

Joint Analysis of PET/MR Data for Improved PET Quantification

Catherine J. Scott

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of
University College London.

Department of Medical Physics and Biomedical Engineering
University College London

April 25, 2021

I, Catherine J. Scott, 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 work.

For Marmalade

Abstract

Quantitative pharmacokinetic analysis of Positron Emission Tomography (PET) data typically requires a dynamic scan of at least one hour, which poses a challenge for both clinical and research studies. Instead, in standard practice, a static 10 minute scan is used to calculate the standardised uptake value ratio (SUVR). SUVR approximates tracer binding but is biased by blood flow changes, rendering it unsuitable for longitudinal studies. In this thesis, the availability of magnetic resonance imaging (MRI) data, simultaneously acquired from a PET-MR scanner is exploited to reduce the time required for accurate PET quantification.

The main body of this work comprises the development of a framework to incorporate blood flow information from arterial spin labelled (ASL) MRI data into the existing simplified reference tissue model (SRTM) to replace the early phase of the PET data, reducing the acquisition time. This reduced acquisition time (RT-) SRTM was evaluated on [^{18}F]-florbetapir data for the estimation of both regional average and voxelwise amyloid burden (BP_{ND}), and was validated against the gold standard BP_{ND} using a 60 minute scan.

The first step of the RT-SRTM requires the PET tracer delivery parameter, R_1 , to be estimated from the ASL cerebral blood flow (CBF) maps. Several methods were evaluated: linear regression using region as a covariate, multi-atlas propagation with image fusion, and deep learning based regression using a convolutional neural network. The RT-SRTM was shown to facilitate accurate regional voxelwise quantification in half the acquisition time (30 minutes).

Additionally, deep learning based regression was used to learn the model which maps ASL-CBF and dynamic PET data to BP_{ND} in a single step (SS-DL). The SS-DL model exploits all available information, and avoids noise sensitive voxelwise fitting. This allows the acquisition time to be cut to 15 minutes, and facilitates accurate voxelwise BP_{ND} quantification on a time-scale manageable for almost all patients and studies.

Impact Statement

The novel methodologies developed in this thesis, along with the analysis performed have several direct impacts on research. This work has quantified the error and bias inherent in using the standardised uptake value ratio (SUVR) to estimate amyloid burden, which is the current standard practice. This may influence how future research studies are designed and analysed, as well as influencing the analysis of on-going studies such as Insight46. Furthermore, the reduced acquisition time methods developed in this work (RT-SRTM and SS-DL) provide a potential solution to the pressing, unmet need for a method of accurately quantifying amyloid burden within an acquisition time which is both tolerable and cost-effective. Furthermore, the comparison of cerebral blood flow estimation using arterial spin labelled (ASL) magnetic resonance imaging (MRI) with that derived from dynamic amyloid positron emission tomography (PET) data highlights the potential to use either method as an imaging biomarker, and provides an insight into the limitations of each technique. All of the methods developed and conclusions drawn in this thesis can potentially be translated into clinical practice in the future.

The work within this thesis has already been disseminated through one published peer reviewed journal paper, with two more under development. Where possible open source code is used and sufficient detail is given to allow other researchers to implement the methods described. This work has also been presented at a series of international conferences, namely the Medical Image Computing and Computer Assisted Intervention Conference (MICCAI 2016,2017,2018), the PET and SPECT/MR Conference (PSMR 2017,2018),

the Functional Neuroreceptor Mapping of the Living Brain Symposium (NRM 2018) and the Human Amyloid Imaging Conference (HAI 2018).

Acknowledgements

Firstly, I would like to offer huge thanks to my supervisors; to Prof. Sébastien Ourselin for supporting me through the highs and particularly the lows of my PhD years, to Prof. Brian Hutton for always pushing me to improve my ideas and strengthen my work, to Prof. Jonathan Schott for providing level headed advice and guidance as well as most of the data used here, and to Dr Jieqing Jiao for always encouraging me to pursue whatever makes me happy.

I am also greatly indebted to the many people that I have consulted, including Dr Michela Antonelli, Dr Marco Borri, Dr Ninon Burgos, Dr Jorge Cardoso, Dr Dave Cash, Dr Kjell Erlandsson, Dr Pawel Markiewicz, Dr Andrew Melbourne, Dr David Thomas, and Dr Enrico De Vita. Thanks also to Dr Gargi Banerjee and Prof. David Werring for the use of and help with the BOCAA data. My gratitude as well goes to all the participants in the Insight46, YOAD and BOCAA studies for making this work possible.

My sincerest thanks to my *compagni di merende* (don't Google that), for teaching me the essential skills required both for research and for life. Huge thanks also to my research group and fellow CDT students, who ensured that the social side of research was never neglected. Particular thanks to Carla Semedo and David Owen, to whom I owe at least one MICCAI paper due to their pep talks spurring me on to the small hours.

Finally, I will never find the words to sufficiently express my gratitude to my family, from the smallest to the biggest, who are far more important than anything written here. Including, of course, my partner Marco, who stuck by my side through everything- *possiamo andare in vacanza ora, per favore?*

Doc: Marty, it's perfect, you're just not thinking fourth dimensionally!

Marty: Right, right. I have a real problem with that.

BACK TO THE FUTURE, PART III

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 30 |
| 1.1 | Motivation and Aim | 32 |
| 1.2 | Outline | 33 |
| 1.3 | Thesis Contributions | 33 |
| 2 | General Background and Theory | 36 |
| 2.1 | Imaging in Alzheimer’s Disease | 36 |
| 2.2 | Positron Emission Tomography (PET) | 38 |
| 2.2.1 | PET detection | 39 |
| 2.2.2 | PET reconstruction | 40 |
| 2.2.3 | PET-CT | 43 |
| 2.3 | Magnetic Resonance Imaging (MRI) | 43 |
| 2.3.1 | Nuclear magnetic resonance- origin of the signal | 44 |
| 2.3.2 | Creating a magnetic resonance image | 46 |
| 2.3.3 | 3D T1-weighted imaging for brain volume measurements | 50 |
| 2.4 | Simultaneous PET-MRI | 50 |
| 2.4.1 | Motivation | 50 |
| 2.4.2 | Challenges | 52 |
| 3 | Review of Amyloid PET Quantification Theory Literature | 54 |
| 3.1 | Introduction to Amyloid PET Tracers | 54 |
| 3.2 | Quantification of Dynamic PET data | 55 |
| 3.2.1 | Background theory | 55 |

- 3.2.2 Pharmacokinetic modelling using arterial blood sampling 59
- 3.2.3 Simplified PK models (without arterial sampling) 62
- 3.3 Quantification of static PET data 67
 - 3.3.1 Standardised uptake value ratio (SUVR) 68
 - 3.3.2 Limitations of SUVR Quantification 69
- 3.4 Selection of a Gold Standard for Amyloid Quantification 71
- 3.5 Conclusion 73

- 4 Blood flow modelling and quantitative ASL 74**
 - 4.1 Introduction 74
 - 4.2 Blood flow modelling 75
 - 4.2.1 Indicator dilution theory 75
 - 4.2.2 The Renkin Crone model 77
 - 4.3 Magnetic Resonance Imaging based measurement of Blood Flow 78
 - 4.3.1 Principles of Arterial Spin Labelling 79
 - 4.3.2 Models for Arterial Spin Labelling 80
 - 4.3.3 ASL quantification considerations 82
 - 4.4 Conclusion 85

- 5 Simplified Reference Tissue Model with reduced acquisition time 87**
 - 5.1 Introduction 87
 - 5.2 Methods 89
 - 5.2.1 Gold standard PET kinetic modelling 89
 - 5.2.2 RT-SRTM framework 90
 - 5.2.3 Data acquisition and pre-processing 92
 - 5.2.4 Data 94
 - 5.2.5 Statistical Analysis 96
 - 5.3 Results 96
 - 5.3.1 Comparison of proposed method with gold standard . . . 96

| | | |
|----------|---|------------|
| 5.3.2 | Influence of CBF-derived R_1 estimation on amyloid quantification | 98 |
| 5.3.3 | Influence of C_R extrapolation on amyloid quantification . | 100 |
| 5.4 | Discussion and Conclusion | 100 |
| 6 | Correlation of PET tracer delivery and ASL cerebral blood flow | 103 |
| 6.1 | Introduction | 103 |
| 6.1.1 | Comparison of ASL-CBF with H_2O^{15} -PET | 104 |
| 6.1.2 | Comparison of amyloid PET- R_1 with H_2O^{15} -PET | 107 |
| 6.1.3 | Correlation of R_1 and ASL for amyloid PET/MR data . | 109 |
| 6.2 | Methods | 110 |
| 6.2.1 | Data | 110 |
| 6.2.2 | CBF estimation with ASL-MRI | 110 |
| 6.2.3 | Dynamic PET acquisition and reconstruction | 111 |
| 6.2.4 | Regional analysis framework | 112 |
| 6.2.5 | Linear regression between ASL-CBF and PET- R_1 | 112 |
| 6.3 | Results | 113 |
| 6.3.1 | Global relationship between ASL-CBF and PET- R_1 | 113 |
| 6.3.2 | Regional relationship between ASL-CBF and PET- R_1 | 114 |
| 6.3.3 | Additional covariates | 120 |
| 6.4 | Conclusions | 121 |
| 7 | Optimisation of Reduced Acquisition Time PET implementation | 122 |
| 7.1 | Introduction | 122 |
| 7.2 | Methods | 123 |
| 7.2.1 | Data | 123 |
| 7.2.2 | Extrapolation of PET reference region, C_R | 123 |
| 7.2.3 | Optimisation of C_R extrapolation and PET acquisition window | 125 |

