC. The voxel-based and surface-based approach are given here in order to demonstrate the evolution of SMPVC. The goal of SMPVC is to find a set of regions which reduces the global regional variance in an image. SMPVC is performed as a post-reconstruction PVC technique that uses a MRI parcellation as a starting point, evolving the region definitions, given the observed PET data. Both algorithms are described in the remainder of this section in chronological order.

5.2.1 Voxel-based SMPVC

SMPVC was first designed as a voxel-based method. The concept was to extend the RBV correction by iteratively modifying region boundaries and calculating changes to regional variability based on estimates of PV-corrected voxel values. The technique operates as follows: first, RBV correction is performed using the initial MRI and PET data. Regional values for the mean and variance are calculated based on the RBV-corrected PET data. All voxels that lie on GM to GM boundaries are then subjected to analysis.

5.2.1.1 Label swapping

Each edge voxel is treated in turn and becomes the *target voxel*. The label of the target voxel is swapped for the label of the neighbouring voxel. Where there exists more than one neighbouring GM region, the assigned swap label becomes the label which is the mode of GM voxels in the $3 \times 3 \times 3$ neighbourhood surrounding the target voxel. When a tie exists between

candidate neighbours, a swap label is chosen at random.

Once the label has been swapped, the neighbourhood surrounding the target voxel is evaluated to ensure that the label change has not resulted in a disconnection of any GM regions. If a disconnection has occured, the label of the target voxel is returned to its original value and no further evaluation of that voxel is performed. A further *connectedness* constraint is placed on the target voxel, which defines the number of voxels of a particular label that must exist in the $3 \times 3 \times 3$ neighbourhood. The number of voxels belonging to the same label as the chosen swap label (m) is calculated. If m is less than the connectedness constraint, the original voxel label is restored and processing of that voxels is stopped.

The connectedness constraint is used to ensure that region surfaces remain intact, with the constraint being relaxed from 18 voxels down to 6 voxels as the algorithm iterates. The values for constraint were found experimentally. Strictly speaking, the initial value for the connectedness constraint of 18 could have been set to 25 ($3 \times 3 \times 3$ - target - neighbour). However, this results in several iterations of the algorithm where very few voxels can be changed and is therefore computationally inefficient. The lower value of 6 was chosen as a minimum connectedness (m_{min}) to ensure that all voxels were at least *6-connected* to their region. This is to prevent regions 'breaking-up' as the algorithm iterates, although it restricts the changing of voxels in areas where the anatomy is very thin. The m_{min} value of 6 is therefore a trade-off between maintaining region integrity and allowing the flexibility to modify regions.

5.2.1.2 Voxel value estimation

All voxels that successfully meet the criteria for swapping in terms of their label are then evaluated according to an estimate of their PV-corrected PET value. To calculate the PV-corrected value of a voxel, it is necessary to perform PVC. However, to calculate a complete RBV correction in order to evaluate a single voxel (or set of voxels), over multiple iterations, is very computationally expensive. With thousands of edge voxels to evaluate, as is the case with a parcellated brain image, directly calculating the PV-corrected voxel value by performing a complete PVC becomes intractable. Therefore, an approximation of the voxel value, given its new label, is made. The calculation of the new value is performed using the estimates of the RBV-corrected regional mean values. The process is described below.

For the purposes of explanation, the equation for RBV correction (equation 3.1, page 56) is given here:

$$f_C(x) \approx f_O(x)\left[\frac{s(x)}{s(x) \otimes h(x)}\right], \tag{5.1}$$

$$s(x) = \sum_{i=1..N} [\mathbf{T}_i p_i(x)].$$

where $\mathbf{T}_i$ is the regional mean value of i calculated by the GTM, p_i is the mask of region i, s is a piece-wise constant image of the regional mean values, h is the PSF of the scanner, f_O is the observed PET image and f_C is the corrected image. Equation 5.1 is then re-written as:

$$\begin{aligned} f_C(x) &= f_O(x)\left[\frac{s(x)}{b(x)}\right], \\ b(x) &= s(x) \otimes h(x). \end{aligned} \tag{5.2}$$

For a voxel x, let u be the index of the voxel's original label and v the index of the new label. The images s and b are then updated with an approximation of their values if the membership of the target voxel had actually been swapped:

$$\begin{aligned} s(x)' &= s(x) + (\mathbf{T}_v - \mathbf{T}_u)\delta(x), \\ b(x)' &= b(x) + (\mathbf{T}_v - \mathbf{T}_u)\delta(x) \otimes h. \end{aligned} \tag{5.3}$$

where $\mathbf{T}_u$ and $\mathbf{T}_v$ are the regional mean values calculated by the GTM. The approximated PV-corrected voxel value can then be computed by:

$$f'_C(x) \approx f_O(x)\left[\frac{s'(x)}{b'(x)}\right] \tag{5.4}$$

5.2.1.3 Acceptance criterion

A label swap is accepted for a particular voxel x if the value $f'_C(x)$ reduces the *global variance* over the set of regions. This is calculated using a region-based cost function, using the global variance before and after $f'_C(x)$ is applied. The measure of global variance (GV) for an image f is given by:

$$GV(f) = \sum_{i=1..N} \left[\frac{\left(\sigma_i^2(f)/\mu_i(f)\right)}{\sqrt{n_i}(f)}\right] \tag{5.5}$$

where i is an index to the region label, N is the number of regions, σ_i^2 is the variance, μ_i is the mean and n_i is the total number of voxels belonging to region i. The difference made to the global variance by swapping voxel x is then calculated by:

$$\Delta GV = GV(f'_C) - GV(f_C) \tag{5.6}$$

A negative value for ΔGV indicates that swapping voxel x from label u to label v reduces the variance across the set of regions. Voxel changes that result in a negative ΔGV value are therefore accepted as valid voxel changes. To reduce computation time required to calculate ΔGV, the mean, variance and size of the regions are computed using an *online* variance calculation [Knuth, 1998, p. 232]. This means that the region properties are updated as voxels are changed rather than being recomputed every time. However, the value of ΔGV is likely to be very small, especially when the regions being modified are large and is sensitive to noise. The noise sensitivity will be discussed later in the section.

5.2.1.4 Iterative scheme

Once all edge voxels have been evaluated, a full RBV correction is performed using the modified mask. ΔGV is calculated between the original f_C and the modified image f'_C. If $\Delta GV < 0$ then the changes to the mask are accepted and f_C is set to equal f'_C. The voxel evaluation process starts again based on the new corrected image and the new mask. If $\Delta GV \geq 0$, the changes are rejected. In this case, the connectedness criterion is decremented by 1. The process is terminated if either the connectedness criterion becomes less than m_{min} or 200 iterations have been performed. Once the stopping criterion has been met, a final RBV correction is performed on the original PET image, using the current mask.

5.2.1.5 Preliminary evaluation

Initial tests were carried out by performing voxel-based SMPVC on digital phantom data. These data were generated using the image generator described in section 3.2.3 (page 67). A full description of the phantom dataset is given in section 5.3 as the same dataset was later used to evaluate the surface-based version of SMPVC. In brief, a parcellation error was introduced into the mask image used for the purposes of PV-correcting the PET.

Ten noise realisations of an AD-like [^{11}C]PIB distribution were evaluated. For some realisations, the voxel-based SMPVC approach would correctly modify the region mask until it was very similar to the ground truth. The boundaries tended be ragged, but this was to be expected given that the correction is based on the changing individual pixels. However, in four of the ten realisations, the voxel-based technique failed to correct the boundary error. Of these, three exhibited a 'tearing' where two regions move against each other, resulting in long strands at the region boundary. The tearing can be seen in figure 5.1.

The other realisation that failed to correct properly was due to an infinite loop caused by a pattern of voxel changes that would repeat until the maximum number of iterations was exceeded. Both the tearing and looping can be attributed to voxel noise sensitivity. The same phantom was tested without noise (other than the resolution blurring) and these effects were

Figure 5.1: An example of tearing observed when applying the voxel-based SMPVC. Images are of the parcellated brain mask. The mask error (left), the ground truth (centre) and result of voxel-based SMPVC (right) are shown. The frontal region is shown in green, central region in purple, WM in red and parietal in yellow.

not observed. It was felt that modifying region surfaces, rather than individual voxels along the surface, would be less sensitive to noise. This led to the development of the surface-based approach detailed in the next section.

5.2.2 Surface-based SMPVC

The noise sensitivity of the voxel-based approach led to the reformulation of SMPVC which modifies boundary surfaces rather than voxels. The core concept of iteratively fitting regions according to the observed PET data remains, although the surface-based approach contains notable differences. The key changes between the voxel- and surface-based approaches are the PVC technique used internally by the algorithm and how the boundary changes are evaluated.

5.2.2.1 Iterative Yang partial volume correction

Voxel-based SMPVC relies on the RBV correction, that in itself is not particularly computationally demanding and can be performed in a few minutes. However, when evaluating thousands of voxels, the overhead quickly becomes very large. This led to the approximation of voxel changes, which while considerably faster to calculate, are less accurate than complete RBV correction.

The computationally intensive part of RBV is the calculation of corrected mean values using GTM. Whereas, the voxel-wise correction step of Yang et al. [1996] is quick to calculate. The previously discussed projection-based PVC method of Erlandsson and Hutton [2010] (section 2.3.2.3, page 36) uses an iterative PVC approach based on the Yang method. This correction is referred to as iterative Yang (iY). The iY correction was applied in the projection domain, although later h-PVC was performed using iY in the image domain. Due to the lower computational cost of iY, this technique was selected for PVC inside the surface-based

SMPVC. iY is described by the equation below:

$$\begin{aligned} f_{k+1}(x) &= f_{PET}(x)\left[\frac{s_k(x)}{s_k(x)\otimes h(x)}\right], \\ s_k(x) &= \sum_{i=1..N}[\mathbf{T}_i p_i(x)], \\ \mathbf{T}_i &= \frac{1}{n_i}\left[\sum_{j\in p_i} f_{kj}\right], \\ f_0 &= f_{PET}. \end{aligned} \tag{5.7}$$

where f_{PET} is the observed PET image, p_i is the binary mask of region i, n_i is the total number of voxels in i and $\mathbf{T}_i$ is the mean value of region i at iteration k. s_k is a piece-wise constant image of the regional mean values and f_k is the PV-corrected image at iteration k. The iY method typically converges in 3 - 5 iterations, producing a PV-corrected image faster than if the GTM were used for the calculation of the mean values. Not needing to calculate the GTM also makes iY arguably easier to implement than RBV.

5.2.2.2 Surface evolution

The surface-based SMPVC uses a fast marching level set method [Sethian, 1996] to evolve region boundaries. The purpose of fast marching methods is to track a moving boundary which moves according to a known *speed* function. This approach has advantages over other methods that rely on control points to define surface meshes in that there are fewer problems in terms of the preservation of topology.

Fast marching methods are a special case of the more general *level sets* [Sethian, 2001], in that fast marching methods are monotonic and therefore only move 'forward'. This makes fast marching methods very quick compared to other numerical methods, however their simplicity may make them unsuitable for some applications. For the purposes of SMPVC, where the aim is to evolve a region boundary a short, constrained distance, fast marching methods were deemed acceptable.

The fast marching method operates by evolving a surface according to the speed function and recording the amount of 'time' it takes the surface boundary to propagate to a particular point. The method is initialised with a set of seed points which define the surface to be evolved. In this instance the output is an image where each voxel represents the time taken for the surface to reach it.

The output image consists of zeros at the seed points, with larger values at voxels where the boundary took longer to reach them. The voxel value is referred to as the *crossing time*.

A new region can then be defined by applying a binary threshold for a desired crossing time. The crossing time at a particular voxel depends on the speed function. In this instance, the speed function is an image and commonly referred to as a *speed image*.

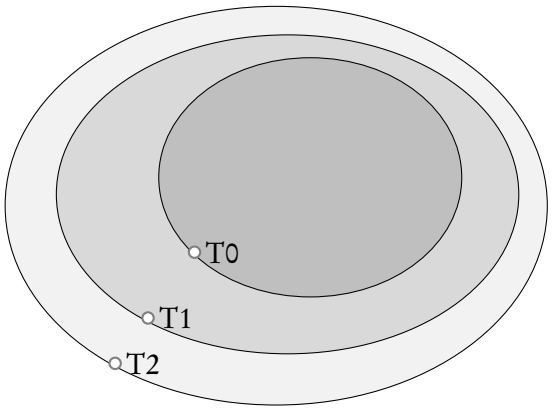


Figure 5.2: An illustrative example of how the initial boundary of an object (crossing time T0) could propagate at crossing times T1 and T2.

In the case of SMPVC, the boundary to be propagated is that of a particular brain region. The speed image is defined according to the regional image statistics from the PV-corrected PET data. More specifically, when evolving the surface of region i, the mean of region i (μ_i) is used as input to a modified sigmoid function. The function is linear transformation that is necessary, first because the fast marching filter of the Insight Segmentation and Registration Toolkit (ITK) expects the speed image to contain values of between 0 and 1 and second to encourage faster surface propagation in regions that are more similar to each other. The surface moves faster in areas of the speed image that are equal to 1 and slows down where the image tends to 0. The function (ϕ) used to calculate the speed image is:

$$\begin{aligned} \phi(x) &= 1 - |\Psi(x) - 0.5|, \\ \Psi(x) &= \left[\frac{1}{1 + e^{-\frac{x-\beta}{\alpha}}} \right], \\ \alpha &= \left[\frac{2 \times \mu_i}{6} \right], \\ \beta &= \mu_i. \end{aligned} \tag{5.8}$$

where μ_i is the iY-corrected PET mean value in region i and ϕ is the value assigned in the speed image. The value of ϕ as a function of the percentage difference from μ_i can be seen in figure 5.3.

In the current implementation, the surface evolution is constrained to GM regions only. Additionally, GM regions are restricted to their respective hemispheres. During initial tests

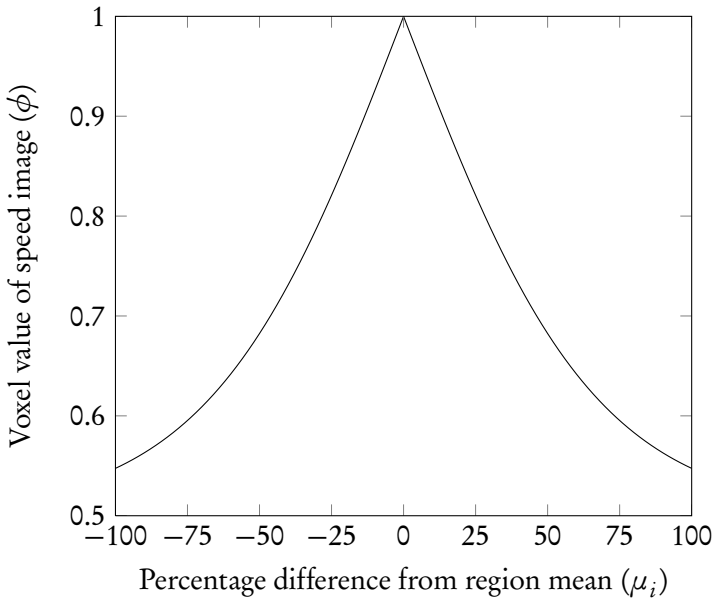


Figure 5.3: The modified sigmoid function (equation 5.8) used to compute the speed image.

of the voxel-based approach, regions such as the cingulate could spread from one hemisphere to another. While it is feasible that inter-hemispheric region definitions may reduce the global variance, they may not accurately describe the PET distribution. This behaviour is therefore prevented by checking the voxel labels in the parcellated brain mask image.

5.2.2.3 Iterative scheme

The surface-based SMPVC algorithm iterates in a different manner to the voxel-based approach. The previous implementation would evaluate voxel changes from all regions during an iteration. However, the surface-based method evolves the surface of a single region and then calculates whether the changes to the region surface improves the PVC. The algorithm iterates over the set of regions, evolving each surface individually. For a region change to be accepted the following three criteria must be met:

1. The acceptance criterion for an iteration is calculated in the same way the voxel-based approach (section 5.2.1.3). ΔGV must be less than 0 for the change to be accepted.

2. The CoV_r in the region must not increase after modification.

3. The size of the region must be within ±20% of the original region size.

The criterion based on the CoV_r is applied to prevent a modification that degrades a region but is not reflected by the ΔGV. As ΔGV is a global measure, it is possible for this situation to occur. Conversely, it is entirely possible that the CoV_r in a region should increase. In this

instance, the change should occur when the neighbouring region is modified as the criterion would be satisfied.

The size restriction was added for two reasons. Firstly, it prevents small regions from disappearing and secondly it limits the problem where two neighbouring regions may be very similar to each other and can therefore change dramatically in terms of their size, while having little effect on the mean or CoV_r of those regions. The latter effect could make any parameter based on volume change quite unstable.

As discussed earlier, the surface-based algorithm iterates over the set of regions. This is performed several times as changes to one region may prompt changes in another. The algorithm therefore has two loops, an inner loop controlling iterations over the set of regions and an outer loop controlling the number of times to evaluate the sets. The surface-based SMPVC algorithm is described in pseudocode in section 5.2.2.4. Once both loops have terminated, the modified mask image is used to perform a RBV correction on the original PET data to produce the final PV-corrected image. RBV is applied because, while results of iY correction is very similar to RBV after 5 iterations, RBV is still considered to be the most accurate correction when assuming regional uniformity. The surface-based SMPVC approach was selected over the voxel-based technique and all subsequent discussion of SMPVC refers to the surface-based approach unless specifically stated.

5.2.2.4 Algorithm pseudocode

Algorithm 1 Surface-based SMPVC

$k \leftarrow 1$ {Number of times to evaluate regions}

$[\ imgPVC\ ] \leftarrow$ iYang($imgPET$, $imgMask$) {Perform first iterative Yang correction}

while $k \leq kMax$ **do**
  $i \leftarrow 1$ {Region index}
  while $i \leq iMax$ **do**

    $[\ imgMaskTmp\ ] \leftarrow$ EvolveRegion($imgMask$, i)
    $[\ imgPVCTmp\ ] \leftarrow$ iYang($imgPET$, $imgMaskTmp$)

    $[\ dGV, dSZ, dCOV\ ] \leftarrow$ EvaluateChanges( )

    if ($dGV < 0$) **and** ($dSZ \leq 20$) **and** ($dCOV[i] \leq 0$) **then**
      {Accept changes and update images}
      $imgPVC \leftarrow imgPVCTmp$
      $imgMask \leftarrow imgMaskTmp$
    else
      {Reject changes and move to next region}
      $i \leftarrow i + 1$
    end if
  end while
  $k \leftarrow k + 1$
end while

$[\ imgPVC\ ] \leftarrow$ RBV($imgPET$, $imgMask$) {Perform final RBV correction}

5.3 Phantom experiments

This section reports the evaluation of SMPVC using digital phantom data. The performance is assessed through a comparison of SMPVC with RBV correction. First, preliminary experiments using a geometric phantom are described. A digital anthropomorphic brain phantom is then evaluated to assess performance in a more realistic object.

5.3.1 Materials and methods

5.3.1.1 Phantom datasets

All phantom data used during these experiments was generated using the phantom image generator described in section 3.2.3 (page 67). The geometric phantom consists of a 64 × 64 × 64 matrix of 1mm^3 voxels. The object is comprised of three conjoined cuboids (figure 5.4). An activity distribution was created where region 1 had 9% lower activity than region 3 and region 2 had 75% lower activity than region 3 (figure 5.5a). A uniform CT image was

created for the purposes of simulating attenuation. Ten noise realisations were generated for two noise levels (1.0e^7 and 2.0e^6 counts) with a resolution of 6mm FWHM. A slice from one realisation can be seen in figure 5.5b. A parcellation error was created in the mask by extending regions 1 and 3 into region 2. The error can be seen in figure 5.5c.

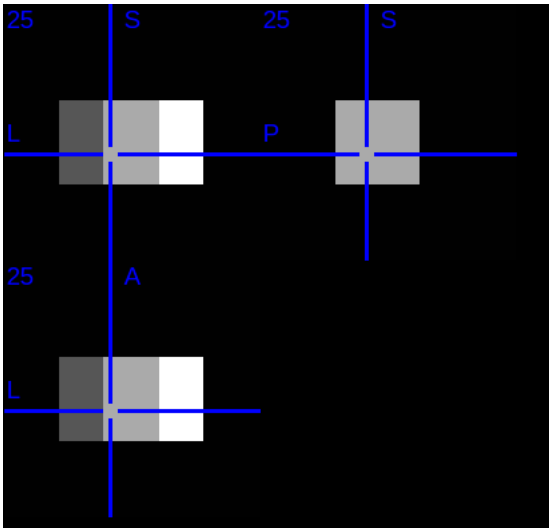


Figure 5.4: The digital geometric phantom in transaxial (bottom left), coronal (top left) and sagittal (top right) planes.