7.2.4 Sensitivity analysis of RT-SRTM 127

7.3 Results 128

7.3.1 Extrapolation of PET reference region, C_R 128

7.3.1.1 Optimisation of *PCA* C_R method 128

7.3.1.2 Comparison of *PCA* C_R and *scaled mean* C_R methods: 130

7.3.2 Optimisation of acquisition time window 132

7.3.3 Optimisation of SUVR acquisition window 134

7.3.4 Comparison of proposed RT-SRTM with gold standard . 135

7.3.5 Comparison of proposed RT-SRTM with *SUVR*-1 . . . 135

7.3.6 Comparison of proposed RT-SRTM with short acquisition time PET 137

7.3.7 Sensitivity analysis of RT-SRTM 138

7.4 Discussion 142

7.5 Conclusions 145

8 Comparison of R_1 estimation methods: towards voxelwise analysis **147**

8.1 Introduction 147

8.2 R_1 estimation approaches 149

8.2.1 Linear regression 150

8.2.2 Multi-atlas propagation with image fusion (IF) (Scott et al., 2017) 151

8.2.2.1 Database construction 151

8.2.2.2 Registration to test subject 152

8.2.2.3 Similarity calculation 152

8.2.2.4 Image fusion 153

8.2.3 Deep learning based regression 154

8.2.3.1 Neural network organisation 154

8.2.3.2 Training a Convolutional Neural Network . . . 156

8.2.3.3 Regularisation. 158

| | | |
|----------|---|------------|
| 8.2.3.4 | ‘Highres3Dnet’ architecture and training | 159 |
| 8.3 | Experimental set-up | 160 |
| 8.3.1 | RT-SRTM implementation | 160 |
| 8.3.2 | Comparison methods | 161 |
| 8.3.3 | Data | 162 |
| 8.3.4 | Data pre-processing | 162 |
| 8.3.5 | Validation and Analysis | 163 |
| 8.4 | Results and discussion | 164 |
| 8.4.1 | Pre-processing and SUVR optimisation | 164 |
| 8.4.1.1 | ASL-CBF map smoothing | 164 |
| 8.4.1.2 | SUVR-1 timing optimisation | 165 |
| 8.4.2 | Regional R_1 estimation comparison | 166 |
| 8.4.3 | Voxelwise R_1 estimation comparison | 169 |
| 8.4.4 | Voxelwise BP_{ND} estimation comparison | 178 |
| 8.5 | Conclusions | 183 |
| 9 | Direct PET-MR binding potential map synthesis using CNNs | 187 |
| 9.1 | Introduction | 187 |
| 9.2 | Methods | 189 |
| 9.2.1 | Deep Learning Framework for BP_{ND} Estimation (SS-DL) | 189 |
| 9.2.2 | Comparative BP_{ND} Estimation Methods | 190 |
| 9.2.3 | Acquisition window definition | 191 |
| 9.2.4 | Data | 191 |
| 9.3 | Results and discussion | 191 |
| 9.3.1 | Method Comparison Over Different Timing Windows | 191 |
| 9.3.2 | Optimised Timing Window Method Comparison | 194 |
| 9.3.2.1 | 30 minute optimised acquisition. | 195 |
| 9.3.2.2 | 15 minute optimised acquisition. | 198 |
| 9.4 | Conclusions | 200 |

| | |
|---|------------|
| 10 General Conclusions | 203 |
| 10.1 Summary | 203 |
| 10.2 Future Research Directions | 205 |
| Appendices | 207 |
| A Additional data information | 207 |
| A.1 Data sets used | 207 |
| A.1.1 Insight46 (1946) | 207 |
| A.1.2 Young Onset Alzheimer’s Disease (YOAD) | 207 |
| A.1.3 Biomarkers and Outcomes in Cerebral Amyloid An- giopathy (BOCAA) | 208 |
| A.2 Data protocols | 208 |
| A.3 Subject amyloid status | 209 |
| B Data processing | 211 |
| B.1 PET reconstruction | 211 |
| B.2 ASL quantification | 212 |
| B.3 pCT generation | 212 |
| B.4 Pharmacokinetic modelling | 212 |
| Bibliography | 213 |

List of Figures

| | | |
|-----|---|----|
| 2.1 | The hypothetical model of biomarker abnormality originally proposed in Jack et al. (2010) and updated in Jack et al. (2013) (Reproduced from Jack et al. (2013) with permission by Elsevier.) | 37 |
| 2.2 | The effect of an RF pulse oscillating at the Larmor frequency on the net magnetisation of spins. The flip angle illustrated here is 90° | 45 |
| 2.3 | Schematics showing how TE and TR can be used to select the contrast in a spin echo (SE) sequence. | 48 |
| 3.1 | 1 tissue compartment model, where C_p and C_T are the tracer concentrations in blood plasma and tissue respectively, K_1 is the transfer rate constant from blood to tissue ($\text{mL} \cdot \text{cm}^{-3} \cdot \text{min}^{-1}$), and k_2 is the transfer rate constant from tissue to blood (min^{-1}). | 58 |
| 3.2 | Schematics showing compartmental models used in Price et al. (2005) | 60 |
| 3.3 | Schematics showing the compartments and exchange constants for reference tissue models | 62 |
| 4.1 | schematic of capillary system | 76 |
| 4.2 | Illustration of the 1 tissue compartment Renkin-Crone model | 77 |

| | | |
|-----|--|-----|
| 4.3 | Illustration of ASL acquisition for cerebral blood flow mapping, showing a) the control image where the arterial blood net magnetisation (red arrow) remains aligned with the external B_0 field, then b) the label image where the arterial blood is tagged using an RF pulse, and finally c) showing that the difference between the control and label images is proportional to perfusion. | 79 |
| 5.1 | Overview- parcellation is registered to PET and ASL to calculate regional average values. ASL data is converted into CBF-derived R_1 values using the linear regression relationship. A population reference tissue time activity curve of 0:60 mins combined with the measured reference tissue data (30:60 mins) is used with the CBF-derived R_1 and the measured PET tissue data (30:60 mins), to apply the modified simplified reference tissue model to estimate BP_{ND} | 95 |
| 5.2 | Division of data for training and testing | 96 |
| 5.3 | Estimated amyloid burden against the gold standard value calculated using full PET time series. | 97 |
| 5.4 | Example amyloid positive ($A\beta^+$) subject (subject 6). Top: Regional average binding potential maps for (left to right) gold standard, proposed method using gold standard R_1 , proposed method using ASL and $SUVR-1$. Bottom: difference maps compared to gold standard. | 98 |
| 5.5 | Example amyloid negative ($A\beta^-$) subject (subject 2). Top: Regional average binding potential maps for (left to right) gold standard, proposed method using gold standard R_1 , proposed method using ASL and $SUVR-1$. Bottom: difference maps compared to gold standard. | 99 |
| 5.6 | Estimated binding potential plotted against the gold standard value calculated using full PET time series | 100 |

| | | |
|-----|---|-----|
| 5.7 | Extrapolated reference regions (C_R) | 101 |
| 6.1 | Pearson correlation coefficient between H_2O^{15} -PET and ASL reported in comparison studies, as a function of the time elapsed between scans. Reproduced from Fan et al. (2016) with permission. | 108 |
| 6.2 | Division of data for optimisation | 111 |
| 6.3 | Correlation of PET- R_1 with ASL-CBF for 20 optimisation set subjects. Linear regression was calculated and applied to the ASL-CBF data to show the residual error in the fit. | 113 |
| 6.4 | Correlation of PET- R_1 with ASL-CBF for 20 optimisation set subjects with a linear regression performed for each tissue type. \triangle indicates $A\beta$ -positive subjects and \circ $A\beta$ -negative | 115 |
| 6.5 | Regional plots of ASL-CBF normalised to the reference region against PET- R_1 | 116 |
| 6.6 | Examples of subjects with flow artefacts affecting the CBF estimation. This shows that the affected regions include the accumbens (red), pallidum (green), putamen (yellow), and cerebral white matter (pink). | 118 |
| 6.7 | Top: CBF map overlaid on T1 showing fat shift artefact, Bottom: parcellation over T1 showing that the artefact appears in the occipital region (purple) | 119 |
| 7.1 | Division of data between testing and optimisation | 124 |
| 7.2 | MSE in the fit of $C_R(t = 30, 60)$ using the <i>PCA</i> C_R method when optimising the number of components used and the upper and lower bounds for the weights using leave-one-out analysis on 39 optimisation set subjects | 129 |

- 7.3 Plot of mean square error in the fit of C_R using the *PCA* C_R method and different ranges of weighting factors for 24 subjects (30:60 mins PET data). Upper and lower bound of the weighting factor are \pm weighting range times standard deviation. Coloured lines represent individual subjects and the black line is the average across all subjects. 130
- 7.4 Mean error for in the fit of $C_R(t = 30, 60)$ using the *PCA* C_R method across timing windows with difference constrains on a leave-one-out analysis on 39 optimisation set subjects 131
- 7.5 ME in estimated C_R using *PCA* C_R method and *scaled mean* C_R method compared to measured C_R calculated using leave-one-out cross validation in 39 optimisation set subjects. 131
- 7.6 Error in BP_{ND} estimates across different timing windows using different estimates of C_R calculated using leave-one-out cross validation in 39 optimisation set subjects.. . . . 133
- 7.7 Plot of average SUVR against time post injection for a representative subset of 10 testing set subjects showing that a plateau is reached at approximately 50 minutes post injection, using cerebellar grey matter as a reference region. 135
- 7.8 Estimated amyloid burden against the gold standard value calculated using full PET time series for 25 testing set subjects. The grey shaded region covers the 95% confidence interval in the regression. 136
- 8.1 Generation of R_1 maps from ASL-CBF maps using linear regression (LR). Result image is for illustration only. 150
- 8.2 Registration transforms input database images into the candidate subject space. The local similarity between the images is calculated and is used to weight the propagation of R_1 information from the output database into candidate space. Images are for illustration only and don't represent actual results. 153

- 8.3 schematics to show the building blocks of feed-forward neural networks where each circle represents a unit or neuron 155
- 8.4 Plot generated by Tensorboard showing the training set (orange) and validation set (blue) loss as a function of training iterations. The arrow indicates the point at which the model starts to overfit to the training data and training should be stopped. 157
- 8.5 Overview of ‘Highres3Dnet’ architecture showing the 20 convolutional layers of the network and the paired residual connections, adapted from Li et al. (2017). 158
- 8.6 The workflow for the DL method, where the model is trained using ASL-CBF and T1-w MRI as the input with PET- R_1 as the output. Inference can then be applied to the test set using the trained model. Images are for illustration only and don’t represent actual results. 160
- 8.7 The voxelwise correlation coefficient between relative ASL-CBF and PET- R_1 as a function of the Gaussian smoothing kernel applied to the ASL-CBF map. This was used to determine the kernel used to smooth the ASL-CBF maps in the database. . . . 165
- 8.8 Parameter maps for subject BOCAA-017, fitted from $t=0:60$ mins of PET data showing that regional analysis (a) does not adequately describe the tracer delivery for this patient due to stroke induced R_1 reductions which is not confined to a single region, shown by the voxelwise R_1 map in b). c) shows reduced tracer delivery in the inferior portion of the cerebellum (outlined in white) affecting much of the grey matter. This demonstrates that grey matter cerebellum may not be a suitable reference region for this subject. 169

| | | |
|------|--|-----|
| 8.9 | MSE in estimated voxelwise R_1 (lines) plotted by subject along the x-axis. Shapes above the lines indicate the subject classification and curved brackets show the study the subject came from. | 170 |
| 8.10 | Parameter maps for subject 1946-156 who is a healthy amyloid- β negative ($A\beta^-$) subject. | 172 |
| 8.11 | Parameter maps for subject 1946-221 who is a healthy amyloid- β positive ($A\beta^+$) subject. | 174 |
| 8.12 | Parameter maps for subject YOAD-057 who is a amyloid- β positive ($A\beta^+$) patient with clinically diagnosed AD. | 176 |
| 8.13 | Parameter maps for subject BOCAA-05 who has cerebral amyloid angiopathy. | 177 |
| 8.14 | Parameter maps for subject BOCAA-17 who has cerebral amyloid angiopathy. | 179 |
| 8.15 | MSE in estimated voxelwise BP_{ND} (lines) plotted by subject along the x-axis. Shapes above the lines indicate the subject classification and curved brackets show the study the subject came from. | 182 |
| 9.1 | Overview of methods tested where true R_1 is R_1^* . Blue boxes indicate input data from the test subject and green boxes show population data from the training database. | 190 |
| 9.2 | MSE for whole brain at different timing windows. | 193 |
| 9.3 | MSE for GM voxels at different timing windows. | 194 |
| 9.4 | Boxplots of voxelwise error using 30 minutes PET data. | 196 |
| 9.5 | Voxelwise estimates of BP_{ND} for a 15 minute acquisition. | 199 |
| 9.6 | Voxelwise error using 15 minutes PET data. | 201 |
| A.1 | Imaging protocol showing simultaneous PET and multi-parametric imaging. | 209 |

List of Tables

| | | |
|-----|---|-----|
| 7.1 | The percentage of the variation explained using increasing number of principal components following PCA on the $C_R(t = 0, 60)$ of 39 optimisation set subjects | 128 |
| 7.2 | MSE and ME between gold standard BP_{ND}^* and BP_{ND} at different 30-minute acquisition windows averaged across 16 regions and 39 optimisation set subjects. True C_R uses the true reference region curve, PCA C_R and mean C_R extrapolate the reference region curve using the <i>PCA C_R</i> and <i>scaled mean C_R</i> methods, respectively. All methods used the gold standard R_1^* | 132 |
| 7.3 | Mean square error (MSE) and bias (mean error) between the gold standard BP_{ND}^* and $SUVR-1$ at different acquisition windows across 25 testing set subjects. | 134 |
| 7.4 | Percentage absolute error in $BP_{ND}+1$ using the RT-SRTM on simulated curves compared to gold standard BP_{ND}^*+1 . Each value was averaged across the 200 noise realisations for simulated data. The dashed lines indicate the upper limits of R_1 % error and noise found for regional analysis. | 140 |
| 7.5 | Mean absolute error in $BP_{ND}+1$ using the RT-SRTM on simulated curves compared to gold standard BP_{ND}^*+1 . SRTM is the gold standard fit using the full SRTM for comparison. Each value was averaged across the 200 noise realisations for simulated data. The dashed lines indicate the upper limits of R_1 % error and noise found for regional analysis. | 141 |

| | | |
|-----|--|-----|
| 8.1 | Mean squared error between $SUVR-1$ and BP_{ND} averaged over 75 subjects and estimated at different timing windows. | 166 |
| 8.2 | Regional mean squared error (MSE) and mean error (ME) in R_1 estimates averaged over 75 subjects. | 168 |
| 8.3 | Voxel mean squared error (MSE) in R_1 estimates averaged over 75 subjects using all brain voxels (WB) and grey matter (GM) voxels only. | 169 |
| 8.4 | Voxel mean squared error (MSE) in BP_{ND} estimates averaged over 75 subjects. | 178 |
| 8.5 | Voxel mean squared error (MSE) in BP_{ND} estimates averaged over 74 subjects (excluding BOCAA17). | 183 |
| 9.1 | Start time (t_s) in minutes since injection which gives the lowest MSE for a given scan duration (t_d) for each method. | 195 |
| 9.2 | Voxel mean squared error (MSE) in BP_{ND} estimates averaged over 74 subjects using all brain voxels (WB) and grey matter (GM) voxels only using 30 minutes of PET data. | 197 |
| 9.3 | Voxel mean squared error (MSE) in BP_{ND} estimates averaged over 74 subjects using all brain voxels (WB) and grey matter (GM) voxels only, for 15 minutes of PET data | 200 |
| A.1 | Summary of mean cortical grey matter SUVR values in the dataset used for two different reference regions with the number of subjects defined as amyloid positive ($a\beta^+$). | 210 |

Nomenclature

General Phamacokinetic modelling

$V_T = \frac{C_T}{C_p}$: volume of distribution in target tissue which is the ratio of the tracer concentration in the target tissue, C_T , to the concentration held in blood plasma, C_p .

$DVR = \frac{V_T}{V_{ND}}$: distribution volume ratio, where V_{ND} is the non-displaceable volume of distribution.

$BP_{ND} = DVR - 1$: non-displaceable binding potential.

$C_{compartment}$: tracer concentration within each compartment.

k_{number} : rate constant of the rate of change between compartments.

Specific notation used in chapters 5-9

$t = t_s, t_e$: this indicates the start (t_s) and end (t_e) times of the scan where $t = 0$ at injection. For the gold standard method $t_s, t_e = 0, 60$ minutes.

$R_1 = \frac{K_1}{K_1'}$: the rate constant of tracer delivery to the target tissue (K_1) relative to the reference tissue (K_1')

k_2 : rate constant from target tissue to blood.

$BP_{ND} = \frac{f_{ND} B_{max}}{k_D}$: the non-displaceable binding potential, where f_{ND} is the tissue free fraction, k_D is the equilibrium dissociation constant and B_{max} is the available target concentration. Since f_{ND} and k_D remain constant across subjects, BP_{ND} is proportional to the target density.