The digital anthropomorphic brain phantom was generated from a T1 MRI image. FreeSurfer parcellation was performed on the MRI and a region mask was defined according to the 24 ROI parcellation (section D.2, page 180). The subject CT was registered to the MRI data for the purposes of simulating an attenuation corrected PET image. The matrix size was $256 \times 256 \times 256$ with 1mm^3 voxels. A typical AD subject [^{11}C]PIB distribution was applied to the region mask for the purpose of generating the phantom data. Ten noise realisations were generated at a count level of 5.4e^8 and with a reconstructed resolution of 6.0mm FWHM. A parcellation error was induced in the region mask by dilating the left central GM region into the frontal GM. The dilation was restricted to GM voxels and manually inspected to ensure that the regions were contiguous. The parcellation error can be seen in figure 5.6, overlaid on the T1 MRI image.

5.3.1.2 Partial volume correction

RBV correction and SMPVC was performed on all phantom images. The RBV correction was performed twice; once with the region mask used to generate the data and once with the mask containing the parcellation error. RBV correction with the original region mask is referred to as *true*, as it is considered to be the ground truth. RBV correction with the mask error is called *RBV (error)*. SMPVC was performed using the erroneous mask only. A resolution of 6.0mm

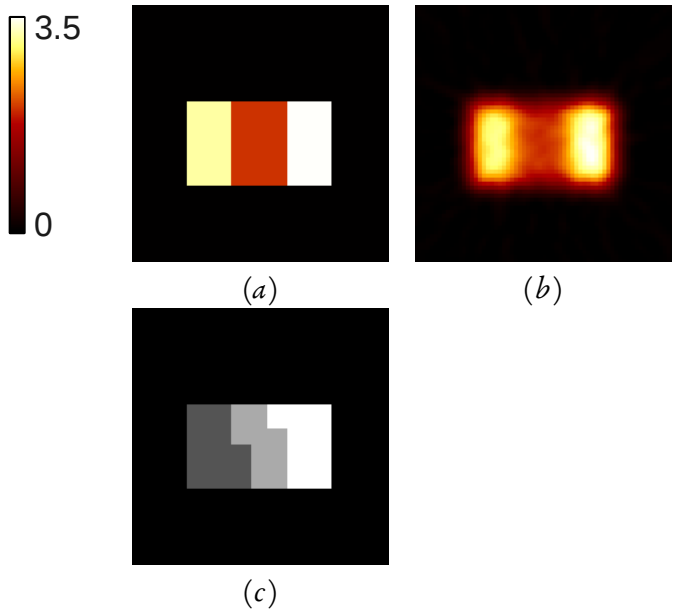


Figure 5.5: A transaxial slice of the geometric phantom with activity distribution applied is shown in (a). The result of the phantom image generator can be seen in (b). Region 1 (left) has a value of 3.2, region 2 (centre) has a value of 2.0 and region 3 (right) is 3.5. The parcellation error induced in the phantom mask can be seen in (c).

FWHM was assumed for the purposes of PVC, this matches the reconstructed resolution of the simulated PET data.

When performing SMPVC on the geometric phantom data, the restriction of preventing a region becoming 20% larger or smaller than its original size, was ignored. This was due to the magnitude of the errors applied in the geometric phantom. The criterion was enforced for the anthropomorphic brain phantom. In addition, region changes during SMPVC were restricted to their respective hemispheres, as described in section 5.2.2.2.

5.3.1.3 Analysis

Regional analysis was performed on the RBV-corrected and SMPVC images. All the images were analysed in MATLAB. In terms of the PET data, mean values were calculated for the ROIs using a 10% trimmed mean value. These trimmed mean values were used to compute RCs. The CoV_r was also calculated for each ROI, given by the regional SD divided by the regional mean. The SD across realisations was found for both the RC and CoV_r. The RBV (error) images were also compared with SMPVC using paired t-tests.

In addition to the PET-related statistics, the accuracy of the parcellation was also evaluated. The Dice Coefficient (DC) [Dice, 1945] is a similarity measure which produces a value of between 0 and 1. Two sets of data are considered to be more similar as the DC tends to 1.

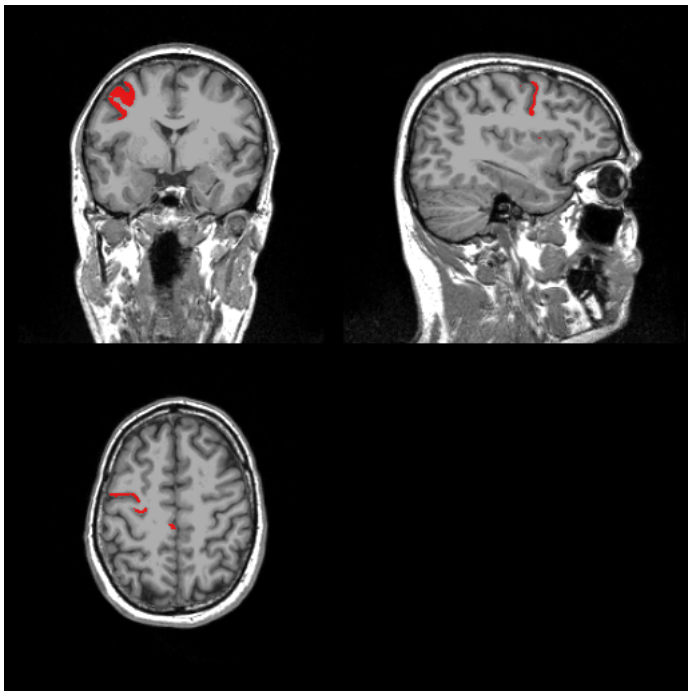


Figure 5.6: Simulated parcellation error between the left central and frontal regions. The red region represents the voxels than have become part of the left central region after dilation.

For two binary masks, p_a and p_b, the DC is calculated as:

$$DC = \frac{2 \times |p_a \cap p_b|}{|p_a| \cup |p_b|} \tag{5.9}$$

In this analysis, the DC measures the similarity between the modified parcellation and the ground truth. The SD of the DC across realisations is also calculated. The DC of the mask containing the parcellation error compared to the ground truth is also stated for reference purposes. Paired t-tests were also performed between the DC of the parcellation error and the DC after SMPVC, for both regions.

5.3.2 Results

5.3.2.1 Geometric phantom

The RC in all regions of the geometric phantom were close to 1 irrespective of correction or the mask used (table 5.1). A RC of 0.97 or greater was observed for all three regions at the $1.0e^7$ count level after SMPVC. The tables of results for the lower count level can be found in appendix A, page 168. Performance in terms of RC for the $2.0e^6$ count level was very similar to the higher count dataset, albeit with wider SD values. The error in recovery was smallest in region 3. This region contained the highest activity of the regions. The parcellation error between region 2 and 3 also is half the size of the error between 1 and 2. Significant differences

($p < 0.05$) in RC between RBV (error) and SMPVC were observed in regions 1 and 2. The difference in RC after SMPVC was not significant in region 3 ($p = 0.059$).

The CoV_r significantly reduces ($p < 0.05$) in all regions with SMPVC compared to RBV with the parcellation error (table 5.2). The largest reduction between the two methods was observed in region 2, where the CoV_r changes from 0.121 ($\pm$ 1.77e^{-3}) after RBV to 0.019 ($\pm$ 1.00e^{-3}) after SMPVC. This equates to a 79% reduction in CoV_r. In regions 1 and 3 the reductions are 39% and 14% respectively. Figure 5.7 shows effects of the PVC techniques in a transaxial slice. When RBV is applied with the erroneous mask, areas in region 2 are over-corrected where the parcellation errors occurs. The remainder of region 2 is under-corrected, hence the large CoV_r in region 2. This over- and under-correction is not however reflected in the RC. The variance is visibly reduced after SMPVC for the areas subjected to the initial parcellation errors.

The DC is reported for all regions in table 5.3. After SMPVC, the DC increases for all regions compared to RBV with the erroneous mask. This suggests that the accuracy of the parcellation improves with SMPVC. The highest DC values were observed in region 3. The percentage improvement in DC between RBV and SMPVC was 10%, 14% and 3% in regions 1, 2 and 3 respectively. As the DC values for SMPVC are less than 1, errors remain in the parcellation compared to the ground truth. These can be observed in figure 5.7 where the boundary between region 1 and region 2 is slightly 'ragged'.

	RBV (error)	True	SMPVC
Region 1	0.981 ($\pm$ 3.41e-03)	0.996 ($\pm$ 3.67e-03)	0.987 ($\pm$ 6.43e-03)
Region 2	0.998 ($\pm$ 4.03e-03)	0.985 ($\pm$ 4.37e-03)	0.977 ($\pm$ 6.11e-03)
Region 3	0.999 ($\pm$ 3.02e-03)	0.998 ($\pm$ 3.74e-03)	0.993 ($\pm$ 3.19e-03)

Table 5.1: The RC ($\pm$SD) for PV-correction of the geometric phantom at 1.0e^7 counts.

	RBV (error)	True	SMPVC
Region 1	0.031 ($\pm$ 9.52e-04)	0.018 ($\pm$ 6.91e-04)	0.019 ($\pm$ 1.00e-03)
Region 2	0.121 ($\pm$ 1.77e-03)	0.023 ($\pm$ 1.03e-03)	0.026 ($\pm$ 8.26e-04)
Region 3	0.021 ($\pm$ 6.10e-04)	0.018 ($\pm$ 1.19e-03)	0.018 ($\pm$ 1.11e-03)

Table 5.2: The CoV_r ($\pm$SD) for PV-correction of the geometric phantom at 1.0e^7 counts.

	RBV (error)	True	SMPVC
Region 1	0.894	1.00	0.982 (± 2.83e-03)
Region 2	0.863	1.00	0.981 (± 2.69e-03)
Region 3	0.967	1.00	0.993 (± 1.24e-03)

Table 5.3: The DC (±SD) for PV-correction of the geometric phantom at $1.0e^7$ counts. Note that the DC for the RBV-corrected data, with and without parcellation error does not change as the same masks are used for each realisation. SD values are therefore not reported for the RBV-corrected data.

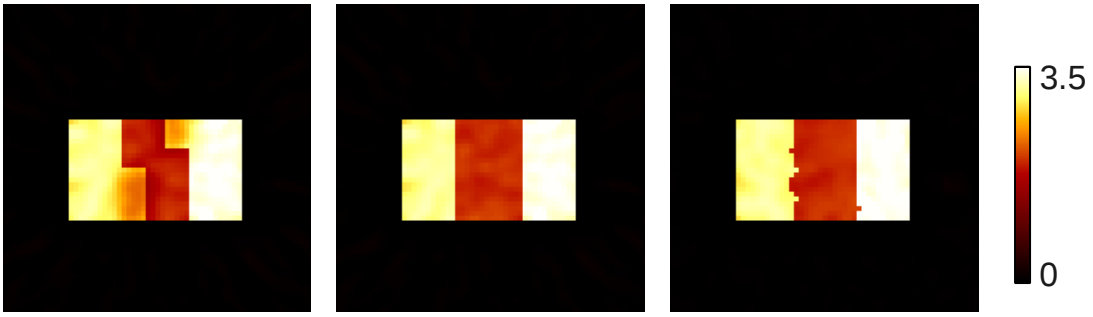


Figure 5.7: Transaxial slice of geometric phantom corrected with RBV using the incorrect mask (left), correct mask (centre) and after SMPVC (right). Images are scaled between 0 and 3.5, where 3.5 is highest value in the phantom (region 3) before noise is added.

5.3.2.2 Anthropomorphic brain phantom

The true activity in the left central region was 25% lower than the left frontal region. The left central region mask was dilated into the left frontal region, this parcellation error resulted in an over-correction ($RC > 1$) in the central region and under-correction in the frontal region when performing RBV with the erroneous mask. The RC for both regions is reported in table 5.4. The left frontal region RC significantly increases ($p < 0.05$) from 0.988 to 0.993 with SMPVC. The over-correction in the central region significantly decreases ($p < 0.05$) from 1.010 to 0.998 when SMPVC is applied.

Significant reductions in CoV_r were observed in both regions after SMPVC compared to RBV (table 5.5). The frontal region exhibited an average 32% reduction in CoV_r ($p < 0.05$) when SMPVC was performed. A 22% average reduction ($p < 0.05$) in observed in the central region when using SMPVC compared to RBV. Despite the reduction for the central region, the CoV_r remained 10% higher with SMPVC than the ground truth. A magnified section of a PV-corrected slice on the border between the left frontal and central regions can be seen in figure 5.8. The boundary between the two regions is more clearly defined after SMPVC compared to RBV. Over-correction in the neighbouring WM can also be observed when RBV was performed with the parcellation errors. The WM over-correction is due to under-

correction for the frontal region. Therefore, the WM is more accurately corrected in this area with SMPVC, although WM was not analysed in this study.

The DC is reported for both regions in table 5.6. The frontal and central region DC increases when SMPVC is applied. These increases were significant ($p < 0.05$) for both regions. This suggests that the SMPVC improves the accuracy of the segmentation in both regions. A lower DC value was observed in the central region compared to the frontal. This is probably due to the magnitude of the parcellation relative to the size of the region. The parcellation error increase the size of the central region by 18%, whereas the frontal region decreased by 7%.

	RBV (error)	True	SMPVC
Left frontal	0.988 (± 2.63e-03)	0.996 (± 2.82e-03)	0.993 (± 2.76e-03)
Left central	1.010 (± 4.81e-03)	0.997 (± 5.36e-03)	0.998 (± 6.60e-03)

Table 5.4: The RC (±SD) for PV-correction of the anthropomorphic brain phantom.

	RBV (error)	True	SMPVC
Left frontal	0.025 (± 4.60e-04)	0.018 (± 2.65e-04)	0.019 (± 2.95e-04)
Left central	0.028 (± 6.10e-04)	0.021 (± 5.33e-04)	0.023 (± 5.32e-04)

Table 5.5: The CoV_r (±SD) for PV-correction of the anthropomorphic brain phantom.

	RBV (error)	True	SMPVC
Left frontal	0.965	1.00	0.985 (± 8.41e-04)
Left central	0.920	1.00	0.965 (± 1.94e-03)

Table 5.6: The DC (±SD) for PV-correction of the anthropomorphic brain phantom.

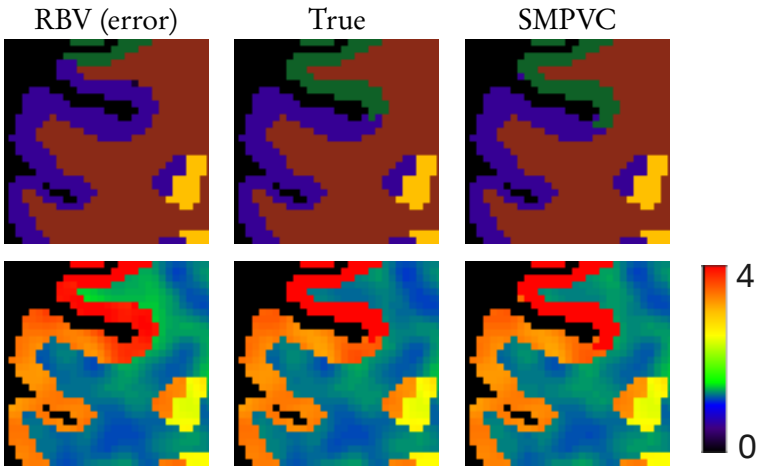


Figure 5.8: A magnified section of a transaxial slice after PV-correction of the anthropomorphic brain phantom. The top row depicts the region mask. The frontal region is shown in green and the central in purple. The bottom row displays the PV-corrected PET data. All PET images are normalised to the cerebellar GM and image intensities are scaled between 0 and 4.

5.3.3 Discussion

SMPVC was evaluated using digital phantom simulations. Two phantoms were created; a geometric phantom and an anthropomorphic brain phantom. The geometric phantom was used during the initial stages of the SMPVC algorithm development to evaluate how the surfaces evolved over iterations and to assess whether the approach was sensitive to noise. The brain phantom data provided a more realistic test in terms of activity distribution, the size of the surfaces to operate on and anatomical complexity. A typical [^{11}C]PIB uptake for an AD subject was modeled during these experiments.

Parcellation errors were induced in the masks used to perform PVC. A comparison between SMPVC and RBV was carried out, evaluating the performance of both techniques, given the parcellation errors. RBV correction with the true region mask was considered to be the ground truth. RBV correction and SMPVC were performed with the erroneous mask and compared to the ground truth in terms of RC, CoV_r and DC. These phantom studies have shown that the SMPVC algorithm is able to modify and reduce the effects of parcellation errors. Correcting the parcellation with SMPVC resulted in a more accurate PVC compared to RBV when parcellation errors were not accounted for. This improved the quantitative accuracy for both the geometric and anthropomorphic phantom.

In terms of RC for the geometric phantom, there is little difference between the RBV with the mask errors, SMPVC and the ground truth. This suggests that the regional trimmed mean value is fairly insensitive to border errors. However, large parcellation errors relative

to the size of the regions were induced, and the errors are reflected in the CoV_r, particularly in region 2. Errors in the RC were observed in the anthropomorphic brain phantom with RBV. The effect of the parcellation error was larger in the central region than the frontal. The inclusion of the 'hotter' frontal voxels in the central region created the observed over-correction in the central region as these voxels caused a positive bias in the estimation of the regional mean value. Under-correction is also observed in the frontal region when RBV is performed with the erroneous mask. Both the over- and under-correction of the brain regions are reduced after SMPVC.

Large significant reductions in the CoV_r for both phantoms were observed after SMPVC. In the geometric phantom the most notable reduction in CoV_r was in region 2, as it was subjected to parcellation errors from both regions 1 and 3. Region 2 was also colder than its neighbours, causing a larger error than in the other regions. SMPVC accurately corrected the geometric phantom over ten realisations, at two noise levels; $1.0e^7$ counts and with five times fewer counts at $2.0e^6$. The CoV_r also significantly reduced when SMPVC was applied to the brain phantom. After SMPVC correction, the average CoV_r values for both brain regions reduced and were comparable to the ground truth. The SD of the CoV_r values with SMPVC is also similar to the ground truth. This suggests SMPVC is reproducible, at least for these phantom datasets. The SMPVC CoV_r values are higher than the ground truth and this is likely to be due to remaining errors in the parcellation.

In addition to the PET-related measures of RC and CoV_r, the DC was also reported for the phantom datasets to evaluate the accuracy of the parcellation. After SMPVC, the DC significantly increased for all regions in both phantoms. This suggests that the parcellation was more similar to the ground truth than the erroneous parcellation. Given the improvements in RC and CoV_r, this is unsurprising as these would not have occurred had the parcellation been poorer. However, there are still errors in the parcellation, albeit relatively small. SMPVC appears to find the correct boundary, but can define a ragged edge along it. The behaviour can be observed in figure 5.7. This is likely due to noise, although noise did not affect the overall movement of the boundaries in the phantom data.

5.4 Assessment of SMPVC reproducibility using clinical data

This section reports the evaluation of PVC in a clinical reproducibility dataset. The study compares RBV correction and SMPVC through regional analysis. Assessment of additional volumetric parameters is also reported.

5.4.1 Materials and methods

The subject information and pre-processing for this study is identical to that described in section 4.2.1 (page 99) and is therefore not reproduced here. In terms of methodology, the studies are the same except that SMPVC is performed on the dataset in addition to RBV.

5.4.1.1 Analysis

RBV correction and SMPVC were also compared through analysis of SUVR and CoV_r. Wilcoxon rank-sum tests were used to evaluate group-wise differences between the two methods for both SUVR and CoV_r in the composite cortical region. Additional analysis was performed on the parcellation after SMPVC. The percentage change in region volume between the RBV-corrected image and the SMPVC image, for each scan, is also reported for the frontal and middle temporal regions. The DC is computed for all movable ROIs, between the two scans, after SMPVC to measure how similar the parcellations of a particular region are.