R_1^*, k_2^*, BP_{ND}^* : these are the gold standard parameter values estimated using the SRTM on PET data from $t = 0, 60$ minutes.

$C_T(t)$: a vector containing the tracer concentration in the target tissue over

time.

$C_R(t)$: a vector containing the tracer concentration in the reference tissue over time.

$C_R^{pop}(t)$: a matrix containing the reference region concentration over time for a population of subjects

$\overline{C}_R^{pop}(t)$: a vector containing the mean population tracer concentration over time.

$C_T^\ddagger(t) = C_T(t) - R_1 C_R(t)$: $C_T^\ddagger(t)$ is a dummy variable to group the pre-defined values together.

Acronyms

| | |
|-------------|---------------------------------|
| 3D | three-dimensional |
| AIF | Arterial Input Function |
| AD | Alzheimer's disease |
| APD | Avalanche photodiodes |
| ASL | Arterial Spin Labelling |
| AC | attenuation correction |
| ATT | arterial transit time |
| BBB | blood brain barrier |
| CASL | continuous ASL |
| CT | Computed Tomography |
| CBF | cerebral blood flow |
| CSF | cerebral spinal fluid |
| CNN | convolutional neural network |
| DSC | dynamic susceptibility contrast |
| DCE | dynamic contrast enhanced |
| DVR | distribution volume ratio |

| | |
|--------------|--------------------------------------|
| DL | deep learning |
| EPI | echo planar imaging |
| FDG | Fluorodeoxyglucose |
| FSE | Fast Spin Echo |
| FBP | filtered back projection |
| FWHM | full width half maximum |
| GPU | graphics processing unit |
| GE | gradient echo |
| GIF | geodesic information flows |
| GRASE | gradient and spin echo |
| HC | healthy control |
| IF | image fusion |
| IR | inversion recovery |
| LR | linear regression |
| LOR | line of response |
| LNCC | locally normalised cross correlation |
| MR | Magnetic resonance |
| MRI | magnetic resonance imaging |
| ME | mean error |
| MCI | mild cognitive impairment |
| MSE | mean squared error |

| | |
|----------------|---|
| MT | magnetisation transfer |
| MP-RAGE | magnetisation prepared rapid acquisition by gradient echo |
| OSEM | ordered subset expectation maximisation |
| PET | Positron Emission Tomography |
| PiB | Pittsburgh compound B |
| PK | Pharmacokinetic |
| pCT | pseudo CT |
| pCASL | pseudo continuous ASL |
| pASL | pulsed ASL |
| PVE | partial volume effect |
| PVC | Partial volume correction |
| PC | phase contrast |
| PLD | post labelling delay |
| PMT | photomultiplier tube |
| PSF | point spread function |
| PD | proton density |
| PCA | Principal component analysis |
| RTM | reference tissue model |
| ROI | region of interest |
| RF | radiofrequency |
| RNN | recurrent neural networks |

| | |
|-------------|-----------------------------------|
| SNR | signal to noise ratio |
| SRTM | simplified reference tissue model |
| SUVR | standardised uptake value ratio |
| SE | spin echo |
| SPGR | spoiled gradient echo |
| SiPM | silicon photomultiplier |
| TAC | time activity curve |
| TR | repetition time |
| TE | echo time |
| TI | inversion time |
| TOF | time-of-flight |

Chapter 1

Introduction

Positron Emission Tomography (PET) facilitates the quantification of a range of important biomarkers through the injection and detection of targeting radiotracers. To interpret the measured signal and derive the biological parameters of interest, data are collected dynamically from injection, covering radiotracer delivery to tissue, interaction with the target, and tracer washout. This provides a map of the spatio-temporal concentration of the tracer *in vivo*. A pharmacokinetic model which describes these processes may then be fitted to these dynamic data to estimate the biological parameters such as radiotracer target density.

Depending on the radiotracer administered, the dynamic data acquisition time required to fit the model may be 60 minutes or more. This is prohibitive in a clinical context due to patient discomfort, restrictions on scanner time availability, and the increased chance of subject motion which corrupts the data. Consequently a simplified technique is commonly employed.

The standardised uptake value ratio (SUVR) is a measure of relative tracer uptake which can be calculated from a static scan lasting approximately 10 minutes. SUVR is calculated by dividing the tracer concentration within the tissue of interest by the concentration in a reference region. The reference region consists of tissue considered to be free of the radiotracer target and represents the non-displaceable (ND) tracer concentration (i.e. tracer in the tissue which is not bound to the intended target). When the ratio of the tracer

concentration within the target tissue and the reference tissue has reached a steady-state, SUVR approximates the distribution volume ratio (DVR). The DVR is related to the density of the imaging target and can also be estimated from pharmacokinetic modelling, see section 3.2.

However, as SUVR is calculated from a single static scan, the tracer concentration present during the acquisition will depend on the delivery and washout rates of the tracer, as well as on target density. Tracer delivery and washout are intrinsically linked to blood flow, and since blood flow can change during the progression of disease (Benedictus et al., 2016), and indeed fluctuate over the course of a day (Parkes et al., 2004), SUVR estimates may be confounded.

The influence of cerebral blood flow (CBF) changes on SUVR estimates has been highlighted in longitudinal Alzheimer’s disease studies in which the target of interest was the protein amyloid- β (van Berckel et al., 2013). Amyloid- β is an early indicator of disease onset and a therapeutic target, hence accurately quantifying amyloid- β density is of paramount importance. Here, variation in blood flow has been shown to cause spurious changes in SUVR which are unrelated to target density (van Berckel et al., 2013; Cselényi and Farde, 2015). Conversely, target density estimates derived from pharmacokinetic modelling of dynamic data starting from radiotracer injection can account for blood flow, as tracer delivery is parameterised within the model.

Dynamic PET data can be divided into two phases: the early phase, in which the signal is dominated by tracer delivery to tissue (Hsiao et al., 2012), and the late phase, which contains information related to tracer binding and washout, and is where SUVR is estimated. The intrinsic correlation between tracer delivery and CBF has been demonstrated for an amyloid- β tracer in Chen et al. (2015). Therefore, if CBF can be measured independently from the PET acquisition, then the data acquisition time may be reduced such that only the late phase data are acquired to estimate the remaining parameters.

The gold standard for the measurement of CBF uses radiolabelled water

as a PET tracer ($\text{H}_2\text{O}^{15}\text{-PET}$). However, this approach requires invasive arterial blood sampling, and the data cannot generally be acquired concurrently with another PET study as the signal from the two tracers cannot easily be distinguished; therefore the overall acquisition time would likely increase.

Arterial Spin Labelling (ASL)-magnetic resonance imaging (MRI) is a non-invasive imaging technique which applies a magnetic ‘tag’ to arterial blood, such that it can be used as an endogenous contrast agent. ASL can be used to estimate CBF, as validated by comparison with the gold standard $\text{H}_2\text{O}^{15}\text{-PET}$ (Fan et al., 2016). Whilst the accuracy of the technique is dependent on the implementation, high quantitative accuracy has been achieved when ASL data are normalised to a reference region (Goetti et al., 2014).

The introduction of combined PET-MRI scanners, which facilitate simultaneous acquisition, means that ASL and PET data can be acquired concurrently. By combining CBF information from ASL, to provide early phase delivery information, with the dynamic late PET data, the total acquisition time can be significantly reduced. Thus amyloid- β burden can be accurately quantified in a clinically feasible time frame, increasing patient comfort and throughput without sacrificing quantitative accuracy.

1.1 Motivation and Aim

The potential to provide robust PET target density estimates within a clinically feasible time frame means that longitudinal studies could be carried out without results being confounded by changes in blood flow. Furthermore, this would also increase the volume and range of patients eligible due to the reduced scan time. The main aim of this work was to explore ways in which the blood flow information from ASL-MRI can be combined with dynamic PET data to improve PET quantification. In this work the methods are applied to amyloid PET data for the imaging of Alzheimer’s disease (AD), as this is a patient population who would find long scans particularly challenging and in which changes in blood flow are inherent both in disease progression and nor-

mal ageing. However, the techniques used are not amyloid specific and could be applied to any PET tracer with similar biokinetics.

1.2 Outline

This thesis is structured as follows: the relevant background and theory of AD and the imaging modalities used are given in chapter 2, while chapter 3 reviews current quantification techniques for PET and their limitations including the susceptibility of simplified techniques to changes in blood flow, and chapter 4 describes how blood flow can be quantified using MRI. Chapter 5 then describes a methodology to exploit the simultaneous measurement of blood flow using MRI and PET to reduce the overall data acquisition time from 60 to 30 minutes, by adapting a standard kinetic model (RT-SRTM). Chapter 6 looks in more detail at the correlation between measures from PET and MRI, which are related to blood flow, such that in chapter 7 the RT-SRTM can be optimised and validated against the gold standard. Chapters 8 and 9 then explore novel image processing techniques to extend the proposed methodology for use on a voxelwise basis.

1.3 Thesis Contributions

The following are a list of outcomes from this PhD.

Journal Papers

- Catherine J Scott, Jieqing Jiao, Andrew Melbourne, Ninon Burgos, David M Cash, Enrico De Vita, Pawel J Markiewicz, Antoinette O'Connor, David L Thomas, Philip Sj Weston, Jonathan M Schott, Brian F Hutton, and Sébastien Ourselin. Reduced acquisition time PET pharmacokinetic modelling using simultaneous ASL-MRI: proof of concept. *Journal of Cerebral Blood Flow and Metabolism*, page 271678X18797343, sep 2018c. ISSN 15597016. doi: 10.1177/0271678X18797343. URL <http://www.ncbi.nlm.nih.gov/pubmed/30182792>

Peer reviewed conference papers

- Catherine J Scott, Jieqing Jiao, M Jorge Cardoso, Kerstin Kläser, Andrew Melbourne, Pawel Markiewicz, Jonathan M. Schott, Brian F. Hutton, and Sébastien Ourselin. Short Acquisition Time PET/MR Pharmacokinetic Modelling using CNNs. *In: Frangi, A., Schnabel, J., Davatzikos, Alberola-López, C., Fichtinger, G. (eds) MICCAI 2018. LNCS, 11070 LNCS:48–56, sep 2018a. ISSN 16113349. doi: 10.1007/978-3-030-00928-1_6. URL https://link.springer.com/chapter/10.1007/978-3-030-00928-1_6*
- Catherine J Scott, Jieqing Jiao, M Jorge Cardoso, Andrew Melbourne, Enrico De Vita, David L Thomas, Ninon Burgos, Pawel Markiewicz, Jonathan M. Schott, Brian F. Hutton, and Sébastien Ourselin. Short Acquisition Time PET Quantification Using MRI-Based Pharmacokinetic Parameter Synthesis. *In: Descoteaux M., Maier-Hein L., Franz A., Jannin P., Collins D., Duchesne S. (eds) MICCAI 2017. LNCS, 10434:737–744, 2017. doi: 10.1007/978-3-319-66185-8. URL <http://link.springer.com/10.1007/978-3-319-66185-8>*
- Catherine J Scott, Jieqing Jiao, Andrew Melbourne, Jonathan M Schott, Brian F Hutton, and Sébastien Ourselin. ASL-incorporated Pharmacokinetic Modelling of PET Data With Reduced Acquisition Time: Application to Amyloid Imaging. *International Conference on Medical Image Computing and Computer-Assisted Intervention, LNCS 9902:406–413, 2016. doi: 10.1007/978-3-319-46720-7. URL <http://link.springer.com/10.1007/978-3-319-46720-7>*

Conference abstracts

- Catherine J Scott, Jieqing Jiao, Kerstin Kläser, M. Jorge Cardoso, Andrew Melbourne, Pawel J Markiewicz, Jonathan M Schott, Brian F Hutton, and Sébastien Ourselin. Reduced acquisition time PET/MR quan-

tification: a comparison of tracer delivery estimation methods. In *Conference on PET and SPECT MR (PSMR) 2018*, 2018b

- Catherine J Scott, Jieqing Jiao, Andrew Melbourne, Pawel J Markiewicz, Jonathan M Schott, Brian F Hutton, and Sébastien Ourselin. Reduced acquisition time amyloid PET quantification: sensitivity analysis. In *The XII International Symposium of Functional Neuroreceptor Mapping of the Living Brain (NRM) 2018*, 2018d
- Catherine J Scott, Jieqing Jiao, Andrew Melbourne, Pawel J Markiewicz, Jonathan M Schott, Brian F Hutton, and Ourselin Sébastien. Reduced acquisition time PET quantification applied to [18f]-florbetapir. In *Human Amyloid Imaging Conference (HAI) 2018*, 2018e
- Catherine J Scott, Jieqing Jiao, Andrew Melbourne, David L Thomas, Enrico De Vita, Pawel J Markiewicz, Jonathan M Schott, Brian F Hutton, and Sébastien Ourselin. Reduced acquisition time PET/MR quantification: optimisation for amyloid-pet data. In *Conference on PET and SPECT MR (PSMR) 2017*, 2018f

Chapter 2

General Background and Theory

This chapter describes the importance of imaging in Alzheimer's disease and the role played by both positron emission tomography (PET) and magnetic resonance imaging (MRI). It also gives the basic theory behind these imaging techniques in addition to the motivation for simultaneous PET-MRI and the instrumentation which facilitates acquisition.

2.1 Imaging in Alzheimer's Disease

Alzheimer's Disease (AD) is the leading cause of dementia, which affects over 45 million people worldwide, a number which is expected to double every 20 years due to an increasing and ageing population (Prince et al., 2015). There is currently no cure, and available treatments may relieve symptoms and slow, but not stop, progression¹.

The leading model for the progression of AD is based on the amyloid cascade hypothesis (Karran et al., 2011). Here, the disease begins with the deposition of the mis-folded protein amyloid-beta ($A\beta$) throughout the brain. This protein, if not cleared, builds up over time, forming plaques which are thought to initiate a cascade including deposition of tau, neuroinflammation

¹National Institute for Health and Care Excellence (NICE): Donepezil, Galantamine, Rivastigmine and Memantine for the treatment of Alzheimer's Disease, <https://www.nice.org.uk/guidance/ta217/resources> 2016

leading to synaptotoxicity and neuronal cell death (atrophy) and cognitive decline. There is much debate about whether the deposition of $A\beta$ and tau is independent, causative or linked via an upstream process (Jack et al., 2013) and whilst other, more complex, theories have been proposed, the accumulation of $A\beta$ is still thought to be a central process within the disease (Herrup, 2015) and $A\beta$ accumulation has been shown to occur approximately 15 years before symptom onset (Bateman et al., 2012).

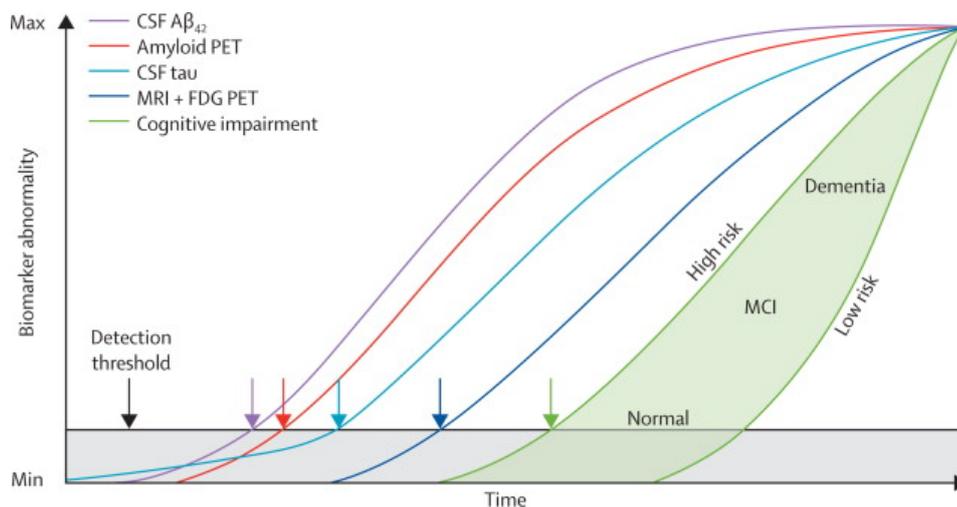


Figure 2.1: The hypothetical model of biomarker abnormality originally proposed in Jack et al. (2010) and updated in Jack et al. (2013) (Reproduced from Jack et al. (2013) with permission by Elsevier.)

According to the theoretical model proposed in Jack et al. (2010), and updated based on available evidence in Jack et al. (2013), there are several biomarkers of AD which become abnormal at different points along the progression of the disease. Figure 2.1 shows these biomarkers and the order in which they appear to become abnormal: 1) $A\beta$ measured in cerebral spinal fluid (CSF) (CSF $A\beta_{42}$), closely followed by 2) $A\beta$ measured using PET (Amyloid PET), then 3) tau protein in CSF (CSF tau), followed much later by 5) atrophy measured using MRI and hypo-metabolism from Fluorodeoxyglucose (FDG) PET, and eventually 6) cognitive impairment.

The model proposed by Jack et al. (2013) takes into account the different detection thresholds associated with each of the biomarkers. This is used to

address the fact that tau deposition is often seen from *post mortem* studies before A β accumulation is seen. Consequently figure 2.1 shows CSF tau as the first biomarker to start to rise, however it reaches the detection threshold after CSF A β_{42} and amyloid PET. This could potentially be revised following the introduction of tau targeting PET tracers (Hahn et al., 2017).

From the model in Jack et al. (2013) it is clear that early detection of A β is essential both for early diagnosis and as a potential target for therapy (Liu-Seifert et al., 2015). Imaging biomarkers are vital for the diagnosis and monitoring of AD, as they allow the disease to be studied in a minimally invasive manner at multiple time points. From the biomarkers highlighted by Jack et al. (2010) it is clear that the use of both PET and MRI are essential for detecting and monitoring the disease.