5.4.2 Results

RBV and SMPVC were compared through regional SUVR analysis. The SUVR values for both scans with RBV and SMPVC correction for the composite cortical region are shown in figure 5.9. The percentage difference in SUVR values of the compostite cortical region, between scans, can be seen in figure 5.10. After SMPVC correction, the overall difference in between scan SUVR values reduces. The between-scan difference of RBV and SMPVC was evaluated with a paired, two-tailed Wilcoxon signed-rank test, using a threshold of $p < 0.05$. This however was not significant ($p = 0.72$, MDD =0.02).

The regional CoV_r was analysed for RBV correction and SMPVC. The CoV_r values of the composite cortical region, for both scans, are shown in figure 5.11. The mean CoV_r for scan 1 with RBV correction was 0.119 ± 0.008, reducing to 0.112 ± 0.005 after SMPVC. For scan 2, the mean CoV_r reduced from 0.119 ± 0.005 to 0.113 ± 0.005. This represents an average reduction in the CoV_r of 5.8% across the group for scan 1 and 5.0% in the case of scan 2. Wilcoxon signed-rank tests were performed between the CoV_r values of RBV and SMPVC, across scan 1 and scan 2. The tests were significant at $p < 0.05$ threshold for both scan 1 ($p = 0.006$) and scan 2 ($p = 0.004$).

In addition to the PET-based statistics, the changes to the parcellation were also exam-

ined. The percentage volume change between the original parcellation and the final output of SMPVC was calculated. The values in the superior and inferior frontal region and left middle temporal region can be seen in table 5.7. These regions are reported as they exhibited large changes in the percentage volume in some scans. There is a consistent trend of reductions in the volume of the superior frontal region, along with increases in the volume of the inferior frontal region. In table 5.7, some of the values, particularly those in the inferior frontal region are close to 20% and may be restricted by the constraint placed on the SMPVC algorithm.

The DC of the SMPVC parcellations was calculated between the two scans for all movable regions. The results can be seen in table 5.8. The lowest (poorest) DC observed was 0.877, found in the right superior temporal region of the 4th subject. Most DC values were above 0.9, which would suggest that there is a strong similarity between the parcellations across scans. It should however be noted that a DC value of 1 suggests that either a region has not moved, or that all voxels in that region have kept their original label and may have expanded into neighbouring regions.

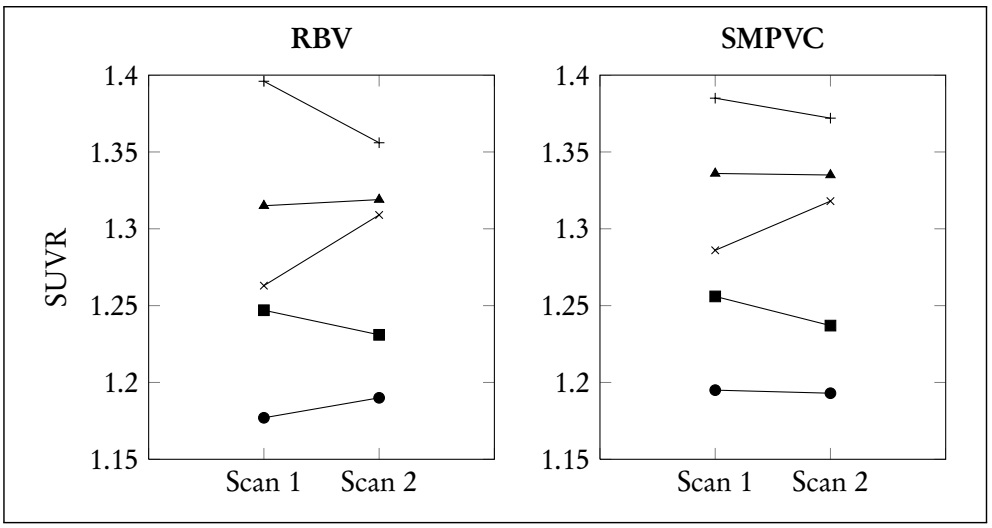


Figure 5.9: SUVR analysis of the composite cortical region after RBV correction (left) and SMPVC (right).

5.4.3 Discussion

This study has investigated the reproducibility of quantitative measurements with [^{18}F]flutemetamol after RBV correction and SMPVC. Having established that RBV does not degrade the reproducibility of SUVR values (section 4.2, page 99) and tends to reduce the CoV_r, especially in the cortical regions, the performance of SMPVC was compared to RBV. SUVR analysis on the composite cortical region showed that SMPVC reduced the between-scan difference in four out of the five subjects, although the reduction did not reach

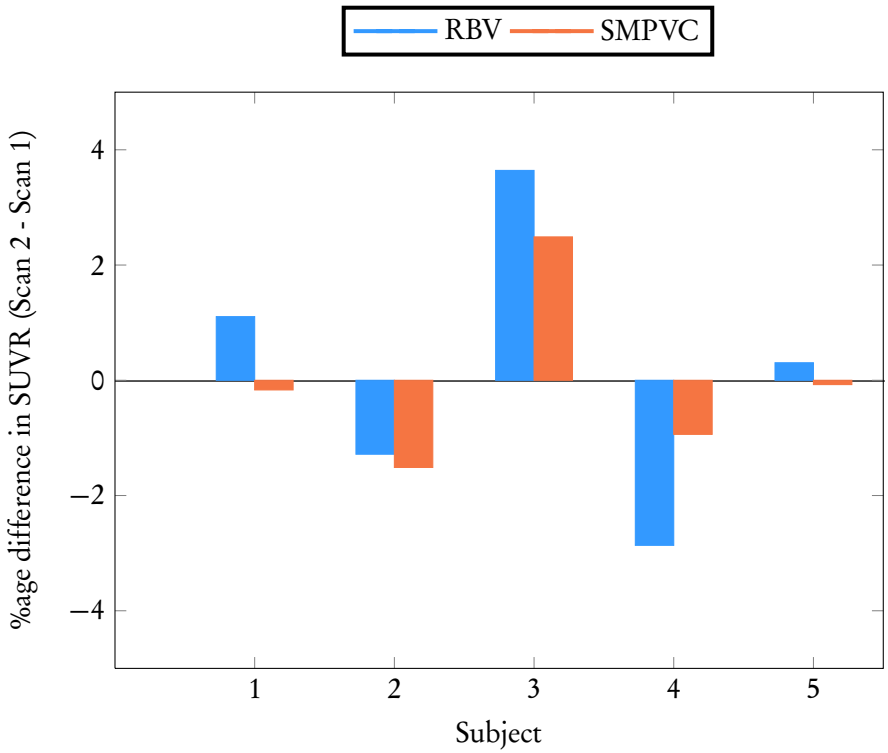


Figure 5.10: The percentage difference in composite cortical SUVR between scans 1 and 2 with RBV correction and SMPVC.

significance. The statistical comparison was powered at 80% to find a MDD of 0.02 at the $p < 0.05$ threshold. However, the mean difference between the two techniques was calculated as 0.010. This may be due to the region being a composite of several ROIs and is therefore less sensitive, or as suggested earlier, the trimmed mean regional SUVR itself may not reflect border changes particularly well.

The CoV_r was also evaluated in the composite cortical region. After SMPVC, significantly lower CoV_r values were observed than with RBV. The lower values suggest a greater degree of uniformity in the ROIs that constitute the composite region. As both PVC techniques rely on the assumption of regional uniformity, this result implies that the PVC performed using SMPVC satisfies this assumption better than RBV. However, this does not necessarily mean that SMPVC more accurately represents the *true* PET distribution.

SMPVC modifies the parcellation to fit the PET data. Changes to the parcellation were assessed in this dataset to see whether the SMPVC consistently modifies the parcellation. The observed percentage volume change in each region was calculated between the original parcellation and the SMPVC parcellation. Large changes were observed, for example, in the frontal region, where the boundary of the inferior frontal region would move into the superior frontal region. This pattern was consistent across the five subjects, for both scans.

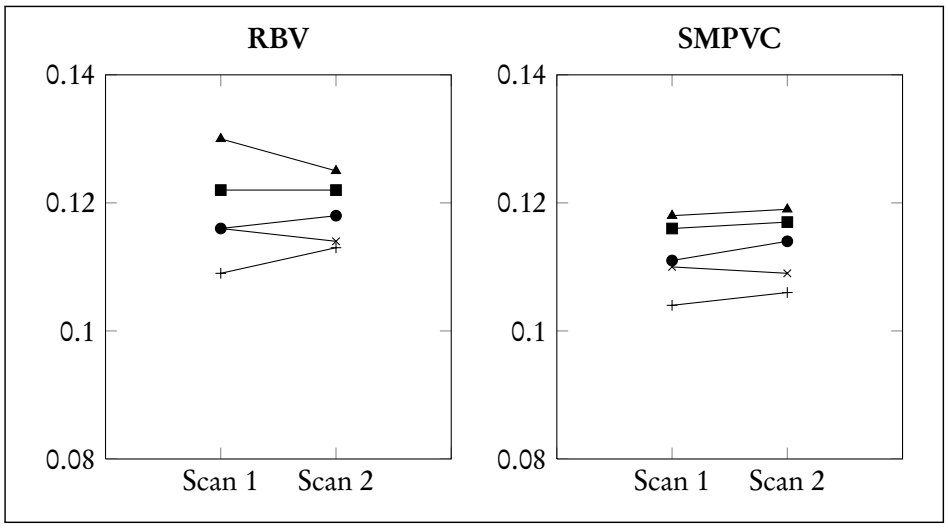


Figure 5.11: The CoV_r of the composite cortical region for both scans, after RBV correction (left) and SMPVC (right).

Some regions exhibited volume changes of close to 20%, these may have been restricted by the ±20% original volume criterion applied to SMPVC. Alteration or removal of this criterion would probably have resulted in a different parcellation and therefore different volume changes.

The DC was computed for all regions between the SMPVC parcellations of the first and second scan, providing a measure of similarity for each region. High DC values were obtained in the majority of regions for all subjects. It should however be noted that in this instance, both parcellations begin with $DC = 1$ in all regions, as SMPVC operates on both scans using the same parcellation. Reductions in the DC value therefore imply that the similarity has reduced. Given that two scans will always have different noise levels, it is unlikely that perfect similarity would be achieved as the algorithm operates on the observed PET data. The DC values obtained for these five subjects suggest that the SMPVC parcellations are similar and reproducible. These findings should be reproduced in a larger cohort with dual scans.

In summary, the performance of SMPVC has been compared to RBV. The further reductions in regional variablility using SMPVC suggests that fitting the parcellation according to the observed PET data improves PVC. The parcellations of the subject scans are shown to be similar and themselves reproducible. Parcellation changes may provide useful additional parameters. While this study has shown that SMPVC reduces the regional variance across a set of regions, the sample size ($n = 5$) is small. In addition, this dataset contains only AD subjects and the technique should be applied to subjects with low amyloid burden. SMPVC should be validated in a larger cohort to demonstrate the improved PVC.

	Subject									
	1		2		3		4		5	
Region	Scan 1	Scan 2	Scan 1	Scan 2	Scan 1	Scan 2	Scan 1	Scan 2	Scan 1	Scan 2
Left sup. Frontal	-7.48 %	-5.30 %	-10.95 %	-8.14 %	-6.97 %	-6.88 %	-8.71 %	-7.42 %	-6.66 %	-6.74 %
Left inf. Frontal	18.68 %	13.36 %	17.97 %	15.04 %	19.17 %	18.99 %	18.40 %	19.82 %	17.72 %	16.96 %
Right sup. Frontal	-8.56 %	-8.39 %	-10.32 %	-7.94 %	-8.62 %	-11.20 %	-6.66 %	-6.66 %	-6.50 %	-6.44 %
Left inf. Frontal	19.38 %	19.04 %	19.11 %	19.48 %	16.84 %	16.90 %	16.38 %	16.42 %	18.24 %	18.06 %
Left mid. Temporal	10.59 %	10.59 %	16.24 %	15.19 %	13.25 %	13.32 %	12.08 %	10.57 %	6.72 %	7.90 %
Right mid. Temporal	10.21 %	10.55 %	16.28 %	17.21 %	13.51 %	8.25 %	9.22 %	9.41 %	9.47 %	11.49 %

Table 5.7: The percentage volume change observed in selected cortical regions after SMPVC was performed.

	Subject				
Region	1	2	3	4	5
Left sup. frontal	0.988	0.974	0.997	0.989	0.956
Left inf. frontal	0.977	0.979	0.995	0.991	0.913
Right sup. frontal	0.997	0.982	0.983	0.998	0.997
Right inf. frontal	0.995	0.993	0.996	0.995	0.993
Left central	0.994	0.945	1.000	0.973	0.949
Right central	0.997	0.954	0.956	0.990	0.998
Left precuneus	0.969	0.973	1.000	0.960	0.952
Right precuneus	0.942	0.980	0.942	0.980	0.985
Left inf. parietal	0.978	0.988	0.997	0.961	0.937
Right inf. parietal	0.967	0.998	0.949	0.952	0.975
Left sup. parietal	0.983	0.962	0.999	0.959	0.959
Right sup. parietal	0.979	0.973	0.950	0.937	0.985
Left sup. temporal	0.997	0.995	0.998	0.969	0.999
Right sup. temporal	0.972	0.996	0.968	0.877	0.972
Left mid. temporal	0.991	0.993	0.998	0.953	0.988
Right mid. temporal	0.972	0.996	0.942	0.944	0.971
Left inf. temporal	0.970	0.990	1.000	0.968	0.950
Right inf. temporal	0.966	0.998	0.909	0.962	0.984
Left occipital	0.957	0.981	1.000	0.972	0.902
Right occipital	0.950	0.997	0.878	0.957	0.981

Table 5.8: The Dice Coefficient (DC) for all moving regions after SMPVC. The DC is calculated between the SMPVC parcellations of the subject scans.

5.5 Longitudinal assessment of amyloid burden

This section reports the evaluation of RBV and SMPVC in a clinical longitudinal cohort using amyloid tracer: [^{11}C]PIB. The aim of this study is to apply PVC to the dataset and investigate whether PVC may improve the discrimination of MCI-converters (those who progress from a diagnosis of MCI to AD) from MCI-non-converters.

5.5.1 Materials and methods

5.5.1.1 Subjects

10 MCI subjects from the Australian Imaging Biomarkers and Lifestyle flagship study of ageing (AIBL) database [Ellis et al., 2009] were used as the basis for this study. All subjects were diagnosed as MCI at baseline. 5 subjects converted from MCI to AD (MCI-converters) within 18 months, with the other 5 continuing to be classified as MCI (MCI-non-converters). All subjects underwent a [^{11}C]PIB PET scan and T1-weighted MRI at baseline. Both PET and MRI were also performed at 18 months follow-up.

5.5.1.2 PET acquisition

The PET data was acquired using a Phillips Allegro. 6×5 minute frames were acquired 40 to 70 minutes post-injection. The PET frames were then summed using the 50 to 70 minute data to adhere to the ADNI protocol, as the AIBL database has been made available as part of ADNI. All PET data was reconstructed using 3D Row action Maximum-Likelihood algorithm (RAMLA) iterative reconstruction. The reconstructed resolution was estimated as 5.4mm FWHM [O'Keefe, 2011].

5.5.1.3 Image processing

Region definition: The subject MRI data for both timepoints was processed using FreeSurfer (FS), in an identical manner to that described in the previous study (section 4.2.1.3, page 99). Briefly, the MRI was parcellated using FS and then the labels were merged into a mask image of 34 regions; 30 GM and 4 non-GM regions.

Registration: The subject summed [^{11}C]PIB image was rigidly registered to T1 MRI associated with that timepoint. This was performed using the registration technique of Ourselin et al. [2001]. Both PET-MR registrations of the baseline and follow-up scans were performed independently of each other. All subsequent processing was performed in the native MRI space.

Partial volume correction: PVC was performed on the registered PET data using the parcellated regions derived from the MRI. A space-invariant 3D PSF with a 5.4mm FWHM was used when applying RBV and SMPVC. As in the previous study, the regions in the parcel-

lation that could be modified were in cortical GM only. This meant that 10 of the 30 GM regions were static when SMPVC was applied. Additionally, modifications were again restricted to their respective hemispheres to ensure sensible region definitions were maintained.

PET normalisation: The brain stem was used as a reference region to normalise the PET data. A 10% trimmed mean value was calculated for the brain stem region. This was used to normalise the PET data, creating SUVR images. Image analysis was then performed using the SUVR images.

5.5.1.4 Analysis

SUVR values were calculated, using a 10% trimmed mean value, for all ROIs after RBV correction and SMPVC. Regional analysis was performed for both the MCI-converters and MCI-non-converters, between the timepoints and for both correction methods. The results of the SUVR analysis for the composite cortical region are reported. The SUVR values are reported for baseline and follow-up for both MCI groups, using both PVC methods. The percentage volume change observed with SMPVC for each timepoint is also reported for selected cortical regions of both groups. Finally, the percentage volume change between timepoints (follow-up - baseline), referred to as *functional volume change*, is reported for both groups, using both correction methods.

5.5.2 Results

The results of the SUVR analysis on the composite cortical region can be seen in figure 5.12. Four of the five subjects in the MCI-converters group exhibit high SUVR values for both timepoints, while lower SUVR values were observed in one subject. The opposite was observed in the MCI-non-converters. The SUVR values after RBV correction and SMPVC are very similar. The values in the composite region after RBV correction and SMPVC were tested with paired, two-tailed Wilcoxon signed-rank tests for both groups. Neither were significant at the $p < 0.05$ threshold (MCI-non-converters: MDD=0.03; MCI-converters: MDD=0.03).

The SUVR values for MCI-converters in the frontal, central and precuneus can be seen in table 5.9 for RBV-corrected data and table 5.10 for SMPVC. For MCI-non-converters, tables 5.11 and 5.12 contain the SUVR values after RBV correction and SMPVC respectively. MCI-non-converters subjects tend to have a lower SUVR than that of the MCI-converters. Some of subjects in both groups exhibit apparent increases in SUVR between baseline and follow-up. However, when baseline and follow-up were tested across groups (for both correction methods) using Wilcoxon signed-rank tests no region reached significance at the $p < 0.05$ threshold.

The percentage volume change from the original region to the region found by SMPVC is reported for both timepoints in tables 5.13 (MCI-converters) and 5.14 (MCI-non-converters). Reductions in the superior frontal regions, coupled with increases in the inferior frontal region was observed in both subject groups after SMPVC was performed. All subjects exhibited an increased SUVR value in the precuneus after SMPVC. The SUVR values were tested between correction methods using a Wilcoxon signed-rank test. This was significant at $p < 0.05$.

The functional volume change is reported for the inferior and superior frontal region, central region and precuneus. RBV-corrected data can be seen in tables 5.15 and 5.16 for the MCI-converters and MCI-non-converters respectively. The data corrected with SMPVC are given in tables 5.17 (MCI-converters) and 5.18 (MCI-non-converters). The mean change across subjects was calculated for each region. This was averaged across the eight regions to obtain a group mean ($\pm$ SD). The regional mean values of the MCI-converters and MCI-non-converters were tested using an unpaired, two-tailed Wilcoxon rank-sum test for both correction methods. The difference between groups using RBV correction was significant ($p = 0.03$) at the $p < 0.05$ threshold. However, no significant difference ($p = 0.48$, MDD=2.6%) was found between the two groups when performing SMPVC. For both correction techniques, larger functional volume changes were observed in the MCI-non-converters group (RBV: -3.1% $\pm$ 0.6; SMPVC: -2.1% $\pm$ 1.6) compared to the MCI-converters (RBV: -1.3% $\pm$ 2.0; SMPVC: -1.3% $\pm$ 2.2).

5.5.3 Discussion

This study has assessed PVC in a longitudinal dataset. 5 MCI-converters and 5 MCI-non-converters were assessed at baseline and 18 month follow-up. Both RBV correction and SMPVC were evaluated. Each group had one subject that appeared to be an outlier. Subject 2 in the MCI-converters group exhibited low cortical uptake, similar to that observed in the MCI-non-converters. Conversely, subject 2 in the MCI-non-converters group exhibited high cortical uptake. The high MCI-non-converter may well be close to conversion, although coginitively normal subjects are often observed with high amyloid burden. Equally, a proportion of AD subjects also exhibit uptake that would usually be classed as normal, which could account for the low uptake observed in one subject that had converted.