2.2 Positron Emission Tomography (PET)

PET uses exogenous tracers which have two properties: a chemical structure such that they target particular biological substances or processes when introduced into the body (usually injected intravenously), and a positron emitting nucleus. When an emitted positron meets an electron they annihilate, producing two 511 keV photons which travel in opposite directions. These photons can be measured by PET scanner detectors, and can be used to reconstruct the spatial distribution of the tracer in the body based on the position of the annihilation events. Due to the small concentration of tracer required to construct an image this technique provides a minimally invasive method to measure biological processes *in vivo* without causing changes to the processes or targets being measured. Common examples of processes imaged using PET include glucose metabolism (FDG), blood flow (H₂O¹⁵ or radio-water) and protein deposition (e.g. amyloid and tau). The following brief explanation of PET imaging is adapted from Bailey et al. (2005).

2.2.1 PET detection

Each annihilation event results in two approximately collinear photons, thus when two photons are detected within a given period of time (coincidence window) they are considered to come from a single event. A line of response (LOR) is drawn between the two points at which the photons were detected as it is assumed that the event occurred somewhere along this line. To maximise the number of photons detected, PET detectors are typically arranged in a ring around the patient. If the time at which each of the photons was detected is measured with sufficient accuracy then the difference in timing can be used to reduce the likely position of the annihilation event along the LOR, as the difference in detection time is directly related to the distance that each photon has to travel. If this information is used in image reconstruction it is known as time-of-flight (TOF) PET.

Standard clinically used PET detectors consist of a series of components:

- The scintillation crystal: converts the 511keV photon to lower energy visible photons, where the number of photons is proportional to the energy deposited in the crystal.
- The photomultiplier tube (PMT): converts the scintillation photons into photoelectrons which are then multiplied to produce a current which is also proportional to the energy of the detected photon. Some modern scanners use alternatives to PMTs, see section 2.4.
- The electronics: which include coincidence, timing and positioning circuits, and write the data to a file. If the position and timing for each detected event are recorded then this is a list-mode acquisition, resulting in a list-mode file. This allows the user to have full control over the image reconstruction.

PET detectors are generally arranged into cylindrical rings such that the detected LORs may be drawn in 3D.

2.2.2 PET reconstruction

To reconstruct the detected photons into images of the tracer distribution the origin of the annihilation photons must be estimated. The most straightforward method of PET image reconstruction is filtered back projection (FBP). Here the detected photons are back projected along the LOR where each position along this line is given equal value. The image is then filtered to reduce blurring (remove low frequencies) and control noise (remove high frequencies). This technique is computationally fast and is linear, thus the reconstruction is not influenced by the tracer distribution. However for FBP there is no way to model the imperfect and stochastic nature of the photon emission and detection process. Instead iterative reconstruction methods can be used, the most common of which is ordered subset expectation maximisation (OSEM).

For iterative reconstruction methods, an estimate of the image is forward projected to the detectors and compared with the detected counts, and an update of the image estimate is generated. This is repeated until convergence or until an adequate image is estimated. For OSEM only a subset of image projections are used to update the estimate each time, making it more computationally efficient. The number of iterations and subsets are defined by the user, where an iteration is completed when all subsets have been used to update the estimate. The higher the number of iterations the closer the estimated image is to the maximum likelihood solution, however the noise also increases such that the choice of the iterations and subsets is a compromise between noise and accuracy. TOF information can be incorporated such that the probability of the location of the emission event along the LOR is no longer uniform but can be isolated to a region in the order of a few cms, speeding up convergence which reduces the noise and increases the contrast to noise ratio.

During the image reconstruction several processes should be modelled to improve accuracy of the reconstructed tracer distribution.

Normalisation. For normalisation the detectors are uniformly irradiated producing a map where the difference in intensities is determined only by the dif-

ferent sensitivities of the detectors, either intrinsic or geometric. The inverse of this map can then be used in the reconstruction to account for this effect.

Randoms. Random coincidences occur when two photons from independent annihilation events are detected within the same coincidence window. This causes the incorrect LOR to be drawn between the detected photons, and the event origin to be incorrectly estimated. The randoms rate depends on the rate of single photon detections (singles) on each detector pair and on the length of the timing window. The higher the singles rate and the longer the coincidence window, the higher the randoms rate. The randoms rate is usually estimated using a delayed coincidence channel where the timing signal from one channel is delayed before going to the coincidence circuits (Markiewicz et al., 2017). In this case there will be no true coincidences, as the data has been delayed, so all coincidences can be counted as randoms and subtracted from the prompt channel to leave true and scattered coincidences.

Scatter. These are detected photons which have undergone scatter which causes a change in photon direction and a loss of energy. Compton scatter is the primary scatter process for photons of 511 keV. Some scatter events can be removed from the reconstruction process by only accepting events within an energy range centred on 511 keV (energy window) however some scattered photons will still fall within this range (scatter fraction) (Zaidi and Koral, 2004). The scatter fraction is generally modelled using an analytical simulation for each LOR using the attenuation map (see attenuation below), the scatter probability at a point is given by the Klein-Nishina formula and the efficiencies of the detectors which define the LOR (Ollinger, 1996). This methodology was extended to 3D voxelwise scatter estimation by Markiewicz et al. (2017) for a parallelised implementation on the graphics processing unit (GPU).

Attenuation. Photons that undergo photoelectric absorption and scatter which are consequently not detected are said to be attenuated. The number of photons lost to attenuation will depend on the electron density of the object and the distance through which the photons travel. This is because Compton

scattering, which is the main contributor to attenuation in PET, is caused by the interaction of a photon with a charged particle, in this case electrons. Attenuation in PET is usually estimated using Computed Tomography (CT) data, as the contrast of these images is dependent on electron density. However in CT the photon energies used are in the order of 100keV, rather than 511keV as in PET, consequently a bilinear mapping process is often employed to convert CT images from Hounsfield units to attenuation (μ) maps (Burger et al., 2002).

Dead time. This is defined as the time required between events to accurately detect them as separate events. This is particularly important in dynamic scanning where high count rates are measured soon after injection.

Decay correction. This accounts for the exponential radioactive decay of the tracer over time which is determined by the half-life of the radionuclide used. The counts recorded in the PET data are corrected to the time of injection ($t = 0$).

Partial Volume Effects. The partial volume effect (PVE) in PET comprises two components; the tissue fraction effect, due to the representation of tissue in voxels in the order of mm, and the intrinsic resolution of the scanner which causes blurring of the signal in the image. The resolution of the scanner is quantified by using the full width half maximum (FWHM) of the point spread function (PSF). Contributions to the size of the FWHM include the size of the scintillation detector elements, the positron range in tissue, acolinearity of the annihilation photons, decoding of multiplexed signal from crystal to photodiode, depth of gamma penetration into scintillator crystals, and the non-uniform sampling of the space within the scanner (Moses, 2011). Partial volume correction (PVC) can be applied within the reconstruction by modelling the PSF, or post reconstruction on the image itself (Erlandsson et al., 2012).

2.2.3 PET-CT

Modern clinical whole-body PET scanners are almost always combined with a CT scanner as a hybrid modality. CT provides anatomical information through three-dimensional (3D) x-ray transmission scans in which the contrast is provided by differences in electron density within the imaged object. This allows the PET uptake to be localised within the anatomy, and the electron density information can be used in PET reconstruction to model attenuation and scatter. Whilst this is a hybrid imaging technique, PET/CT scanners are designed for sequential acquisition as CT is a significantly shorter acquisition. This can cause issues for PET reconstruction if the patient position is different between scans.

Since the contrast in CT is determined by the electron densities within the imaged object, there is excellent contrast between bone and soft tissue. However, there is poor contrast between different soft tissues; this is a problem for studying conditions such as Alzheimer's disease, which involve soft tissue changes in the brain which cannot be detected using CT, so additional imaging is required.

2.3 Magnetic Resonance Imaging (MRI)

MRI is a highly versatile imaging modality which is used for a wide range of tasks including both structural and functional imaging. MRI is widely acknowledged to have superior soft tissue contrast compared to CT and is routinely used in neuroimaging. MRI exploits the effects of magnetic fields in manipulating hydrogen nuclei, or spins, which are abundant in the human body. The signal recorded in MRI is related to the proton density, binding of the spins in tissue, and the chemical environment surrounding the spins and can therefore be used to create images in which the contrast is dependent on any combination of these effects.

Furthermore, MRI is not a purely anatomical imaging modality like CT. MRI can also be used to measure a wide range of parameters including tissue

diffusivity and structural connectivity (diffusion weighted MRI), the concentration of deoxyhaemoglobin to infer brain activation and functional connectivity (functional MRI), the relative concentrations of metabolites (spectroscopy), and blood flow (arterial spin labelled MRI). Each of these measurements has the potential to be combined with PET images to give complementary information and to improve the characterisation of disease. A brief description of the basic theory behind the image acquisition is given here.