The difference in SUVR values across the time points tends to become more similar after SMPVC, although this similarity did not reach significance in either the cortical composite region or in other cortical regions. The same behaviour was observed during the assessment of the reproducibility data (section 5.4 page 142). It was anticipated that SUVR values would increase over time, but this was not observed consistently in this study. This may be due

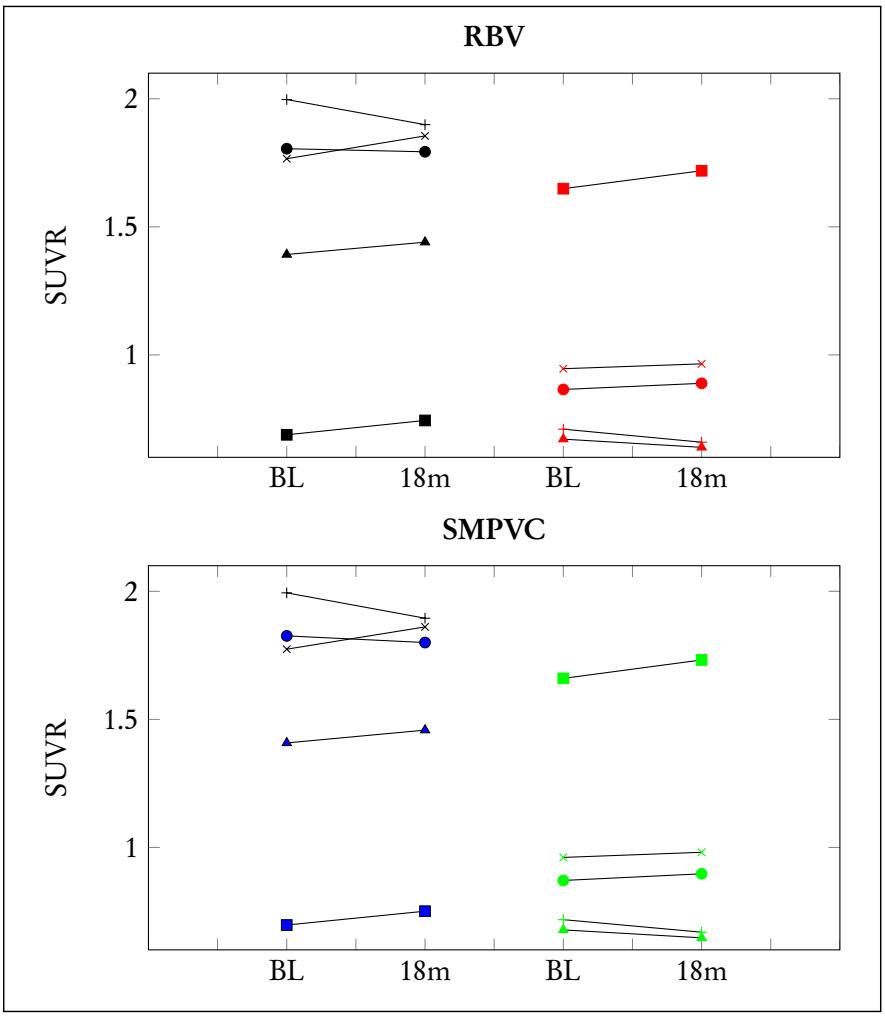


Figure 5.12: The mean SUVR of the composite cortical region for both scans, after RBV correction (top) and SMPVC (bottom). The black markers denote MCI-converters after RBV correction; MCI-converters after SMPVC are shown in blue. The red makers are the MCI-non-converters after RBV correction; MCI-non-converters after SMPVC are shown in green. Values are shown for baseline (BL) and 18 month follow-up (18m).

to the scanning interval. 18 months may not be sufficient to observe noticable changes, although Villemagne et al. [2011a] reported an average 2.1% increase in MCI subjects with a scanning interval of 20 ± 3 months. Further timepoints would be advantageous for this analysis although these data are not currently available. In addition, the previous study suggested that the regional SUVR may not be particularly sensitive, even when the regions are modified using SMPVC.

The percentage volume change after SMPVC was also evaluated. As was observed in the reproducibility study, the inferior frontal region tends to extend into the superior frontal region. This was seen in all subjects and suggests that the amyloid distribution is not accurately partitioned by a superior/inferior boundary. It should also be noted that the frontal region often exhibits high activity in diseased subjects. The tracking of this boundary could

be useful as movement may represent the initial spread of amyloid from inferior to superior cortical regions as described by Braak and Braak [1991].

The precuneus was identified as a region where SUVR increased after SMPVC. The precuneus often exhibits high amyloid burden in diseased subjects, but has also be observed in up to 30% of HC subjects [Rowe et al., 2007]. In terms of SUVR, all subjects exhibited a small but significant increase in the precuneus after SMPVC correction. This suggests modification to the parcellation consistently increases SUVR values in the precuneus.

Functional volume change was assessed in both groups using RBV correction and SMPVC. Volume changes were overall negative, indicating a volume loss over time which might logically be attributed to atrophy. Larger changes were observed in the MCI-non-converters than MCI-converters. However, the sample size in this study is low and therefore caution should be taken when drawing conclusions about the observed group differences.

The difference between the two groups was significant after RBV correction and not significant when applying SMPVC. This could suggest that SMPVC is less sensitive to volumetric changes over time, but equally the volumes defined by SMPVC are governed by the PET distribution. Therefore, after SMPVC, the functional volume change reflects atrophy over time but also volume changes in each timepoint which reduce variability in terms of SUVR. This makes the interpretation of functional volume change less clear than a pure volumetric change. The potential for a functional volume change parameter provided by SMPVC requires more extensive study. Restricting the SMPVC algorithm to allow for the assessment of individual surface changes rather than multiple surfaces, may improve the interpretability of functional volume change.

Several of the statistical comparisons in this study did not reach significance. When evaluating the similarity of the composite cortical SUVR values between the RBV-corrected and SMPVC data, neither group demonstrated a significant difference. The MDD was calculated as 0.03 for both comparisons. However, the mean difference between methods was 0.01. When testing between SUVR values between baseline and 18 month follow-up, no significant difference was observed in either group, with RBV correction or SMPVC. For example, the MDD in the left precuneus for the MCI-converters after RBV correction was 0.06. The average increase in SUVR was 0.04. Similarly, no significance was found between groups when assessing functional volume change after SMPVC was applied. The MDD was calculated to be 2.6%, whereas the mean group difference was found to be 0.8%. The small sample size in each subject group ($n = 5$) has likely resulted in under-powered statistical analyses.

In summary, this study has evaluated the performance of RBV and SMPVC in a clinical

longitudinal dataset. Increases in SUVR over time were not observed in this study and is probably down to sample size. Large volume changes were seen between the superior and inferior frontal regions. SUVR changes were also observed in the precuneus with SMPVC that was not observed in RBV, and could be an indicator of subtle changes due to the modification of the parcellation, although this cannot be proven in this study. Further investigation using a larger dataset, with multiple timepoints is necessary to see if SMPVC improves detection of subtle changes.

	Subject									
	1		2		3		4		5	
Region	BL	18m	BL	18m	BL	18m	BL	18m	BL	18m
Left sup. frontal	1.95	1.92	0.62	0.67	1.87	1.95	2.16	2.05	1.79	1.71
Left inf. frontal	2.01	1.93	0.74	0.84	1.82	1.98	2.08	2.07	1.59	1.47
Right sup. frontal	2.02	1.95	0.66	0.75	2.06	2.16	2.09	1.97	1.80	1.77
Right inf. frontal	2.04	2.00	0.75	0.85	2.05	2.20	2.06	1.88	1.78	1.73
Left central	1.58	1.49	0.65	0.71	1.45	1.52	1.84	1.72	1.08	1.17
Right central	1.55	1.49	0.68	0.82	1.60	1.62	1.78	1.66	1.14	1.22
Left precuneus	2.06	2.08	0.67	0.74	1.93	2.05	2.25	2.16	1.67	1.77
Right precuneus	2.00	2.07	0.72	0.63	1.95	2.18	2.18	2.13	1.73	1.76

Table 5.9: SUVR values observed in MCI-converters for selected cortical regions with RBV correction.

	Subject									
	1		2		3		4		5	
Region	BL	18m	BL	18m	BL	18m	BL	18m	BL	18m
Left sup. frontal	1.96	1.92	0.63	0.67	1.87	1.95	2.16	2.06	1.80	1.74
Left inf. frontal	2.00	1.93	0.74	0.84	1.82	1.98	2.08	2.04	1.59	1.48
Right sup. frontal	2.04	1.96	0.66	0.75	2.05	2.15	2.09	1.97	1.83	1.77
Right inf. frontal	2.04	2.00	0.76	0.84	2.06	2.21	2.07	1.89	1.76	1.72
Left central	1.58	1.49	0.66	0.73	1.45	1.53	1.84	1.72	1.09	1.18
Right central	1.55	1.50	0.69	0.83	1.61	1.61	1.77	1.64	1.15	1.21
Left precuneus	2.12	2.13	0.68	0.75	1.96	2.05	2.25	2.15	1.72	1.81
Right precuneus	2.03	2.08	0.73	0.65	1.96	2.18	2.18	2.14	1.80	1.90

Table 5.10: SUVR values observed in MCI-converters for selected cortical regions with SMPVC.

	Subject									
	1		2		3		4		5	
Region	BL	18m	BL	18m	BL	18m	BL	18m	BL	18m
Left sup. frontal	1.18	1.20	1.78	1.92	1.17	1.10	0.70	0.68	0.65	0.65
Left inf. frontal	1.09	1.24	1.74	1.83	1.09	1.12	0.73	0.66	0.74	0.72
Right sup. frontal	1.20	1.18	1.71	1.75	1.28	1.16	0.79	0.73	0.69	0.63
Right inf. frontal	1.33	1.38	1.72	1.81	1.11	1.20	0.75	0.75	0.80	0.72
Left central	0.74	0.72	1.45	1.54	0.81	0.74	0.73	0.67	0.67	0.67
Right central	0.77	0.74	1.41	1.40	0.90	0.78	0.77	0.70	0.72	0.69
Left precuneus	0.91	0.86	2.07	2.19	1.27	1.32	0.61	0.57	0.63	0.66
Right precuneus	0.75	0.74	2.06	1.99	1.26	1.37	0.67	0.59	0.62	0.57

Table 5.11: SUVR values observed in MCI-non-converters for selected cortical regions with RBV correction.

	Subject									
	1		2		3		4		5	
Region	BL	18m	BL	18m	BL	18m	BL	18m	BL	18m
Left sup. frontal	1.19	1.20	1.80	1.92	1.17	1.11	0.71	0.68	0.65	0.66
Left inf. frontal	1.09	1.26	1.74	1.84	1.09	1.14	0.74	0.68	0.75	0.73
Right sup. frontal	1.19	1.18	1.71	1.77	1.28	1.17	0.80	0.73	0.69	0.64
Right inf. frontal	1.32	1.37	1.72	1.81	1.13	1.21	0.77	0.76	0.80	0.73
Left central	0.74	0.72	1.46	1.54	0.83	0.75	0.74	0.68	0.68	0.68
Right central	0.78	0.75	1.42	1.42	0.92	0.79	0.78	0.71	0.74	0.70
Left precuneus	0.91	0.90	2.09	2.22	1.33	1.37	0.62	0.59	0.64	0.67
Right precuneus	0.77	0.76	2.11	2.06	1.31	1.42	0.68	0.61	0.63	0.58

Table 5.12: SUVR values observed in MCI-non-converters for selected cortical regions with SMPVC.

	Subject									
	1		2		3		4		5	
Region	BL	18m	BL	18m	BL	18m	BL	18m	BL	18m
Left sup. frontal	-5.8%	-6.2%	-4.9%	-6.2%	-4.6%	-4.7%	0.0%	-6.2%	-3.0%	-5.4%
Left inf. frontal	16.1%	17.1%	11.4%	15.7%	13.3%	13.4%	0.0%	16.1%	9.5%	0.5%
Right sup. frontal	-3.6%	-2.5%	-10.7%	-6.0%	-4.5%	-6.2%	-6.2%	-6.3%	-10.6%	-5.4%
Right inf. frontal	0.0%	0.0%	17.6%	14.6%	10.6%	16.0%	16.0%	16.9%	18.2%	15.3%
Left central	-3.1%	-3.4%	-1.7%	-3.3%	-2.7%	-2.8%	0.0%	-2.9%	-2.1%	14.6%
Right central	12.3%	8.2%	1.3%	-3.1%	-2.8%	-4.2%	-4.6%	-5.4%	9.7%	-4.1%
Left precuneus	-8.7%	-6.1%	0.0%	0.0%	-3.7%	-1.4%	0.0%	0.0%	-5.8%	-8.0%
Right precuneus	-2.5%	-2.0%	-4.9%	-3.4%	-3.3%	-1.4%	13.7%	17.6%	-12.7%	-17.8%

Table 5.13: The percentage volume change observed in MCI-converters selected cortical regions after SMPVC was performed.

	Subject									
	1		2		3		4		5	
Region	BL	18m	BL	18m	BL	18m	BL	18m	BL	18m
Left sup. frontal	-7.3%	-5.8%	-5.4%	-3.1%	0.0%	-1.5%	-5.8%	-5.4%	-6.5%	0.0%
Left inf. frontal	20.0%	15.8%	1.2%	9.3%	0.0%	-1.2%	17.1%	15.2%	19.5%	0.0%
Right sup. frontal	-6.4%	-8.0%	-9.4%	-6.1%	-4.7%	-2.8%	-3.7%	-6.5%	-6.3%	-8.1%
Right inf. frontal	16.9%	16.5%	18.9%	-0.1%	9.8%	6.1%	8.1%	15.3%	13.9%	18.7%
Left central	-4.4%	-3.6%	13.5%	-2.1%	-2.1%	6.2%	-2.6%	-4.5%	-3.7%	0.0%
Right central	-4.1%	3.4%	-0.2%	18.8%	-2.1%	-1.4%	-2.3%	-3.2%	-7.0%	-4.3%
Left precuneus	0.0%	-6.4%	-5.3%	-5.0%	-8.4%	-8.2%	0.0%	-4.0%	0.0%	0.0%
Right precuneus	-1.6%	3.4%	-12.3%	-16.4%	-6.9%	-7.4%	-7.6%	-7.3%	-12.5%	-2.6%

Table 5.14: The percentage volume change observed in MCI-non-converters selected cortical regions after SMPVC was performed.

	Subject						
Region	1	2	3	4	5	**Reg. mean**	**Reg. SD**
Left sup. frontal	2.8 %	2.6 %	-0.7 %	-0.9 %	-4.4 %	-0.1 %	3.0 %
Left inf. frontal	3.6 %	-4.8 %	-3.2 %	-3.9 %	-2.0 %	-2.1 %	3.3 %
Right sup. frontal	2.8 %	6.8 %	-1.7 %	-0.4 %	0.1 %	1.5 %	3.4 %
Right inf. frontal	5.1 %	-11.7 %	-6.1 %	-6.2 %	-4.0 %	-4.6 %	6.1 %
Left central	4.1 %	-1.1 %	-3.6 %	0.6 %	0.2 %	0.0 %	2.8 %
Right central	1.4 %	3.1 %	-6.5 %	-3.5 %	4.6 %	-0.2 %	4.7 %
Left precuneus	-1.0 %	0.8 %	-3.8 %	-6.8 %	-3.4 %	-2.9 %	2.9 %
Right precuneus	-8.4 %	0.1 %	2.0 %	-2.1 %	-3.7 %	-2.4 %	4.0 %
Subject mean	1.3 %	-0.5 %	-3.0 %	-2.9 %	-1.6 %	**-1.3**%	
Subject SD	4.3 %	5.6 %	2.8 %	2.7 %	3.1 %	**2.0**%	

Table 5.15: Functional volume change in MCI-converters with RBV correction.

	Subject						
Region	1	2	3	4	5	**Reg. mean**	**Reg. SD**
Left sup. frontal	-8.0 %	-5.2 %	-0.6 %	0.3 %	-1.0 %	-2.9 %	3.5 %
Left inf. frontal	-7.4 %	-3.3 %	0.2 %	-3.3 %	-3.3 %	-3.4 %	2.7 %
Right sup. frontal	-6.8 %	-3.5 %	-7.3 %	-0.3 %	-1.8 %	-3.9 %	3.1 %
Right inf. frontal	-5.5 %	-7.3 %	-7.4 %	1.4 %	-0.6 %	-3.9 %	4.1 %
Left central	-6.5 %	-3.1 %	1.5 %	-1.0 %	-2.1 %	-2.3 %	2.9 %
Right central	-6.1 %	-0.8 %	1.3 %	-4.4 %	-2.6 %	-2.5 %	2.9 %
Left precuneus	-6.6 %	-2.8 %	0.8 %	-4.5 %	-3.0 %	-3.2 %	2.7 %
Right precuneus	-6.7 %	2.5 %	-2.1 %	2.2 %	-8.4 %	-2.5 %	5.0 %
Subject mean	-6.7 %	-2.9 %	-1.7 %	-1.2 %	-2.8 %	**-3.1**%	
Subject SD	0.8 %	2.9 %	3.7 %	2.6 %	2.4 %	**0.6**%	

Table 5.16: Functional volume change in MCI-non-converters with RBV correction.

	Subject						
Region	1	2	3	4	5	**Reg. mean**	**Reg. SD**
Left sup. frontal	2.4 %	1.2 %	-0.8 %	-7.0 %	-3.8 %	-1.6 %	3.8 %
Left inf. frontal	4.6 %	-1.0 %	-3.1 %	11.6 %	-4.3 %	1.6 %	6.6 %
Right sup. frontal	3.3 %	12.4 %	-3.5 %	-0.5 %	0.3 %	2.4 %	6.1 %
Right inf. frontal	5.1 %	-13.9 %	-1.5 %	-5.5 %	-4.2 %	-4.0 %	6.9 %
Left central	3.9 %	-2.8 %	-3.7 %	-2.3 %	0.8 %	-0.8 %	3.1 %
Right central	-1.9 %	-1.4 %	-7.8 %	-4.2 %	-0.8 %	-3.2 %	2.9 %
Left precuneus	1.8 %	0.8 %	-1.5 %	-6.8 %	-5.7 %	-2.3 %	3.8 %
Right precuneus	-8.0 %	1.7 %	4.0 %	1.3 %	-9.3 %	-2.1 %	6.1 %
Subject mean	1.4 %	-0.4 %	-2.2 %	-1.7 %	-3.4 %	**-1.3**%	
Subject SD	4.4 %	7.2 %	3.3 %	6.1 %	3.4 %	**2.2**%	

Table 5.17: Functional volume change in MCI-converters with SMPVC.

	Subject						
Region	1	2	3	4	5	**Reg. mean**	**Reg. SD**
Left sup. frontal	-6.4 %	-2.9 %	-2.2 %	0.7 %	0.6 %	-2.0 %	2.9 %
Left inf. frontal	-10.6 %	4.4 %	-1.0 %	-1.5 %	-1.6 %	-2.1 %	5.4 %
Right sup. frontal	-7.3 %	0.1 %	-5.5 %	-1.0 %	-1.6 %	-3.1 %	3.2 %
Right inf. frontal	-5.8 %	-2.1 %	-10.6 %	3.2 %	1.1 %	-2.8 %	5.5 %
Left central	-5.7 %	-1.5 %	2.5 %	-2.0 %	1.6 %	-1.0 %	3.3 %
Right central	1.3 %	0.9 %	2.1 %	-5.2 %	0.3 %	-0.1 %	2.9 %
Left precuneus	-12.6 %	-2.5 %	1.0 %	-8.3 %	-3.0 %	-5.1 %	5.4 %
Right precuneus	-2.1 %	-2.3 %	-2.6 %	2.6 %	2.0 %	-0.5 %	2.5 %
Subject mean	-6.2 %	-0.7 %	-2.0 %	-1.4 %	-0.1 %	**-2.1**%	
Subject SD	4.4 %	2.5 %	4.4 %	3.8 %	1.8 %	**1.6**%	

Table 5.18: Functional volume change in MCI-non-converters with SMPVC.

5.6 Conclusion

This chapter has described a novel PVC technique, SMPVC, that modifies the MRI-based parcellation to fit the observed PET data. Initially, the parcellation modification in SMPVC was formulated as a voxel-based operation. This was later adapted to a surface-based approach, which proved to be more resilient to noise. The performance of SMPVC has been compared to RBV in a digital phantom study, clinical reproducibility data and a longitudinal study of MCI subjects.

In the phantom study, SMPVC was demonstrated to reduce bias and regional variability compared to performing RBV with known parcellation errors. When applied to clinical reproducibility data, SMPVC was shown to reduce the CoV_r, while maintaining the reproducibility of regional SUVR values. A further preliminary investigation using longitudinal MCI data was undertaken. Initial results indicate that SMPVC may provide volumetric indicators based on activity in the PET data. Evaluation of SMPVC in a large clinical cohort, such as ADNI or AIBL, is a necessary step to fully validate the technique.

Chapter 6

Conclusions

6.1 Summary and conclusions

The work in this thesis has concentrated on the development and application of PVC techniques for quantitative amyloid brain PET imaging. PVEs are known to induce large errors in quantification, but corrections are not commonly applied in clinical practice. While PVC methods have been proposed [Müller-Gärtner et al., 1992, Rousset et al., 1998a, Teo et al., 2007] for the purpose of correcting brain PET, each have limitations and/or make inappropriate assumptions about the PET distribution. The focus of this work has been on the development of techniques to overcome these issues.

A summary of the chapters of this thesis, and the main conclusions drawn from each, are given below:

Chapter 3: Evaluating partial volume correction techniques using phantom data

The first stage of this project was to compare existing PVC techniques that are applied in brain PET to a modified PVC approach; RBV correction. Phantom data were created that represent typical AD and control amyloid PET distributions. These data were corrected using three techniques, two anatomy-based approaches and one purely data-driven. The RBV and MG correction both rely on parcellated MRI data. In this study, the parcellation used for PVC was identical to the one used to generate the data.

Near perfect recovery was observed for all brain regions when performing RBV correction and demonstrates that the method is very accurate when the parcellation and resolution are accurately measured. The MG and mMG (a variation on the MG method) achieved good recovery in cortical regions, although their performance was inferior to that of RBV. However, some GM regions, in particular the hippocampi exhibited a positive bias for AD-like distributions when applying the MG methods. This was due to the inability of the MG ap-

proaches to account for *within*-tissue variability in GM.

The data-driven approach evaluated in this study was the VC iterative deconvolution. The VC technique improved recovery of the phantom images, but was not as accurate as the anatomy-based approaches. This was attributed to premature termination of the iterative process. Early termination is necessary when applying the VC method due to noise amplification issues.

The second half of this chapter investigates the use of filters and the Huber prior as part of the VC deconvolution in order to reduce noise amplification. Digital phantom PET data were generated with realistic noise. These data were then deconvolved with VC, VC-smooth and VC-prior. VC-smooth consisted of pre-smoothing the image data before deconvolution. The Huber prior was incorporated into the VC deconvolution and is referred to as VC-prior. Both noise suppression techniques reduced noise compared to the standard VC. This did however tend to be associated with a reduction in recovery. Neither VC-smooth or VC-prior resulted in notable improvements compared to VC and suggests that there is a limit to the amount of recovery that can be achieved with the VC approach.

Chapter 4: Comparing partial volume correction techniques using clinical data

This chapter evaluates PVC in a clinical cohort. The first section reports the effects of PVC on the quantification of [^{18}F]flutemetamol. An evaluation of the quantitative reproducibility of [^{18}F]flutemetamol with and without PVC is reported in the second section. The third section investigates amyloid PET normalisation. WM variability in AD and control subjects is assessed and a comparison between [^{11}C]PIB and [^{18}F]flutemetamol is also reported. Alternative reference regions for image normalisation are also assessed as part of the comparison.

RBV and mMG correction were applied to a clinical [^{18}F]flutemetamol dataset consisting of control, MCI and AD subjects. Regional SUVR analysis was performed on the uncorrected and PV-corrected data. The key finding was that cortical GM regions are severely affected by the PVE. In addition, this study also confirmed *within*-tissue PVEs, as was observed in the phantom study of the previous chapter. Differences between RBV and mMG were again observed in the hippocampi. This suggests that when applying mMG, bias may occur due to *within*-tissue PVEs, which cannot be accounted for. RBV appears to be superior to mMG in this respect.

A subset of the AD subjects ($n = 5$) underwent a second [^{18}F]flutemetamol scan for the purposes of assessing quantitative reproducibility. Regional SUVR analysis was performed for both scans. Statistical analysis was performed on the SUVR values of both scans in a composite cortical region. No significant difference between scans was found for either the

uncorrected or RBV-corrected data, suggesting that [^{18}F]flutemetamol SUVR values are reproducible and that RBV correction does not degrade the reproducibility. The regional coefficient of variation (CoV_r) was also assessed. Reductions in the CoV_r were observed in the majority of cortical regions after RBV correction and suggest that regions are more uniform after PVC.

The third section reports an investigation into amyloid PET normalisation. After RBV correction, a significant difference in SUVR between the cerebellar and sub-cortical WM was found in both groups, although no difference was found between the subject groups. This finding suggests that while WM appears to be non-specific, variability exists between WM region. This variability has implications for the application of PVC techniques, such as MG, which assume that WM uptake is uniform and prompted the subsequent evaluation of alternative reference regions.

A comparison between [^{18}F]flutemetamol and [^{11}C]PIB was also carried out in 40 subjects from the MCI and AD groups. Regression analysis was performed between regional SUVR values for both tracers without PVC and with RBV correction. In addition, alternative reference regions were evaluated (cerebellar GM, sub-cortical WM and brain stem), with and without PVC. Strong agreement between the two tracers was observed in cortical GM regions for all three reference regions, with weaker agreement in WM regions. Brain stem normalisation produced the strongest agreement between tracers in GM for the PV-corrected data.

Chapter 5: Segmentation-modifying partial volume correction

This chapter reports the development and application of a novel PVC technique: SMPVC. The method iteratively modifies a MRI-derived parcellation to fit the observed PET data. The first section reports the algorithm development process. SMPVC is then evaluated using digital phantom data, [^{18}F]flutemetamol reproducibility data and a longitudinal [^{11}C]PIB dataset.

SMPVC was initially designed as a voxel-based approach where region labels in the parcellation would be changed on a voxel-by-voxel basis. The algorithm was later redesigned to operate using surfaces, as this was less sensitive to noise and computationally less expensive. As part of the surface-based approach, a further PVC technique called iterative Yang (iY) is described, which is similar to RBV.

SMPVC was first assessed using digital phantom data. Two phantoms were evaluated; a geometric phantom and an anthropomorphic brain phantom. Large reductions in CoV_r were observed for the geometric phantom after SMPVC. The accuracy of the parcellation

also increased after SMPVC. 10 instances of a typical AD subject [^{11}C]PIB distribution were generated to create the anthropomorphic brain phantom. Parcellation errors were introduced and SMPVC was applied. SMPVC reduced the under- and over-correction in recovery of the two brain regions and reductions in CoV_r were also observed. The relatively small changes in regional means suggests that this measurement may be fairly insensitive to border changes.

The technique was then assessed using the clinical [^{18}F]flutemetamol reproducibility dataset described in the preceding chapter. RBV and SMPVC were compared through regional analysis of SUVR and CoV_r. Additional analysis of the parcellation was reported in terms of a similarity measure: the Dice Coefficient (DC). The between-scan difference in SUVR values tended to reduce, however this reduction did not result in a significant difference between RBV and SMPVC. Consistently lower CoV_r values were observed in the composite cortical region when applying SMPVC compared to RBV, suggesting a higher degree of regional uniformity with SMPVC correction. The DC was calculated between the subject scans after SMPVC and high values were observed, implying that the parcellations were similar and reproducible.

Finally, a longitudinal dataset of MCI subjects was evaluated using RBV and SMPVC. 5 MCI-converters (those who convert from MCI to AD) and 5 MCI-non-converters were scanned with [^{11}C]PIB at baseline and at 18 months follow-up. Consistent increases over time in cortical SUVR values were not observed in this study, although this could be attributed to sample size. The percentage volume change after SMPVC was also assessed and consistent changes across timepoints were observed in some cortical regions, particularly in the frontal lobe. A new parameter of *functional volume change* is evaluated in this study. An overall trend towards reductions in volume over time was observed with RBV and SMPVC, however this parameter requires further assessment in a larger cohort.

6.1.1 Summary of significant findings

The significant findings of the thesis work can be summarised as follows:

- The evaluation of existing PVC techniques in phantom and clinical data revealed brain regions that can suffer from PVC-induced bias due to inappropriate assumptions about tracer distribution.
- A modified PVC technique (RBV) was adopted that is more robust than methods normally employed for brain PET PVC. White matter variability was identified through applying the correction. This variability has implications both for PVC and image normalisation.
- Alternative reference regions were investigated in two amyloid PET tracers. The brain stem, in combination with PVC, was found to result in the strongest agreement between tracers.
- A novel anatomy-based PVC technique (SMPVC) has been developed that reduces regional variability by modifying the anatomical parcellation to fit the observed PET data.

6.2 Future directions

To extend on the work undertaken in this thesis, it is necessary to fully validate the new methodology through assessing a large clinical longitudinal dataset. Multiple time points, covering several years are required to evaluate the ability to detect changes over time. These data are also key to concluding whether or not 'functional volume' changes are useful biomarkers. Additionally, it may be necessary to fine-tune the algorithm depending on factors such as scanning protocol or the type of reconstruction used. Improvements in the quantitative accuracy of brain PET have been demonstrated, and further validation and extension to the methods contained in this thesis could reveal valuable information in PET data that may otherwise be obscured by PVEs.

This work has focussed on amyloid PET tracers; namely [^{11}C]PIB and [^{18}F]flutemetamol. An important development would be to assess the methodology in other tracers. [^{18}F]FDG is currently the most commonly used PET ligand and it would therefore be advantageous to be able to perform the methods described here in [^{18}F]FDG data. In amyloid PET, uptake in cortical GM appears to 'plateau' during later stages of AD, while cognitive decline progresses. [^{18}F]FDG may provide additional data, although this would be severely affected by the PVE, hence the necessity to perform PVC.

In terms of the SMPVC algorithm itself, the current implementation relies on an initial parcellation and optimises the existing regions. Extending the method to sub-divide and merge volumes would be beneficial as the algorithm could then be independent of the parcellation. This independence could lead to developments such as an atlas-based parcellation and/or PVC without MRI data. PET PVC without the requirement for structural information would be a desirable development in order to apply techniques in a routine clinical environment.

While this thesis has centred on brain imaging, PVEs are a major source of quantification error in other PET, and indeed SPECT applications. For example, in oncology, PVEs may result in large errors in the measurement of tumour activity. In addition, the tumour can be poorly delineated on structural data, making existing approaches to PVC unsuitable. Modifying information from structural imaging to fit the observed PET data could improve quantification in this area. Similarly, cardiac images are affected by PVEs and the parcellation of cardiac images is non-trivial due to organ motion. Again, optimising the PVC in terms of the measured PET data, could be advantageous.

Appendix A

Additional figures

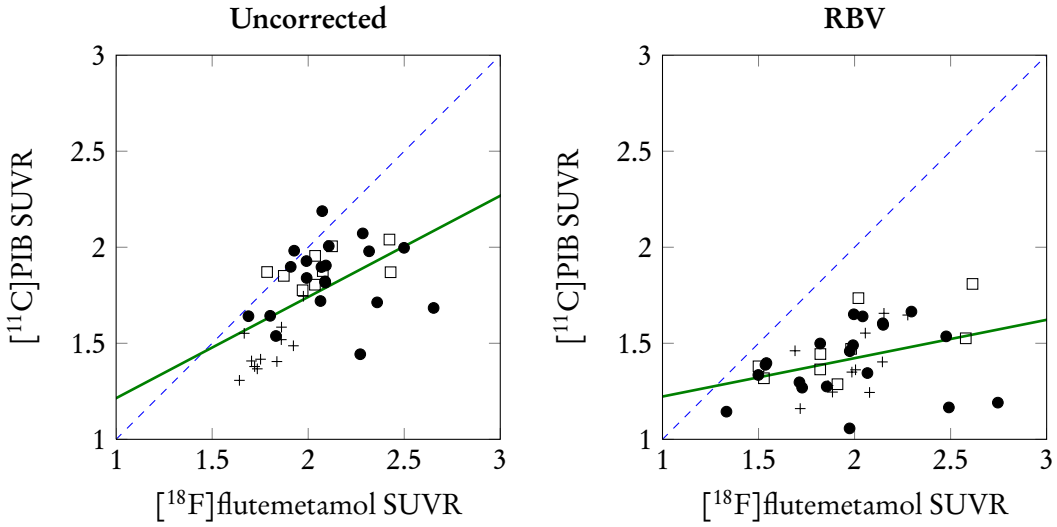


Figure A.1: Scatter plots of mean [^{18}F]flutemetamol SUVR and mean [^{11}C]PIB SUVR for the sub-cortical white matter region without PVC (left) and after RBV correction (right). AD (black circles), amyloid-positive (white squares) and amyloid-negative (black crosses). Cerebellar GM was used as the reference region. The solid line represents the line of best fit. The dashed line is the line of identity.

	RBV (error)	True	SMPVC
Region 1	0.987 (± 6.32e-03)	0.998 (± 6.19e-03)	0.981 (± 7.18e-03)
Region 2	0.993 (± 7.22e-03)	0.981 (± 6.37e-03)	0.984 (± 6.95e-03)
Region 3	0.991 (± 5.64e-03)	0.994 (± 6.38e-03)	0.997 (± 6.12e-03)

Table A.1: The RC (±SD) for PV-correction of the geometric phantom at 2.0e^6 counts.

	RBV (error)	True	SMPVC
Region 1	0.052 (± 1.20e-03)	0.031 (± 1.04e-03)	0.035 (± 1.29e-03)
Region 2	0.144 (± 2.24e-03)	0.047 (± 1.68e-03)	0.052 (± 1.26e-03)
Region 3	0.044 (± 9.26e-04)	0.033 (± 1.39e-03)	0.034 (± 1.19e-03)

Table A.2: The CoV_r (±SD) for PV-correction of the geometric phantom at 2.0e^6 counts.

	RBV (error)	True	SMPVC
Region 1	0.894	1.00	0.977 (± 3.42e-03)
Region 2	0.863	1.00	0.978 (± 4.01e-03)
Region 3	0.967	1.00	0.991 (± 2.68e-03)

Table A.3: The DC (±SD) for PV-correction of the geometric phantom at 2.0e^6 counts. Note that the DC for the RBV-corrected data, with and without parcellation error does not change as the same masks are used for each realisation.

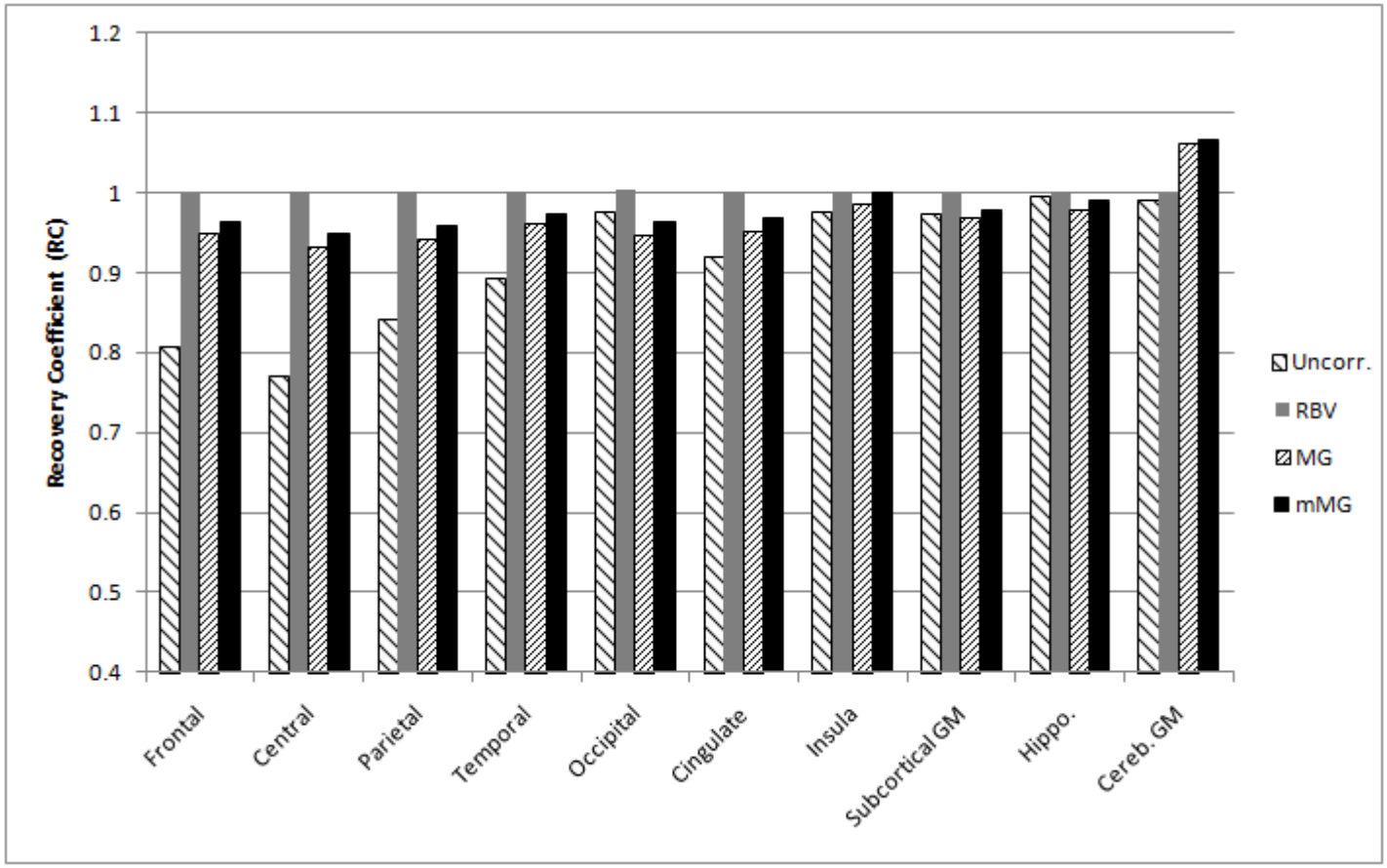


Figure A.2: The effects of partial volume correction on recovery of grey matter regions of the healthy control phantom with realistic WM variability.

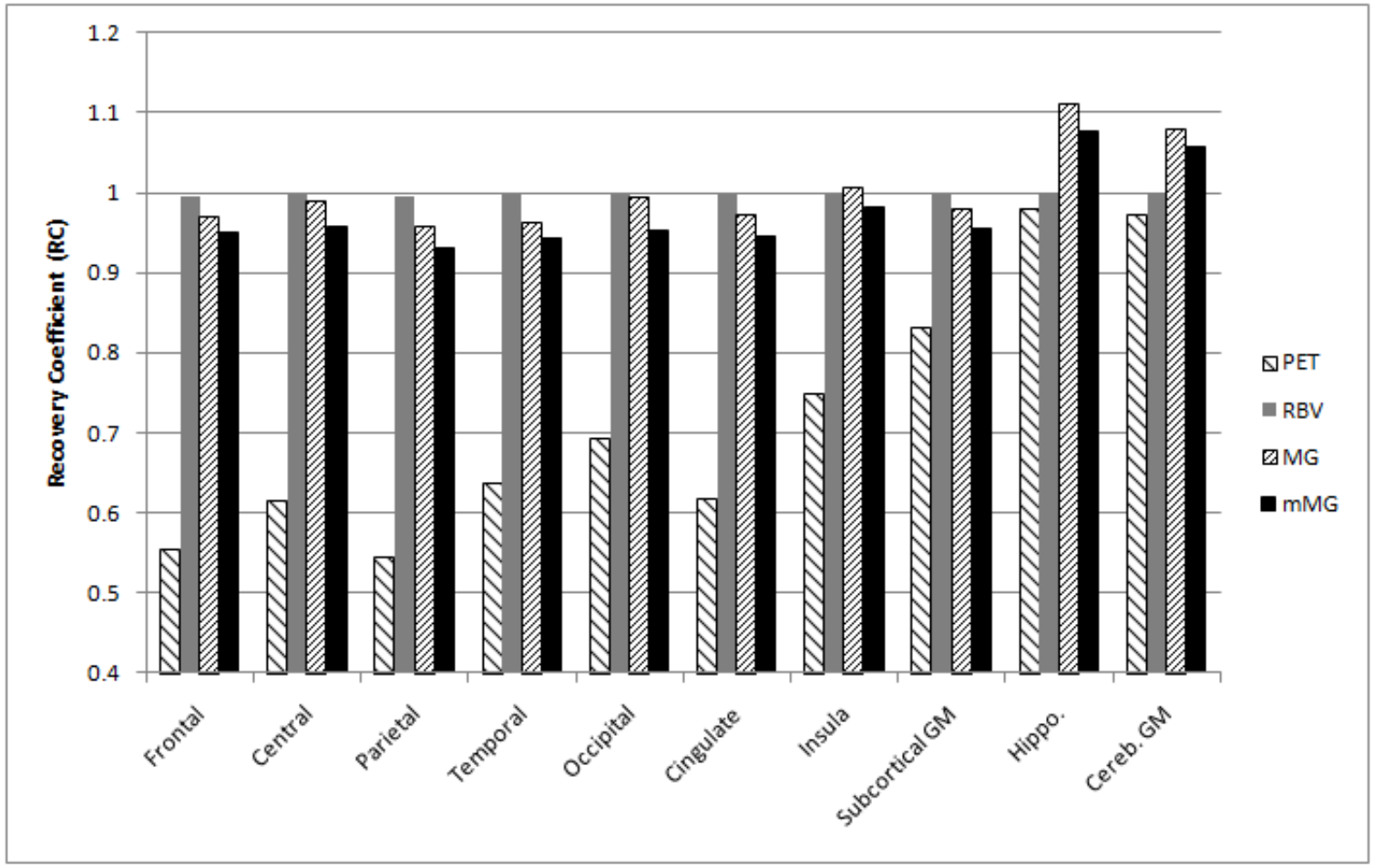


Figure A.3: The effects of partial volume correction on recovery of grey matter regions of the Alzheimer's disease phantom with realistic WM variability.