2.3.1 Nuclear magnetic resonance- origin of the signal

MRI uses magnetic fields to manipulate the spins of nuclei, or ‘spins’, with unpaired protons. Most commonly hydrogen nuclei are targeted, largely due to their abundance in the human body. MRI utilises three magnetic fields 1) the main, uniform magnetic field, B_0 , which remains constant and is usually 1.5 or 3T in whole-body clinical scanners, 2) the spatially varying directional gradients ($G_{x/y/z}$) and 3) the radiofrequency (RF) pulses (B_1 field).

Under normal conditions, the magnetic moments of the spins within the body are randomly oriented such that they cancel out and the net magnetisation is zero. However, when placed within a magnetic field, the spins will either align or anti-align their intrinsic magnetic moment with the external field, where the proportion of spins in each orientation is described by the Boltzmann distribution. Aligning with the magnetic field requires less energy than opposing it, and the Boltzmann distribution determines that the stronger the magnetic field, the greater the proportion of aligned spins. The resulting magnetisation determines the available signal.

$$2\pi f = -\gamma_{gyro}B \quad (2.1)$$

The frequency at which the nuclei precess in the magnetic field is known as the Larmor frequency (f) and is related to the magnetic field strength, B , through the gyromagnetic ratio, γ_{gyro} , which is a constant for each type of nucleus, equation 2.1. The Larmor frequency is the resonant frequency of the

system, and a circularly polarised RF pulse oscillating at the Larmor frequency can tip the net magnetisation of the spins from the direction of the B_0 field (usually denoted z) to produce a transverse component of the signal in the $x - y$ plane, see figure 2.2. This allows the signal from the spins to be detected as their signal in the z -direction is negligible compared to the magnitude of the B_0 -field. The angle through which the magnetisation is tipped is known as the flip angle.

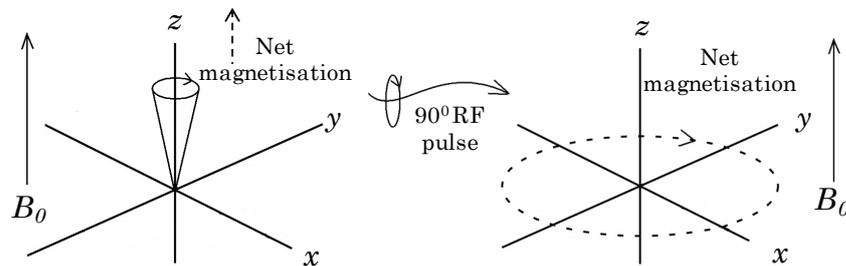


Figure 2.2: The effect of an RF pulse oscillating at the Larmor frequency on the net magnetisation of spins. The flip angle illustrated here is 90° .

The net magnetisation of the spins in the x-y plane can then be detected in a receive coil, as it undergoes free inductive decay, until all the spins return to their equilibrium position. The signal decays through two simultaneous and independent mechanisms; spin-lattice relaxation and spin-spin relaxation.

Spin-lattice relaxation is the longitudinal recovery of the magnetisation in the direction of B_0 . Here the energy of the anti-aligned spins absorbed from the RF pulse is dissipated via the surrounding tissue, or ‘lattice’, to realign with the external B_0 field. This random process can be described as an exponential recovery and T1 is defined as the time taken for the magnetisation in the z direction to recover to $1/e$ of its equilibrium value.

Spin-spin relaxation is the decay of the transverse magnetisation due to the dephasing of spins. Here the energy is transferred between spins as they interact with one another. Pure spin-spin relaxation occurs when the magnetic fields of two nearby nuclei combine to produce a local magnetic field, with which they realign themselves. This process is repeated many times across all

of the spins until they are all completely dephased, and the transverse magnetisation is zero. This process is described by an exponential decay where T2 is defined as the time it takes after the RF pulse for the transverse magnetisation to decay to $1/e$ of its value.

Further dephasing of spins occurs due to local inhomogeneities in the static field caused by intrinsic defects in the magnet and differences in magnetic susceptibilities in the imaged object, such as at air-tissue interfaces. In this case T2* is used in the place of T2, to indicate that dephasing is due to both spin-spin interaction and B_0 inhomogeneities.

2.3.2 Creating a magnetic resonance image

Phase and frequency encoding. To create an image, the measured signals must contain spatial information. Frequency and phase encoding use the relationship between the Larmor frequency and the applied magnetic field to encode this information into the image.

For phase encoding, a spatially varying magnetic field gradient is applied along one of the scanner axes, for example the y-axis, in addition to the constant, homogeneous, B_0 field (z direction). This gradient field, G_y ensures that the spins experience different magnetic field strengths depending on their y-position, and hence they precess at different frequencies, as defined by equation 2.1. When the gradient is switched off the spins return to their original Larmor frequencies, however their phase is dependent on their position along the y-axis.

Frequency encoding works by the same principle, except that the gradient is not turned off before the signal is recorded. This means that during the acquisition, the spins precess at a frequency determined by their position (in this case along the x-direction) within the gradient field.

The encoding gradients are applied orthogonally to one another such that along the y-axis the spins will have different phases, and along the x-axis they will precess at different frequencies. This gives each position on the x-y plane a unique phase-frequency signature, which is used to calculate its position in

space.

Basic pulse sequences. Magnetic resonance (MR) images are acquired using a combination of RF pulses and gradients arranged into a pulse sequence. The pulse sequence dictates the ordering and duration of the RF pulses and gradients which will determine the contrast seen in the resulting image.

The measured signal acquired during the read out portion of the pulse sequence is encoded into the frequency space (k-space) through the frequency encoding gradient. The pulse sequence needs to be repeated a number of times to fill k-space. Here the duration of the sequence is defined as the repetition time (TR). The final image is produced by applying a Fourier transform to the k-space data.

One of the most basic pulse sequences is the spin echo (SE) sequence. In a SE sequence a 90° RF excitation pulse flips the magnetisation in the z-direction (M_z) into the transverse plane, and a subsequent 180° RF pulse refocuses the spins to compensate for $T2^*$ decay. This produces a signal ‘echo’ at the echo time (TE) after the initial 90° pulse, having applied the 180° pulse at TE/2.

In the SE sequence the 180° pulse compensates for the dephasing due to local B_0 inhomogeneities, therefore the signal decay is characterised by the relaxation constant T2 rather than $T2^*$. However, overall the measured signal will depend on the proton density (PD), T1 and T2 of the tissue being imaged. The weighting of the image, which determines which property dominates the image contrast, can be defined by altering TE and TR, as shown by figure 2.3.

The gradient echo (GE) is another basic pulse sequence in which the spin rephasing which creates the echo is achieved using gradient reversal. Here a negative gradient is applied directly after the RF excitation pulse and this causes rapid spin dephasing. A positive gradient is then applied which reverses the dephasing to form an echo. It should be noted that only the dephasing due to the negative gradient is reversed, and not the dephasing due to local field inhomogeneities, so the transverse signal decay is described by $T2^*$ rather than T2.

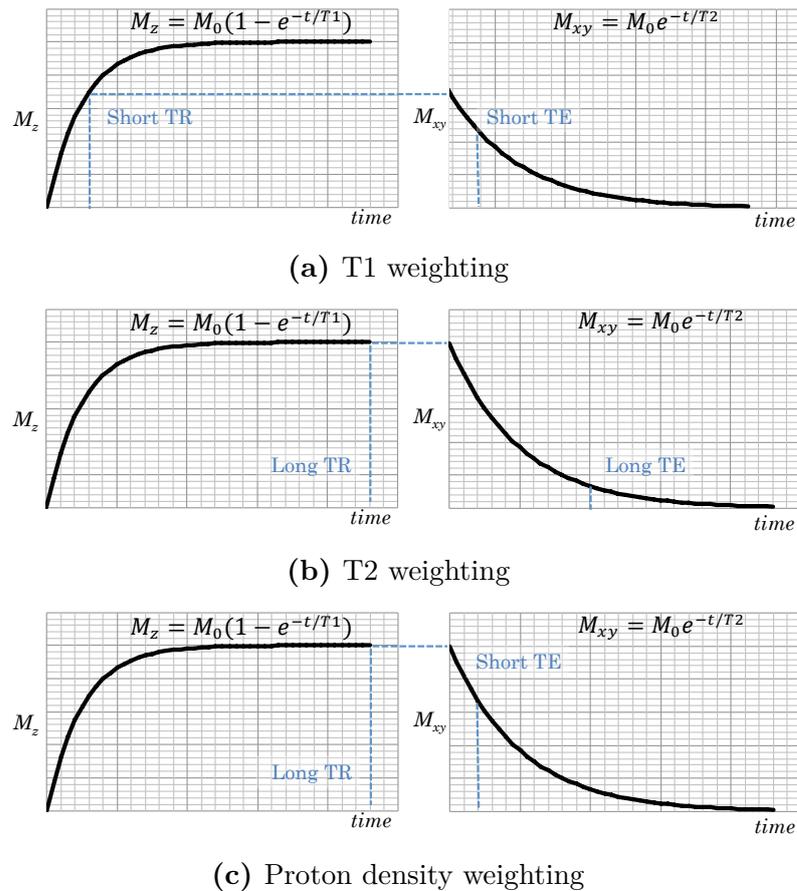


Figure 2.3: Schematics showing how TE and TR can be used to select the contrast in a spin echo (SE) sequence.