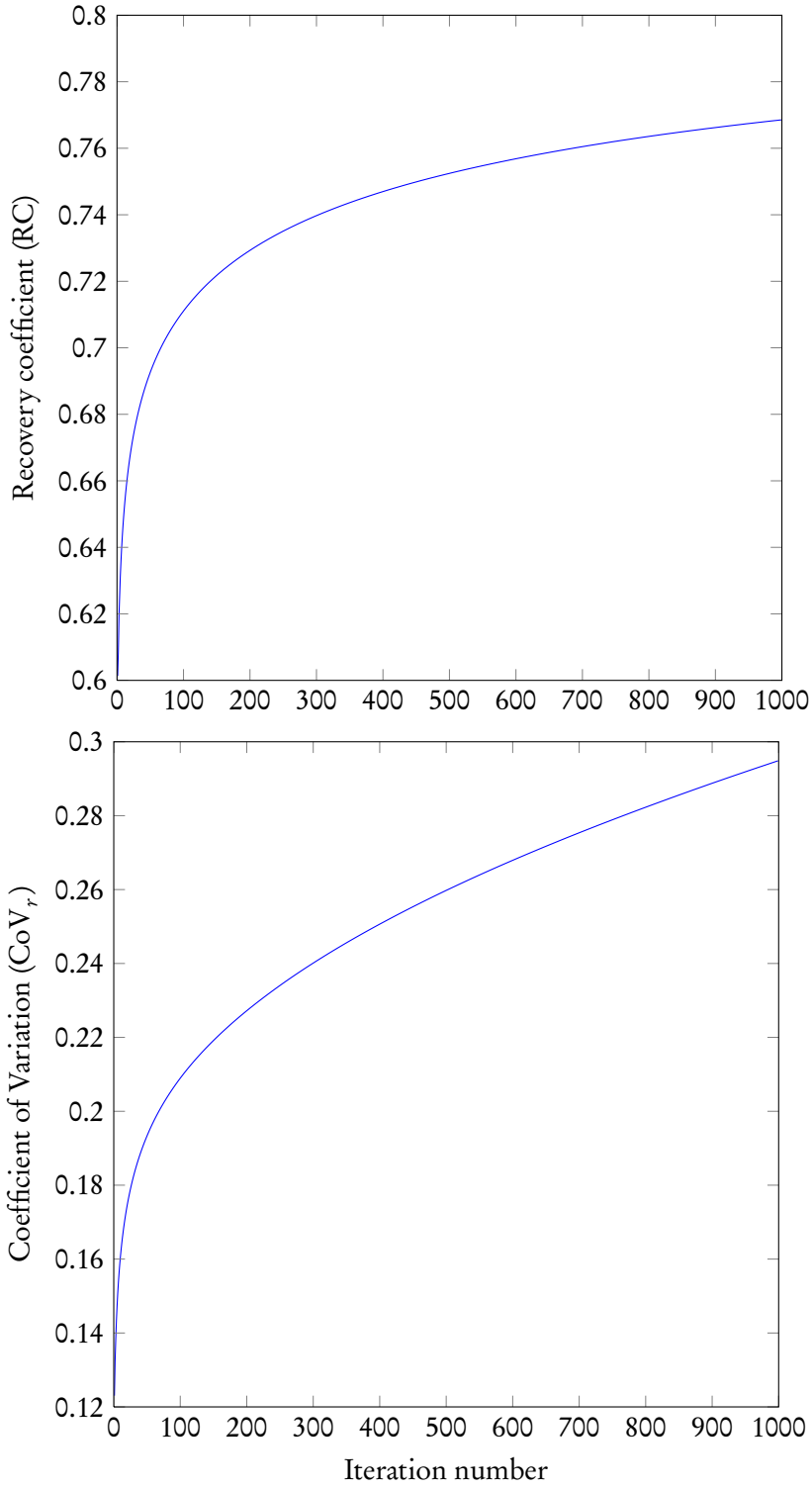


Figure A.4: VC deconvolution of the left frontal region in a single realisation of the AD phantom for iterations 1 to 1000. Graphs are of the mean recovery coefficient (RC) (top) and regional coefficient of variation (CoV$_r$) (bottom).

Appendix B

Publications arising from thesis work

B.1 The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease

Thomas, B. A., Erlandsson, K., Modat, M., Thurfjell, L., Vandenberghe, R., Ourselin, S., and 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.

B.2 Partial volume correction in SPECT reconstruction with OSEM

Erlandsson, K., Thomas, B., Dickson, J., and Hutton, B. F. (2011). Partial volume correction in SPECT reconstruction with OSEM. *Nucl Instrum Meth A*, 648, Supplement 1:S85-S88. NIMA 4th International Conference on Imaging techniques in Subatomic Physics, Astrophysics, Medicine, Biology and Industry.

B.3 Choice of alternative reference region and partial volume correction method improves agreement between amyloid PET tracers

Thomas, B. A., Erlandsson, K., Thurfjell, L., Vandenberghe, R., Ourselin, S., and Hutton, B. F. (2011) Choice of alternative reference region and partial volume correction method improves agreement between amyloid PET tracers. *Submitted*.

B.4 A novel iterative partial volume correction technique for emission tomography: optimising MR parcellations to improve brain PET quantification

Thomas, B. A., Erlandsson, K., Thurfjell, L., Ourselin, S., and Hutton, B. F. (2011) A novel iterative partial volume correction technique for emission tomography: optimising MR par-

cellations to improve brain PET quantification. *Accepted abstract, to appear*.

B.5 White matter variability for two amyloid brain PET tracers: implications for partial volume correction

Thomas, B. A., Erlandsson, K., Thurfjell, L., Vandenberghe, R., Ourselin, S., and Hutton, B. F. (2011) White matter variability for two amyloid brain PET tracers: implications for partial volume correction *Accepted abstract, to appear*.

B.6 Application of partial-volume correction in 123I-Ioflupane SPECT imaging

Thomas, B., Dickson, J., Erlandsson, K., and Ourselin, S., and Hutton, B. (2010) Application of partial-volume correction in 123I-Ioflupane SPECT imaging. *J Nucl Med Meeting Abstracts*, 51(2):1342 (abs).

B.7 The effects of iterative deconvolution partial-volume correction on beta-amyloid PET

Thomas, B., Thurfjell, L., Rinne, J.O., Ourselin, S. and Hutton, B. (2009) The effects of iterative deconvolution partial-volume correction on beta-amyloid PET *J Cereb Blood F Met*, 29:S329-S345 (abs).

B.8 The application of partial-volume correction to Alzheimer's disease discrimination

Thomas, B., Thurfjell, L., Rinne, J.O., Ourselin, S., Hutton B. (2009) The application of partial-volume correction to Alzheimer's disease discrimination. *Alzheimers Dement*, 5(4):P20-P21 (abs).

B.9 The use of priors for deconvolution-based partial volume correction

Thomas, B., Erlandsson, K., Thurfjell, L., Ourselin, S., Hutton B. (2008) The use of priors for deconvolution-based partial volume correction. *Nucl Med Commun*, 500-501 (abs).

Appendix C

Software developed during the thesis

C.1 INMRBV

The RBV correction technique has been implemented as a C++ command-line application. This was created to process the large clinical cohort (chapter 4) using the cluster computing resources at the Centre for Medical Image Computing (CMIC), UCL. The ITK (Kitware Inc., New York, NY, USA) [Yoo et al., 2002] is used extensively throughout the application.

`INMRBV` employs ITK file readers and writers, and can handle different types of input and output files. All images processed during this work were in either NIfTI file format or its compressed equivalent. The application has also undergone successful rudimentary testing with ANALYZE 7.5 files. At present, Digital Imaging and Communications in Medicine (DICOM) files are not supported in either the 2D or 3D format.

Two input files are required to use the application. The first must contain the PET image, with the second being a file containing the segmented region definitions. The regions must be enumerated as discrete integers from $1 \ldots N$, where N is the number of regions. The background region should be set to `0`. The parameter list is given below:

```
INMRBV <inputImageFile>  <maskFileName>  <outputImageFile>  <noRegions>
       <FWHM> [-q]
```

where `inputImageFile` is the PET image filename, `maskFileName` is the filename of the mask, `outputImageFile` is the name of the file to output (N.B. this file must have a suffix such as `.hdr`, `.nii` or `.nii.gz`), `noRegions` is the value N and `FWHM` is the FWHM of the PSF given in mm. The optional parameter `-q` tells `INMRBV` to operate in 'quite mode', suppressing any commmand-line output.

C.2 INMPVC

INMPVC is a command-line application that can perform either a MG, mMG, GTM or RBV correction. The application was also written in C++ and uses ITK. As with INMRBV, the first input file must be the PET image. The region masks are handled differently, in that a separate file is required for each masked region. This allows non-binary region definitions. Regions can be defined as probability masks, although it is important that each voxel must sum to the value of 1 over the set of masks. This functionality was included so that the results of probablistic segmentation algorithms could be readily used and also to allow for estimates of MRI PVEs to be taken into account. Probablistic segmentations produced by SPM were evaluated to test the functionality.

Each region mask filename is assumed to prefixed with cX where X is the region number between 1 and N. For example, the filename of the first region definition could be c1_myfile.nii.gz. When performing MG or mMG, c1 must be the GM, c2 must be the WM and c3 must be the CSF mask. The order is important as the MG methods only correct the GM voxels. If the MG correction is selected, an erosion operator is used to calculate the mean values. When performing mMG, a three-compartment GTM is evalutated using the region masks, and the mean values are then used to perform a MG correction.

Selecting the GTM option will result in the contents of the GTM matrix being displayed along with corrected (and uncorrected) mean values. No image will be produced. If RBV correction is selected, the creation of the 'synthetic PET' mask is performed differently to account for the possibility probablistic masks, although the execution of the correction is the same. The parameter list is given below:

```
INMPVC <inputImageFile>  <maskDir>  <maskFileName>  <outputImageFile>
       <noRegions> <FWHM> <corrType> [-q]
```

where inputImageFile is the PET image filename, maskDir is the path to the directory containing the mask files and maskFileName is portion of the mask filename after the region identifier; e.g. _myfile.nii.gz. outputImageFile is the name of the file to output, noRegions is the value N and FWHM is the FWHM of the PSF given in mm. corrType is the desired PVC technique to apply. For MG, the parameter is -MG, mMG is -MMG, GTM is called with -G, and RBV using -RBV. The optional parameter -q tells INMPVC to operate in 'quiet mode'.

C.3 SMPVC

The SMPVC technique has been implemented as a C++ command-line application. `SMPVC` employs ITK file readers and writers, and can handle different types of input and output files. All images processed during this work were in either NIfTI file format or its compressed equivalent. DICOM files are not supported in either the 2D or 3D format.

Three input files are required to use the application. The first must contain the PET image, with the second being a file containing the segmented region definitions. The regions must be enumerated as discrete integers from $1 \ldots N$, where N is the number of regions. The background region should be set to `0`. The third is a binary file of unsigned integers that represent an $N \times N$ matrix. The matrix describes which regions are allowed to be modified, and also which regions can move into each other. This is used to restrict modifications to GM regions only and to prevent inter-hemispheric changes. A value of `0` at location (i, j) in the matrix will prevent region i from modifying region j. Whereas, a value of 1 will permit modification. The parameter list for `SMPVC` is given below:

```
SMPVC <inputImageFile>  <maskFileName>  <outputImageFile>
       <FWHM>  <noRegions>  <noIterations> <matrixFileName>
```

where `inputImageFile` is the PET image filename, `maskFileName` is the filename of the mask, `outputImageFile` is the name of the file to output (N.B. this file must have a suffix such as `.hdr`, `.nii` or `.nii.gz`), `FWHM` is the FWHM of the PSF given in mm, `noRegions` is the number of regions (N), `noIterations` is the number of times to iterate SMPVC (referred to as $kMax$ in section 5.2.2.4, page 133) and `matrixFileName` is the file containing the matrix of allowed region modifications.

C.4 Phantom image generator

The phantom image generator described in section 3.2.3 was created in MATLAB. All convolutions were performed using the MATLAB fast fourier transform (FFT) routines. The forward- and back-projections use the Radon and inverse Radon transforms that are part of the Image Processing toolbox. The function `MakeImage` generates the phantom images. `MakeImage` requires a minimum of four parameters, which are the input files and the desired FWHM used to convolve the image. The following shows the parameter list for `MakeImage`:

```
[Sino]=MakeImages(inputFile,distFile,dFWHM,ctFile,Projections,dCountLevel);
```

`inputFile` is a string holding the filename of the brain mask image.`distFile` is the filename of the CSV file containing the regional values associated with the brain regions. `dFWHM` is the

FWHM of the PSF to be simulated, given in mm. `ctFile` is a string containing the filename of the associated CT image. `Projections` is a vector containing the angles to be used by the projector. The default for `Projections` is `0:0.8:179` (or from 0° to 179° in 0.8° increments). The parameter `dCountLevel` scales the image to the desired count level when adding Poisson noise. The default value for `dCountLevel` is $1e^8$.

`MakeImage` produces the sinogram data. This can then be reconstructed with a desired reconstruction algorithm. In this work, all phantom data was reconstructed using FBP, which in MATLAB is provided by the `iradon` function.

C.5 RegionViewer

`RegionViewer` is an Adobe Integrated Runtime (AIR) application written for the purposes of data visualisation. The application reads image statistics that have been summarised as XML documents. A sample from top of one of the documents can be seen below:

```
<?xml version="1.0" encoding="utf-8"?>
  <subject>
      <region>
          <regionID>1</regionID>
          <mean>2.227692</mean>
          <std>0.213757</std>
      </region>
      <region>
          <regionID>2</regionID>
          <mean>1.981361</mean>
          <std>0.179788</std>
      </region>
```

Each region node contains the mean and SD of regional SUVR values. In this case all images were normalised to cerebellar GM. `RegionViewer` does not access the actual imaging data and therefore normalisation is determined by the creator of the XML documents. The data can also be collated into subject groups in order to perform group analysis. `RegionViewer` can perform the following analysis:

- Region-by-region SUVR analysis on a per subject basis.
- Group-wise plots of a single region SUVR for one or more subject groups.

- Region-against-region SUVR scatter plots for one or more subject groups.
- MRI-volume (mm^3) against SUVR scatter plots for one or more subject groups.

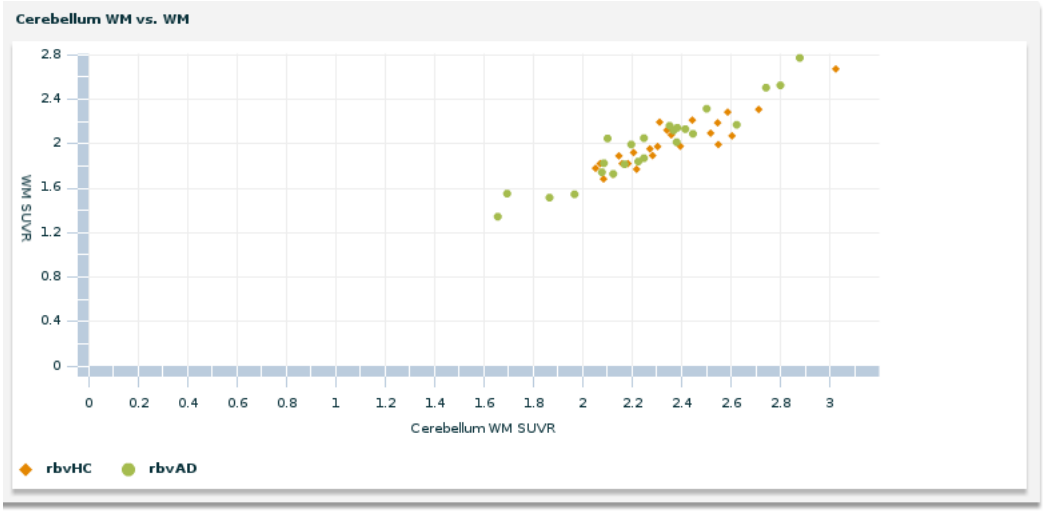


Figure C.1: Example of a `RegionViewer` scatter plot between cerebellar WM and sub-cortical WM.

The application was primarily used to investigate possible relations between regions, but could be further developed for multi-modal parameter evaluation.

Appendix D

Region definitions

FreeSurfer produces a file called `aparc+aseg.mgz` for each subject which contains the results of the parcellation. The labels in this file were used to create larger region masks according to the tables below:

D.1 Region definitions for clinical [^{18}F]flutemetamol study

Region	FreeSurfer label
Frontal	1028 2028 1027 2027 1003 2003 1018 2018 1019 2019 1020 2020 1012 2012 1014 2014 1032 2032
Central	1024 2024 1017 2017 1022 2022
Parietal	1029 2029 1008 2008 1031 2031 1025 2025
Temporal	1030 2030 1015 2015 1009 2009 1007 2007 1034 2034 1006 2006 1033 2033 1016 2016 1001 2001
Occipital	1011 2011 1013 2013 1005 2005 1021 2021
Cingulate	1026 2026 1002 2002 1023 2023 1010 2010
Insula	1035 2035
Deep GM	10 49 11 50 12 51 13 52 18 54 26 58 27 59
Hippocampus	17 53
Sub-cortical WM	1004 2004 2 41 77 78 79 250 251 252 253 254 255
Cerebellar WM	7 46
Cerebellar GM	8 47
Ventricles and CSF	4 5 14 72 75 76 43 44 24 31 63
Brain stem	16

Table D.1: The region definitions used during the analysis of clinical [^{18}F]flutemetamol data in section 4.1.

D.2 Region definitions for the left \ right brain parcellation

Region	FreeSurfer label
Left frontal	1028 1027 1003 1018 1019 1020 1012 1014 1032
Right frontal	2028 2027 2003 2018 2019 2020 2012 2014 2032
Left central	1024 1017 1022
Right central	2024 2017 2022
Left parietal	1029 1008 1031 1025
Right parietal	2029 2008 2031 2025
Left temporal	1030 1015 1009 1007 1034 1006 1033 1016 1001
Right temporal	2030 2015 2009 2007 2034 2006 2033 2016 2001
Left occipital	1011 1013 1005 1021
Right occipital	2011 2013 2005 2021
Left ant. cingulate	1026 1002
Right ant. cingulate	1023 1010
Left post. cingulate	2026 2002
Right post. cingulate	2023 2010
Left insula	1035
Right insula	2035
Deep GM	10 49 11 50 12 51 13 52 18 54 26 58 27 59
Left hippocampus	17
Right hippocampus	53
Sub-cortical WM	1004 2004 2 41 77 78 79 250 251 252 253 254 255
Cerebellar WM	7 46
Cerebellar GM	8 47
CSF	4 5 14 72 75 76 43 44 24 31 63
Brain stem	16

Table D.2: The lateralised FreeSurfer region definitions used during the amyloid normalisation section 4.3.

D.3 Region definitions for the extended brain parcellation

Region	FreeSurfer label
Left superior frontal	1028 1027 1003
Left inferior frontal	1018 1019 1020 1012 1014 1032
Right superior frontal	2028 2027 2003
Right inferior frontal	2018 2019 2020 2012 2014 2032
Left central	1024 1017 1022
Right central	2024 2017 2022
Left precuneus	1025
Right precuneus	2025
Left inferior parietal	1008
Right inferior parietal	2008
Left superior parietal	1029 1031
Right superior parietal	2029 2031
Left superior temporal	1030 1001 1034
Right superior temporal	2030 2001 2034
Left mid temporal	1015
Right mid temporal	2015
Left inferior temporal	1016 1033 1006 1007 1009
Right inferior temporal	2016 2033 2006 2007 2009
Left occipital	1011 1013 1005 1021
Right occipital	2011 2013 2005 2021
Left ant. cingulate	1026 1002
Right ant. cingulate	1023 1010
Left post. cingulate	2026 2002
Right post. cingulate	2023 2010
Left insula	1035
Right insula	2035
Deep GM	10 49 11 50 12 51 13 52 18 54 26 58 27 59
Left hippocampus	17
Right hippocampus	53
Sub-cortical WM	1004 2004 2 41 77 78 79 250 251 252 253 254 255
Cerebellar WM	7 46
Cerebellar GM	8 47
CSF	4 5 14 72 75 76 43 44 24 31 63
Brain stem	16

Table D.3: The extended FreeSurfer region definitions used for SMPVC (chapter 5).

Bibliography

Aisen, P. S. (2002). Development of antiinflammatory therapy for Alzheimer's disease. *Drug Develop Res*, 56(3):421–427.

Aizenstein, H. J., Nebes, R. D., Saxton, J. A., Price, J. C., Mathis, C. A., Tsopelas, N. D., Ziolko, S. K., James, J. A., Snitz, B. E., Houck, P. R., Bi, W., Cohen, A. D., Lopresti, B. J., DeKosky, S. T., Halligan, E. M., and Klunk, W. E. (2008). Frequent amyloid deposition without significant cognitive impairment among the elderly. *Arch Neurol*, 65(11):1509–1517.

Aljabar, P., Heckemann, R., Hammers, A., Hajnal, J., and Rueckert, D. (2009). Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. *NeuroImage*, 46(3):726 – 738.

Andersson, J. L. R. and Thurfjell, L. (1997). Implementation and validation of a fully automatic system for intra- and interindividual registration of PET brain scans. *J Comp Assist Tomo*, 21:136–144.

Anger, H. O. (1958). Scintillation camera. *Rev Sci Instrum*, 29(1):27–33.

Archer, H. A., Edison, P., Brooks, D. J., Barnes, J., Frost, C., Yeatman, T., Fox, N. C., and Rossor, M. N. (2006). Amyloid load and cerebral atrophy in Alzheimer's disease: An 11-C-PIB positron emission tomography study. *Ann Neurol*, 60(1):145–147.

Ashburner, J. and Friston, K. J. (1999). Nonlinear spatial normalization using basis functions. *Hum Brain Mapp*, 7:254–266.

Aston, J. A., Cunningham, V. J., Asselin, M. C., Hammers, A., Evans, A. C., and Gunn, R. N. (2002). Positron emission tomography partial volume correction: Estimation and algorithms. *J Cereb Blood F Met*, 22(8):1019–1034.

Barnes, J., Scahill, R. I., Boyes, R. G., Frost, C., Lewis, E. B., Rossor, C. L., Rossor, M. N.,

and Fox, N. C. (2004). Differentiating AD from aging using semiautomated measurement of hippocampal atrophy rates. *NeuroImage*, 23(2):574–581.

Barthel, H., Gertz, H.-J., Dresel, S., Peters, O., Bartenstein, P., Buerger, K., Hiemeyer, F., Wittemer-Rump, S. M., Seibyl, J., Reininger, C., and et al. (2011). Cerebral amyloid-β PET with florbetaben ((18)F) in patients with Alzheimer's disease and healthy controls: a multicentre phase 2 diagnostic study. *Lancet*, 10(5):424–435.

Bataille, F., Comtat, C., Jan, S., Sureau, F. C., and Trebossen, R. (2007). Brain PET Partial-Volume Compensation Using Blurred Anatomical Labels. *IEEE T Nucl Sci*, 54(5):1606–1615.

Bergström, M., Eriksson, L., Bohm, C., Blomqvist, G., and Litton, J. (1983). Correction for scattered radiation in a ring detector positron camera by integral transformation of the projections. *J Comp Assist Tomo*, 7(1):42–50.

Bourgeat, P., Chételat, G., Villemagne, V. L., Fripp, J., Raniga, P., Pike, K., Acosta, O., Szoeke, C., Ourselin, S., Ames, D., and et al. (2010). Beta-amyloid burden in the temporal neocortex is related to hippocampal atrophy in elderly subjects without dementia. *Neurology*, 74(2):121–127.

Boussion, N., Cheze Le Rest, C., Hatt, M., and Visvikis, D. (2009). Incorporation of wavelet-based denoising in iterative deconvolution for partial volume correction in whole-body PET imaging. *Eur J Nucl Med Mol Imaging*, 36:1064–1075.

Boussion, N., Hatt, M., Lamare, F., Bizais, Y., Turzo, A., Rest, C., and Visvikis, D. (2006). A multiresolution image based approach for correction of partial volume effects in emission tomography. *Phys Med Biol*, 51(7):1857–1876.

Bowsher, J., Yuan, H., Hedlund, L., Turkington, T., Akabani, G., Badea, A., Kurylo, W., Wheeler, C., Cofer, G., Dewhirst, M., and Johnson, G. (2004). Utilizing MRI information to estimate F18-FDG distributions in rat flank tumors. In *IEEE Nucl Sci Conf R*, volume 4, pages 2488 – 2492.

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

Boxer, A., Rabinovici, G., Kepe, V., Goldman, J., Furst, A., Huang, S.-C., Baker, S., O'Neil, J., Chui, H., Geschwind, M., Small, G., Barrio, J., Jagust, W., and Miller, B. (2007). Amyloid imaging in distinguishing atypical prion disease from Alzheimer disease. *Neurology*, 69(3):283–290.

Braak, H. and Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica*, 82(4):239–259.

Braak, H. and Braak, E. (1995). Staging of Alzheimer's Disease-related neurofibrillary changes. *Neurobiol Aging*, 16(3):271–278.

Caramanos, Z., Fonov, V. S., Francis, S. J., Narayanan, S., Pike, G. B., Collins, D. L., and Arnold, D. L. (2010). Gradient distortions in MRI: Characterizing and correcting for their effects on SIENA-generated measures of brain volume change. *NeuroImage*, 49(2):1601 – 1611.

Carson, R. E., Barker, W. C., Liow, J.-S., and Johnson, C. A. (2003). Design of a Motion-Compensation OSEM List-mode Algorithm for Resolution-Recovery Reconstruction for the HRRT. In *IEEE Nucl Sci Conf R*, volume 5, pages 3281–3285.

Casey, M. E. and Nutt, R. (1986). A multicrystal two-dimensional BGO detector system for positron emission tomography. *IEEE T Nucl Sci*, 33:460–463.

Cherry, S., Sorenson, J., and Phelps, M. (2003). *Physics in nuclear medicine*. Saunders, 3rd edition.

Chow, V. W., Savonenko, A. V., Melnikova, T., Kim, H., Price, D. L., Li, T., and Wong, P. C. (2010). Modeling an Anti-Amyloid Combination Therapy for Alzheimer's Disease. *Sci Tranl Med*, 2(13):13ra1.

Clark, C. M., Schneider, J. A., Bedell, B. J., Beach, T. G., Bilker, W. B., Mintun, M. A., Pontecorvo, M. J., Hefti, F., Carpenter, A. P., Flitter, M. L., Krautkramer, M. J., Kung, H. F., Coleman, R. E., Doraiswamy, P. M., Fleisher, A. S., Sabbagh, M. N., Sadowsky, C. H., Reiman, E. M., Zehntner, S. P., and Skovronsky, D. M. (2011). Use of Florbetapir-PET for Imaging β-Amyloid Pathology. *JAMA-J Am Med Assoc*, 305(3):275–283.

Comtat, C., Kinahan, P. E., Fessler, J. A., Beyer, T., Townsend, D. W., Defrise, M., and Michel, C. (2002). Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels. *Phys Med and Biol*, 47(1):1–20.

Comtat, C., Sureau, F. C., Sibomana, M., Hong, I. K., Sjoholm, N., and Trebossen, R. (2008). Image based resolution modeling for the HRRT OSEM reconstructions software. In *IEEE Nucl Sci Conf R*, pages 4120–4123.

Cootes, T. F., Twining, C. J., and Taylor, C. J. (2004). Diffeomorphic statistical shape models. In *Proceedings of BVMC*, pages 447–456.

Dantzig, G. (1998). *Linear programming and extensions*. Landmarks in Physics and Mathematics. Princeton University Press.

De Bernardi, E., Mazzoli, M., Zito, F., and Baselli, G. (2007). Resolution Recovery in PET During AWOSEM Reconstruction: A Performance Evaluation Study. *IEEE T Nucl Sci*, 54(5):1626–1638.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., and et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31(3):968–980.

Dey, N., Blanc-Feraud, L., Zimmer, C., Roux, P., Kam, Z., Olivo-Marin, J.-C., and Zerubia, J. (2006). Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Micros Res and Techniq*, 69(4):260–266.

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302.

Drzezga, A., Lautenschlager, N., Siebner, H., Riemenschneider, M., Willoch, F., Minoshima, S., Schwaiger, M., and Kurz, A. (2003). Cerebral metabolic changes accompanying conversion of mild cognitive impairment into Alzheimer's disease: a PET follow-up study. *Eur J Nucl Med Mol Imaging*, 30:1104–1113.

Du, Y., Du, Y., Tsui, B., and Frey, E. (2005). Partial Volume Effect Compensation for Quantitative Brain SPECT Imaging Partial Volume Effect Compensation for Quantitative Brain SPECT Imaging. *IEEE T Med Imaging*, 24(8):969–976.

Edison, P., Archer, H. A., Hinz, R., Hammers, A., Pavese, N., Tai, Y. F., Hotton, G., Cutler, D., Fox, N., Kennedy, A., Rossor, M., and Brooks, D. J. (2007). Amyloid, hypometabolism, and cognition in Alzheimer disease. *Neurology*, 68(7):501–508.

Ellis, K. A., Bush, A. I., Darby, D., De Fazio, D., Foster, J., Hudson, P., Lautenschlager, N. T., Lenzo, N., Martins, R. N., Maruff, P., and et al. (2009). The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *Int Psychogeriatr*, 21(4):672–687.

Erlandsson, K. and Hutton, B. F. (2010). Partial Volume Correction in SPECT Using Anatomical Information and Iterative FBP. *Tsinghua Sci Technol*, 15(1):50–55.

Erlandsson, K. and Hutton, B. F. (2011). Hyper-plane based partial volume correction for PET or SPECT. *Nucl Med Comm*, 32(5):427.

Erlandsson, K., Thomas, B., Dickson, J., and Hutton, B. F. (2011). Partial volume correction in SPECT reconstruction with OSEM. *Nucl Instrum Meth A*, 648(Supplement 1):S85 – S88. NIMA 4th International Conference on Imaging techniques in Subatomic Physics, Astrophysics, Medicine, Biology and Industry.

Erlandsson, K., Wong, A. T., van Heertum, R., Mann, J. J., and Parsey, R. V. (2006). An improved method for voxel-based partial volume correction in PET and SPECT. *NeuroImage*, 31:T84.

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., and Peters, T. M. (1993). 3D statistical neuroanatomical models from 305 MRI volumes. *IEEE Nucl Sci Conf R*, 3:1813–1817.

Ferrant, M., Nabavi, A., Macq, B., Black, P. M., Jolesz, F. A., Kikinis, R., and Warfield, S. K. (2002). Serial registration of intraoperative MR images of the brain. *Medical Image Anal*, 6(4):337–359.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., and Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–55.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., and Dale, A. M. (2004). Automatically parcellating the human cerebral cortex. *Cereb Cortex*, 14(1):11–22.

Foley, J., Van Dam, A., Feiner, S., and Hughes, J. (1996). *Computer graphics: principles and practice*. The Systems Programming Series. Addison-Wesley.

Foster, N. L., Heidebrink, J. L., Clark, C. M., Jagust, W. J., Arnold, S. E., Barbas, N. R., DeCarli, C. S., Scott Turner, R., Koeppe, R. A., Higdon, R., and Minoshima, S. (2007). FDG-PET improves accuracy in distinguishing frontotemporal dementia and Alzheimer's disease. *Brain*, 130(10):2616–2635.

Foster, N. L., Wang, A. Y., Tasdizen, T., Fletcher, P., Hoffman, J. M., and Koeppe, R. A. (2008). Realizing the potential of positron emission tomography with 18F-fluorodeoxyglucose to improve the treatment of Alzheimer's disease. *Alzheimers Dement*, 4(1, Supplement 1):S29–S36.

Fox, N. C. and Schott, J. M. (2004). Imaging cerebral atrophy: normal ageing to Alzheimer's disease. *Lancet*, 363(9406):392–394.

Fripp, J., Bourgeat, P., Acosta, O., Raniga, P., Modat, M., Pike, K. E., Jones, G., O'Keefe, G., Masters, C. L., Ames, D., Ellis, K. A., Maruff, P., Currie, J., Villemagne, V. L., Rowe, C. C., Salvado, O., and Ourselin, S. (2008). Appearance modeling of 11C PiB PET images: Characterizing amyloid deposition in Alzheimer's disease, mild cognitive impairment and healthy aging. *NeuroImage*, 43(3):430–439.

Frisoni, G., Laakso, M., Beltramello, A., Geroldi, C., Bianchetti, A., Soininen, H., and Trabucchi, M. (1999). Hippocampal and entorhinal cortex atrophy in frontotemporal dementia and Alzheimer's disease. *Neurology*, 52(1):91–100.

Frisoni, G. B., Pievani, M., Testa, C., Sabattoli, F., Bresciani, L., Bonetti, M., Beltramello, A., Hayashi, K. M., Toga, A. W., and Thompson, P. M. (2007). The topography of grey matter involvement in early and late onset Alzheimer's disease. *Brain*, 130(3):720–730.

Hardy, J. A. and Higgins, G. A. (1992). Alzheimer's disease: the amyloid cascade hypothesis. *Science*, 256(5054):184–185.

Heckemann, R. A., Keihaninejad, S., Aljabar, P., Rueckert, D., Hajnal, J. V., and Hammers, A. (2010). Improving intersubject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation. *NeuroImage*, 51(1):221 – 227.

Huber, P. (1981). *Robust statistics*. Wiley.

Hudson, H. and Larkin, R. (1994). Accelerated image reconstruction using ordered subsets of projection data. *IEEE T Med Imaging*, 13(4):601–609.

Imran, M. B., Kawashima, R., Awata, S., Sato, K., Kinomura, S., Ono, S., Sato, M., and Fukuda, H. (1999). Tc-99m HMPAO SPECT in the evaluation of Alzheimer's disease: correlation between neuropsychiatric evaluation and CBF images. *J Neurol Neurosur Ps*, 66(2):228–232.

Ishii, K., Kono, A., Sasaki, H., Miyamoto, N., Fukuda, T., Sakamoto, S., and Mori, E. (2006). Fully automatic diagnostic system for early- and late-onset mild Alzheimer's disease using FDG PET and 3D-SSP. *Eur J Nucl Med Mol Imaging*, 33(5):575–583.

Ishii, K., Willoch, F., Minoshima, S., Drzezga, A., Ficaro, E. P., Cross, D. J., Kuhl, D. E., and Schwaiger, M. (2001). Statistical Brain Mapping of 18F-FDG PET in Alzheimer's Disease: Validation of Anatomic Standardization for Atrophied Brains. *J Nucl Med*, 42(4):548–557.

Jack, C. R. (1997). Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. *Neurology*, 49(3):786–794.

Jack, C. R. J., Shiung, M. M., Weigand, S. D., O'Brien, P. C., Gunter, J. L., Boeve, B. F., Knopman, D. S., Smith, G. E., Ivnik, R. J., Tangalos, E. G., and Petersen, R. C. (2005). Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. *Neurology*, 65(8):1227–1231.

Jack, Clifford R., J., Lowe, V. J., Weigand, S. D., Wiste, H. J., Senjem, M. L., Knopman, D. S., Shiung, M. M., Gunter, J. L., Boeve, B. F., Kemp, B. J., Weiner, M., Petersen, R. C., and the Alzheimer's Disease Neuroimaging Initiative (2009). Serial PIB and MRI in normal, mild cognitive impairment and Alzheimer's disease: implications for sequence of pathological events in Alzheimer's disease. *Brain*, 132(5):1355–1365.

Jagust, W. (2008). Is amnestic mild cognitive impairment always AD? *Neurology*, 70(7):502–503.

Jagust, W., Thisted, R., Devous, M.D., S., Van Heertum, R., Mayberg, H., Jobst, K., Smith, A., and Borys, N. (2001). SPECT perfusion imaging in the diagnosis of Alzheimer's disease: A clinical-pathologic study. *Neurology*, 56(7):950–956.

Joachim, C. L., Morris, J. H., and Selkoe, D. J. (1989). Diffuse senile plaques occur commonly in the cerebellum in Alzheimer's disease. *Am J Pathol*, 135(2):309–319.

Juottonen, K., Laakso, M., Insausti, R., Lehtovirta, M., Pitkanen, A., Partanen, K., and Soininen, H. (1998). Volumes of the Entorhinal and Perirhinal Cortices in Alzheimer's Disease. *Neurobiol Aging*, 19(1):15–22.

Kemppainen, N., Aalto, S., Wilson, I., Nagren, K., Helin, S., Bruck, A., Oikonen, V., Kailajarvi, M., Scheinin, M., Viitanen, M., Parkkola, R., and Rinne, J. (2006). Voxel-based analysis of PET amyloid ligand [11C]PIB uptake in Alzheimer Disease. *Neurology*, 67(9):1575–1580.

Kinahan, P. E., Townsend, D. W., Beyer, T., and Sashin, D. (1998). Attenuation correction for a combined 3D PET / CT scanner. *Med Phys*, 25:2046–2053.

Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M.-C., Christensen, G. E., Collins, D. L., Gee, J., Hellier, P., Song, J. H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R. P., Mann, J. J., and Parsey, R. V. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, 46(3):786 – 802.

Klein, A., Mensh, B., Ghosh, S., Tourville, J., and Hirsch, J. (2005). Mindboggle: Automated brain labeling with multiple atlases. *BMC Med Imaging*, 5(1):7.

Klunk, W. E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D. P., Bergström, M., Savitcheva, I., Huang, G.-F., Estrada, S., Ausen, B., Debnath, M. L., Barletta, J., Price, J. C., Sandell, J., Lopresti, B. J., Wall, A., Koivisto, P., Antoni, G., Mathis, C. A., and Längström, B. (2004). Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Ann Neurol*, 55(3):306–319.

Knuth, D. (1998). *The Art of Computer Programming: Seminumerical algorithms*. The Art of Computer Programming. Addison-Wesley Pub. Co.

Koivunen, J., Verkkoniemi, A., Aalto, S., Paetau, A., Ahonen, J.-P., Viitanen, M., Någren, K., Rokka, J., Haaparanta, M., Kalimo, H., and Rinne, J. O. (2008). PET amyloid ligand [11C]PIB uptake shows predominantly striatal increase in variant Alzheimer's disease. *Brain*, 131(7):1845–1853.

Kono, A., Ishii, K., Sofue, K., Miyamoto, N., Sakamoto, S., and Mori, E. (2007). Fully automatic differential diagnosis system for dementia with Lewy bodies and Alzheimer's disease using FDG-PET and 3D-SSP. *Eur J Nucl Med Mol Imaging*, 34(9):1490–1497.

Långström, B. (2007). In vitro imaging techniques in Neurodegenerative diseases. *Mol Imaging Biol*, 9(4):161–175.

Larner, A. J. (1997). The cerebellum in Alzheimer's disease. *Dement Geriatr Cogn Disord*, 8(4):203–209.

Le Pogam, A., Boussion, N., Hatt, M., Turkheimer, F., Prunier-Aesch, C., Guilloteau, D., Baulieu, J., and Visvikis, D. (2008). A 3D multi resolution local analysis approach for correction of partial volume effects in emission tomography. In *IEEE Nucl Sci Conf R*, pages 5300–5303.

Li, Y., Rinne, J. O., Mosconi, L., Pirraglia, E., Rusinek, H., DeSanti, S., Kemppainen, N., Någren, K., Kim, B.-C., Tsui, W., and de Leon, M. J. (2008). Regional analysis of FDG and PIB-PET images in normal aging, mild cognitive impairment, and Alzheimer's disease. *Eur J Nucl Med Mol Imaging*, 35(12):2169–81.

Logan, J. (2000). Graphical analysis of pet data applied to reversible and irreversible tracers. *Nuclear Medicine and Biology*, 27(7):661–70.

Logan, J., Fowler, J. S., Volkow, N. D., Wang, G. J., Ding, Y. S., and Alexoff, D. L. (1996). Distribution volume ratios without blood sampling from graphical analysis of PET data. *J Cerebr Blood F Met*, 16(5):834–840.

Lopresti, B. J., Klunk, W. E., Mathis, C. A., Hoge, J. A., Ziolko, S. K., Lu, X., Meltzer, C. C., Schimmel, K., Tsopelas, N. D., DeKosky, S. T., and Price, J. C. (2005). Simplified Quantification of Pittsburgh Compound B Amyloid Imaging PET Studies: A Comparative Analysis. *J Nucl Med*, 46(12):1959–1972.

Lucy, L. (1974). An iterative technique for the rectification of observed distributions. *Astron J*, 79:745–754.

Maintz, J. B. and Viergever, M. A. (1998). A survey of medical image registration. *Med Image Anal*, 2(1):1–36.

Mawlawi, O., Podoloff, D. A., Kohlmyer, S., Williams, J. J., Stearns, C. W., Culp, R. F., and Macapinlac, H. (2004). Performance Characteristics of a Newly Developed PET/CT Scanner Using NEMA Standards in 2D and 3D Modes. *J Nucl Med*, 45(10):1734–1742.

Mazziotta, J. C., Toga, A. W., Evans, A., Fox, P., and Lancaster, J. (1995). A Probabilistic

Atlas of the Human Brain: Theory and Rationale for Its Development : The International Consortium for Brain Mapping (ICBM). *NeuroImage*, 2(2, Part 1):89 – 101.

McEvoy, L. K., Fennema-Notestine, C., Roddey, J. C., Hagler, Donald J., J., Holland, D., Karow, D. S., Pung, C. J., Brewer, J. B., and Dale, A. M. (2009). Alzheimer Disease: Quantitative Structural Neuroimaging for Detection and Prediction of Clinical and Structural Changes in Mild Cognitive Impairment. *Radiology*, 251(1):195–205.

Meltzer, C. C., Kinahan, P. E., Greer, P. J., Nichols, T. E., Comtat, C., Cantwell, M. N., Lin, M. P., and Price, J. C. (1999). Comparative Evaluation of MR-Based Partial-Volume Correction Schemes for PET. *J Nucl Med*, 40(12):2053–2065.

Meltzer, C. C., Leal, J. P., Mayberg, H. S., Wagner, H. N. J., and Frost, J. J. (1990). Correction of PET Data for Partial Volume Effects in Human Cerebral Cortex by MR Imaging. *J Comput Assist Tomo*, 14(4):561–570.

Messa, C., Perani, D., Lucignani, G., Zenorini, A., Zito, F., Rizzo, G., Grassi, F., Del Sole, A., Franceschi, M., Gilardi, M. C., and Fazio, F. (1994). High-Resolution Technetium-99m-HMPAO SPECT in Patients with Probable Alzheimer's Disease: Comparison with Fluorine-18-FDG PET. *J Nucl Med*, 35(2):210–216.

Mintun, M. A., Larossa, G. N., Sheline, Y. I., Dence, C. S., Lee, S. Y., Mach, R. H., Klunk, W. E., Mathis, C. A., DeKosky, S. T., and Morris, J. C. (2006). [11C]PIB in a nondemented population: potential antecedent marker of Alzheimer disease. *Neurology*, 67(3):446–452.

Modat, M., Ridgway, G. R., Taylor, Z. A., Lehmann, M., Barnes, J., Hawkes, D. J., Fox, N. C., and Ourselin, S. (2010). Fast free-form deformation using graphics processing units. *Comput Meth Prog Bio*, 98(3):278–284.

Mormino, E. C., Kluth, J. T., Madison, C. M., Rabinovici, G. D., Baker, S. L., Miller, B. L., Koeppe, R. A., Mathis, C. A., Weiner, M. W., and Jagust, W. J. (2008). Episodic memory loss is related to hippocampal-mediated beta-amyloid deposition in elderly subjects. *Brain*, 132:1310–1323.

Mosconi, L., Tsui, W., Rusinek, H., De Santi, S., Li, Y., Wang, G. J., Pupi, A., Fowler, J., and de Leon, M. (2007). Quantitation, regional vulnerability, and kinetic modeling of brain glucose metabolism in mild Alzheimer's disease. *Eur J Nucl Med Mol Imaging*, 34(9):1467–1479.

Müller-Gärtner, H. W., Links, J. M., Prince, J. L., Bryan, R. N., McVeigh, E., Leal, J. P., Davatzikos, C., and 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 F Met*, 12(4):571–583.

Natterer, F. (2001). *The mathematics of computerized tomography*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

Nunez, J. and Llacer, J. (1990). A fast Bayesian reconstruction algorithm for emission tomography with entropy prior converging to feasible images. *IEEE T Med Imaging*, 9(2):159–171.

O'Keefe, G. (2011). personal communication.

Ourselin, S., Roche, A., Subsol, G., Pennec, X., and Ayache, N. (2001). Reconstructing a 3D structure from serial histological sections. *Image Vision Comput*, 19(1-2):25–31.

Ourselin, S., Stefanescu, R., and Pennec, X. (2002). Robust Registration of Multi-modal Images: Towards Real-Time Clinical Applications. In *MICCAI'02*, pages 140–147. Springer.

Pike, K. E., Savage, G., Villemagne, V. L., Ng, S., Moss, S. A., Maruff, P., Mathis, C. A., Klunk, W. E., Masters, C. L., and Rowe, C. C. (2007). β-amyloid imaging and memory in non-demented individuals: evidence for preclinical Alzheimer's disease. *Brain*, 130(11):2837–2844.

Powell, M. J. D. (1964). An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Comput J*, 7(2):155–162.

Price, J. C., Klunk, W. E., Lopresti, B. J., Lu, X., Hoge, J. A., Ziolko, S. K., Holt, D. P., Meltzer, C. C., DeKosky, S. T., and Mathis, C. A. (2005). Kinetic modeling of amyloid binding in humans using PET imaging and Pittsburgh Compound-B. *J Cerebr Blood F Met*, 25(11):1528–47.

Quarantelli, M., Berkouk, K., Prinster, A., Landeau, B., Svarer, C., Balkay, L., Alfano, B., Brunetti, A., Baron, J.-C., and Salvatore, M. (2004). Integrated Software for the Analysis of Brain PET/SPECT Studies with Partial-Volume-Effect Correction. *J Nucl Med*, 45(2):192–201.

Reader, A. J., Erlandsson, K., Flower, M. A., and Ott, R. J. (1998). Fast accurate iterative reconstruction for low-statistics positron volume imaging. *Phys Med Biol*, 43(4):835–846.

Reader, A. J., Julyan, P. J., Williams, H., Hastings, D. L., and Zweit, J. (2003). Em algorithm system modeling by image-space techniques for pet reconstruction. *IEEE Trans Nucl Sci*, 50(5):1392–1397.

Reilhac, A., Lartizien, C., Costes, N., Sans, S., Comtat, C., Gunn, R. N., and Evans, A. C. (2004). PET-SORTEO: A Monte Carlo-Based Simulator With High Count Rate Capabilities. *IEEE T Nucl Sci*, 51(1):46–52.

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., and Davatzikos, C. (2003). Longitudinal Magnetic Resonance Imaging Studies of Older Adults: A Shrinking Brain. *J Neurosci*, 23(8):3295–3301.

Resnick, S. M., Sojkova, J., Zhou, Y., An, Y., Ye, W., Holt, D. P., Dannals, R. F., Mathis, C. A., Klunk, W. E., Ferrucci, L., Kraut, M. A., and Wong, D. F. (2010). Longitudinal cognitive decline is associated with fibrillar amyloid-beta measured by [11C]PiB. *Neurology*, 74(10):807–815.

Richardson, W. H. (1972). Bayesian-based iterative method of image restoration. *J Opt Soc Am*, 62(1):55.

Rousset, O., Rahmim, A., Alavi, A., and Zaidi, H. (2007). Partial Volume Correction Strategies in PET. *PET Clinics*, 2(2):235–249.

Rousset, O. G., Ma, Y., and Evans, A. C. (1998a). Correction for partial volume effects in PET: principle and validation. *J Nucl Med*, 39(5):904–11.

Rousset, O. G., Ma, Y., Wong, D. F., and Evans, A. C. (1998b). Pixel- versus region-based partial volume correction in PET. In Carson, R. E., Daube-Witherspoon, M. E., and Herscovitch, P., editors, *Quantitative Functional Brain Imaging with Positron Emission Tomography*, pages 67–75. Academic Press.

Rowe, C., Ng, S., Ackermann, U., Gong, S., Pike, K., Savage, G., Cowie, T., Dickinson, K., Maruff, P., Darby, D., Smith, C., Woodward, M., Merory, J., Tochon-Danguy, H., O'Keefe, G., Klunk, W., Mathis, C., Price, J., Masters, C., and Villemagne, V. (2007). Imaging beta-amyloid burden in aging and dementia. *Neurology*, 68(20):1718–1725.

Rowe, C. C., Ackerman, U., Browne, W., Mulligan, R., Pike, K. L., O'Keefe, G., Tochon-Danguy, H., Chan, G., Berlangieri, S. U., Jones, G., Dickinson-Rowe, K. L., Kung, H. P., Zhang, W., Kung, M. P., Skovronsky, D., Dyrks, T., Holl, G., Krause, S., Friebe, M.,

Lehman, L., Lindemann, S., Dinkelborg, L. M., Masters, C. L., and Villemagne, V. L. (2008). Imaging of amyloid β in Alzheimer's disease with 18F-BAY94-9172, a novel PET tracer: proof of mechanism. *Laneur*, 7(2):129–135.

Rudin, L., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4):259–268.

Rueckert, D., Aljabar, P., Heckemann, R. A., Hajnal, J. V., and Hammers, A. (2006). Diffeomorphic registration using b-splines. In *MICCAI (2)'06*, pages 702–709.

Saha, G. (2010). *Basics of PET Imaging: Physics, Chemistry, and Regulations*. Springer.

Sakamoto, S., Ishii, K., Sasaki, M., Kitagaki, H., Yamaji, S., Matsuda, K., and Mori, E. (1997). Reduction of Cerebellar Glucose Metabolism in Advanced Alzheimer's Disease. *J Nucl Med*, 38(6):925–928.

Scahill, R. I. and Fox, N. C. (2007). Longitudinal imaging in dementia. *Brit J Radiol*, 80:S92–S98.

Schiepers, C. (2006). *Diagnostic nuclear medicine*. Birkhäuser, Boston, MA, USA, 2nd edition.

Schmidtlein, C. R., Kirov, A. S., Nehmeh, S. A., Erdi, Y. E., Humm, J. L., Amols, H. I., Bidaut, L. M., Ganin, A., Stearns, C. W., McDaniel, D. L., and et al. et al. (2006). Validation of GATE Monte Carlo simulations of the GE Advance/Discovery LS PET scanners. *Med Phys*, 33(1):198–208.

Segobin, S., Matthews, J., Markiewicz, P., and Herholz, K. (2010). A hybrid between region-based and voxel-based methods for Partial Volume correction in PET. In *IEEE Nucl Sci Conf R*, pages 3073–3078.

Sethian, J. A. (1996). A fast marching level set method for monotonically advancing fronts. *P Natl Acad Sci USA*, 93(4):1591–1595.

Sethian, J. A. (2001). Evolution, implementation, and application of level set and fast marching methods for advancing fronts. *J Comput Phys*, 169(2):503–555.

Shaw, L. M., Vanderstichele, H., Knapik-Czajka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., Blennow, K., Soares, H., Simon, A., Lewczuk, P., Dean, R., Siemers, E., Potter, W., Lee, V. M.-Y., and Trojanowski, J. Q. (2009). Cerebrospinal fluid biomarker signature in alzheimer's disease neuroimaging initiative subjects. *Ann Neurol*, 65(4):403–413.

Shidahara, M., Tsoumpas, C., Hammers, A., Boussion, N., Visvikis, D., Suhara, T., Kanno, I., and Turkheimer, F. E. (2009). Functional and structural synergy for resolution recovery and partial volume correction in brain PET. *NeuroImage*, 44(2):340–348.

Singh, V., Chertkow, H., Lerch, J. P., Evans, A. C., Dorr, A. E., and Kabani, N. J. (2006). Spatial patterns of cortical thinning in mild cognitive impairment and Alzheimer's disease. *Brain*, 129(11):2885–2893.

Sloane, P. D., Zimmerman, S., Suchindran, C., Reed, P., Wang, L., Boustani, M., and Sudha, S. (2002). The public health impact of Alzheimer's Disease, 2000-2050: Potential implication of treatment advances. *Annu Rev Publ Health*, 23(1):213–231.

Soret, M., Bacharach, S. L., and Buvat, I. (2007). Partial-Volume Effect in PET Tumor Imaging. *J Nucl Med*, 48(6):932–945.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87(2):245–251.

Studholme, C., Hill, D. L. G., and Hawkes, D. J. (1999). An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recogn*, 32(1):71–86.

Talairach, J. and Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging*. Thieme Medical Publishers, New York, NY.

Teo, B.-K., Seo, Y., Bacharach, S. L., Carrasquillo, J. A., Libutti, S. K., Shukla, H., Hasegawa, B. H., Hawkins, R. A., and Franc, B. L. (2007). Partial-Volume Correction in PET: Validation of an Iterative Postreconstruction Method with Phantom and Patient Data. *J Nucl Med*, 48(5):802–810.

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

Thompson, P. M., Hayashi, K. M., de Zubicaray, G., Janke, A. L., Rose, S. E., Semple, J., Herman, D., Hong, M. S., Dittmer, S. S., Doddrell, D. M., and Toga, A. W. (2003). Dynamics of Gray Matter Loss in Alzheimer's Disease. *J Neurosci*, 23(3):994–1005.

Thompson, P. M., MacDonald, D., Mega, M. S., Holmes, C. J., Evans, A. C., and Toga,

A. W. (1997). Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces. *J Comp Assist Tomo*, 21(4):567–581.

Tohka, J. and Reilhac, A. (2006). A Monte Carlo study of deconvolution algorithms for partial volume correction in quantitative PET. In *IEEE Nucl Sci Conf R*, volume 6, pages 3339–3345.

Tohka, J. and Reilhac, A. (2008). Deconvolution-based partial volume correction in Raclopride-PET and Monte Carlo comparison to MR-based method. *NeuroImage*, 39(4):1570–1584.

van Cittert, P. (1931). Zum Einfluß der Spaltbreite auf die Intensitätsverteilung in Spektrallinien. II. *Zeitschrift für Physik*, 69(5-6):298–308.

van de Pol, L., Barnes, J., Scahill, R., Frost, C., Lewis, E., Boyes, R., van Schijndel, R., Scheltens, P., Fox, N., and Barkhof, F. (2007). Improved reliability of hippocampal atrophy rate measurement in mild cognitive impairment using fluid registration. *NeuroImage*, 34(3):1036–1041.

Vandenberghe, R., Thurfjell, L., Van Laere, K., Buckley, C. J., Owenius, R., and Brooks, D. J. (2010a). Comparison between cerebellum and pons as reference regions for quantification of the amyloid imaging agents [18F]flutemetamol and [11C]PIB. *Human Amyloid Imaging 2010 Meeting Abstracts. 2010 April 9, Toronto*.

Vandenberghe, R., Van Laere, K., Ivanoiu, A., Salmon, E., Bastin, C., Triau, E., Hasselbalch, S., Law, I., Andersen, A., Korner, A., Minthon, L., Garraux, G., Nelissen, N., Bormans, G., Buckley, C., Owenius, R., Thurfjell, L., Farrar, G., and Brooks, D. J. (2010b). 18F-flutemetamol amyloid imaging in Alzheimer disease and mild cognitive impairment: A phase 2 trial. *Ann Neurol*, 68(3):319–329.

Verkkoniemi, A., Kalimo, H., Paetau, A., Somer, M., Iwatsubo, T., Hardy, J., and Haltia, M. (2001). Variant alzheimer disease with spastic paraparesis: Neuropathological phenotype. *J Neuropathol Exp Neurol*, 60(5):483–492.

Videen, T. O., Perlmutter, J. S., Mintun, M. A., and Raichle, M. E. (1988). Regional correction of positron emission tomography data for the effects of cerebral atrophy. *J Cereb Blood F Met*, 8(5):662–670.

Villemagne, V. L., Ong, K., Mulligan, R. S., Holl, G., Pejoska, S., Jones, G., O'Keefe, G., Ackerman, U., Tochon-Danguy, H., Chan, J. G., Reininger, C. B., Fels, L., Putz, B., Rohde, B., Masters, C. L., and Rowe, C. C. (2011a). Amyloid Imaging with 18F-Florbetaben in Alzheimer Disease and Other Dementias. *J Nucl Med*, 52(8):1210–1217.

Villemagne, V. L., Pike, K. E., Chételat, G., Ellis, K. A., Mulligan, R. S., Bourgeat, P., Ackermann, U., Jones, G., Szoeke, C., Salvado, O., Martins, R., O'Keefe, G., Mathis, C. A., Klunk, W. E., Ames, D., Masters, C. L., and Rowe, C. C. (2011b). Longitudinal assessment of Aβ and cognition in aging and Alzheimer disease. *Ann Neurol*, 69(1):181–192.

Villemagne, V. L. and Rowe, C. C. (2010). Amyloid PET Ligands for Dementia. *PET Clinics*, 5(1):33–53.

Viola, P. and Wells III, W. M. (1997). Alignment by maximization of mutual information. *Int J Comput Vision*, 24(2):16–23.

Waragai, M., Okamura, N., Furukawa, K., Tashiro, M., Furumoto, S., Funaki, Y., Kato, M., Iwata, R., Yanai, K., Kudo, Y., and Arai, H. (2009). Comparison study of amyloid PET and voxel-based morphometry analysis in mild cognitive impairment and Alzheimer's disease. *J Neurol Sci*, 285(1-2):100–108.

Weiner, M. W., Aisen, P. S., Jack Jr, C. R., Jagust, W. J., Trojanowski, J. Q., Shaw, L., Saykin, A. J., Morris, J. C., Cairns, N., Beckett, L. A., and et al. (2010). The Alzheimer's disease neuroimaging initiative: progress report and future plans. *Alzheimers Dement*, 6(3):202–11.

Wong, K.-P., Wardak, M., Shao, W., Dahlbom, M., Kepe, V., Liu, J., Satyamurthy, N., Small, G., Barrio, J., and Huang, S.-C. (2010). Quantitative analysis of [18F]FDDNP PET using subcortical white matter as reference region. *Eur J Nucl Med Mol Imaging*, 37:575–588.

Yakushev, I., Hammers, A., Fellgiebel, A., Schmidtmann, I., Scheurich, A., Buchholz, H.-G., Peters, J., Bartenstein, P., Lieb, K., and Schreckenberger, M. (2009). SPM-based count normalization provides excellent discrimination of mild Alzheimer's disease and amnestic mild cognitive impairment from healthy aging. *NeuroImage*, 44(1):43 – 50.

Yang, J., Huang, S., Mega, M., Lin, K., Toga, A., Small, G., and Phelps, M. (1996). Investigation of partial volume correction methods for brain FDG PET studies. *IEEE T Nucl Sci*, 43(6):3322–3327.

Yoo, T. S., Ackerman, M. J., and Lorensen, W. E. (2002). Engineering and algorithm design for an image processing API: A technical report on ITK-the insight toolkit. *Proc. of Medicine Meets Virtual Reality*, pages 586–592.

Zitova, B. (2003). Image registration methods: a survey. *Image Vision Comput*, 21(11):977–1000.