In the inversion recovery (IR) sequence the RF excitation pulse is preceded by an inversion pulse, which is normally 180° , where the time between the inversion and the excitation pulse is called the inversion time (TI). Here the initial longitudinal magnetisation (M_0) is inverted such that $M_z = -M_0$, and will recover along the z direction at a rate determined by T1. By choosing an appropriate TI, it is possible to null the initial magnetisation of a specific tissue of a given T1, and therefore eliminate its contribution to the image signal. For instance, TI can be chosen to null the signal from fluid (long TI) or fat (short TI) based on their T1.

Clinical pulse sequences. In clinical applications basic sequences are rarely adopted, as covering adequate portions of anatomy with sufficient spatial resolution would require impractically long acquisition times. A number of strate-

gies have been devised to decrease the time required to complete the image volume acquisition. These include multi-slice imaging, where multiple non-contiguous 2D slices are acquired at the same time, and parallel imaging and partial Fourier approaches, which allow k-space undersampling using different strategies to reconstruct the image. Since the encoding of the image is based on the production of signal echoes, another way to accelerate conventional imaging is by acquiring a number of echoes during a single TR, and using each echo to encode a different line of k-space. For a SE sequence this is achieved by adding n refocusing 180° pulses after the first one, thus producing n consecutive echoes; this sequence is called Turbo or Fast Spin Echo (FSE).

It is also possible to collect multiple echoes by reversing the frequency encoding gradient multiple times during a TR, and use the phase encoding to cover multiple lines of k-space. This approach is referred to as echo planar imaging (EPI), and can be applied to both SE and GE sequences. This type of sequence allows very short acquisitions but is prone to a number of artefacts, largely due to its sensitivity to susceptibility effects and B_0 inhomogeneity.

Instead of a stack of 2D slices, image volumes can be obtained by performing 3D acquisitions, and this enables isotropic (or close to isotropic) voxels. In 3D imaging an additional phase encoding gradient is added such that a volumetric k-space can be acquired. Covering 3D k-space requires a large number of echoes to be encoded, and therefore fast imaging acquisitions are employed.

One way to accelerate image acquisition is to use spoilers which are either RF pulses or gradients applied to dephase residual transverse magnetisation before the next excitation pulse is applied. The spoiled gradient echo (SPGR) is a basic GE sequence with additional spoilers to destroy the residual transverse magnetisation after readout. This sequence reaches a steady state for the longitudinal magnetisation, and the T1 weighting increases with flip angle.

2.3.3 3D T1-weighted imaging for brain volume measurements

In the model of progression for Alzheimer’s disease proposed by Jack et al. (2013) brain atrophy, measured using structural MRI, is a key biomarker which tracks closely with progressive cognitive impairment. To measure atrophy in different brain regions, volume measurements of the regions must be made from high resolution structural scans. T1-weighted 3D structural imaging is widely performed using magnetisation-prepared variants of SPGR sequences such as the magnetisation prepared rapid acquisition by gradient echo (MP-RAGE), as recommended by the Alzheimer’s Disease Neuroimaging Initiative (ADNI)². MP-RAGE includes an additional 180° inversion ‘preparation’ pulse to ensure T1 weighting at short TRs, and allows 3D high spatial resolution images to be obtained in approximately 5 minutes. The brain regions can then be delineated either manually on the images, or using automated parcellation algorithms which segment and label the brain images based on an atlas. Common approaches to automated T1 parcellation include multi-atlas propagation techniques such as geodesic information flows (GIF) (Cardoso et al., 2015). Here multi-atlas propagation with image fusion is used to generate tissue label maps in the space of the patient’s T1 image, see section 8.2.2.

2.4 Simultaneous PET-MRI

2.4.1 Motivation

Whilst CT images are ideal for providing electron density information for image reconstruction, the clinical information obtained can be limited, particularly when applied in neuroimaging. By contrast, MRI has the capability to provide structural images with excellent soft tissue contrast as well as a wealth of other clinically useful information. Each of the different types of MR imaging has the potential to be combined with PET images to give complementary information and to improve the disease detection and characterisation.

²<http://adni.loni.usc.edu/methods/mri-tool/mri-analysis/>

Combining PET and MRI images has several other advantages, including improvements in PET image quality using structural MRI data for partial volume correction (Thomas et al., 2011), and reconstruction with anatomical priors (Vunckx et al., 2012). There is also a potential dose reduction if MRI is used in place of CT, since MRI uses non-ionising rather than ionising radiation, which is particularly important for paediatric or pregnant patients.

PET and MRI data from separate sessions can be combined through image registration to remap the imaging data into a common space, and a whole-body sequential PET-MRI scanner has been developed by Philips (Zaidi et al., 2011) to minimise movement between the scans. Further advantages associated with simultaneous acquisition include:

- Temporal alignment of the scans: this ensures that the biological state of the patient (e.g. blood flow) has not changed between scans (Fan et al., 2016).
- Optimised acquisition time: the longer acquisition time required for MRI (typically 30 minutes for a neurological brain protocol or up to 60 minutes for research studies (Lane et al., 2017)) allows the PET acquisition time (approximately 10 minutes for a brain scan) to be extended to improve quantification, see chapter 3.
- Motion estimation: patient motion can be estimated from MRI data, due to its high temporal resolution, and applied to PET data (Catana, 2015).
- Inherent spatial registration of the two modalities: this makes it easier to combine the information from the datasets. This is particularly important outside the brain, where differences in positioning between separate scans result in non-rigid deformations which can be difficult to estimate.
- Reduced number of hospital visits for the patient: as two scans are performed in a single session. This is particularly important for patients who require general anaesthesia for scans such as young children.

2.4.2 Challenges

There are several challenges which must be addressed when considering simultaneous PET-MRI acquisition, such as developing compatible instrumentation, performing attenuation correction of the PET data, and the clinical implementation of this new workflow.

Instrumentation. The largest obstacle to overcome is how to physically combine PET and MRI hardware into a single bore system (Vandenberghe and Marsden, 2015). Firstly, there are physical constraints due to the space requirements for both systems and the difficulty of fitting one inside the other whilst ensuring adequate space for the patient.

Furthermore, once the systems are combined, interference from one system to another must be minimised. The MRI scanner produces a high strength static magnetic field, time varying magnetic fields and RF pulses to generate images. The static magnetic field will interfere with the functionality of a standard PMT's due to the influence of the field on the path of the charged electrons. Furthermore, the changing fields causes eddy currents in any conductive components which also leads to interference, vibration and heating, and the RF pulses cause further interference which can reduce the detected PET count rate.

In turn the PET hardware can affect the operation of the MRI scanner, as any changes in magnetic susceptibility affect the uniformity of the static field and spatial linearity of the changing field. This effect is compounded by eddy currents induced in PET hardware, which produces further magnetic distortion. Finally, since the MRI scanner detects the signal to create the image as RF from the patient, any other sources of RF, such as unshielded power cables, can affect the recorded signal.

Driven by the advantages of simultaneous PET-MRI acquisition, whole-body scanners are now commercially available. This first was the Siemens Biograph mMR (Delso et al., 2011), which uses MR compatible Avalanche photodiodes (APD)s in place of traditional PMTs, mounted between the MR

body coil and the gradient coils. More recently the GE Signa PET/MR (Levin et al., 2016) was released, which uses silicon photomultiplier (SiPM)s instead of APDs; these have a better timing resolution, facilitating TOF PET reconstruction (Roncali and Cherry, 2011).

Attenuation and Scatter correction. As described in section 2.2.2, PET image reconstruction requires photon attenuation and scatter to be taken into account. In PET-CT both effects were estimated using the electron density information from the CT data, which is unavailable for PET-MR. Instead, synthetic CTs, or pseudo CTs, can be generated from MR data or estimated from the PET emission data (Nuyts et al., 1999). MRI based techniques include assigning population attenuation coefficients for bone, soft tissue and air to segmented MR images (Keereman et al., 2010), and synthesising the CT using a database of MRI-CT pairs using deep learning (Ladefoged et al., 2019) or based on their similarity to the target MRI scan using multi-atlas propagation (Burgos et al., 2015), see chapter 8 for details. In a review of available techniques conducted by Ladefoged et al. (2017) it was concluded that the issue of MR attenuation correction (AC) for adult brains has been solved to a quantitatively acceptable degree.

Clinical Implementation. As with any new imaging modality, the final challenge for PET-MRI is practical implementation. Combining these two modalities requires new imaging protocols, room design, safety procedures, different patient workflows, and specialised staff training as dual training in both modalities is rare (Parikh et al., 2015). As the number of simultaneous scanners installed increases, 7 scanners in the UK as of 2017³, there is a growing need to understand both modalities, in order to determine the key applications for simultaneous PET-MR imaging (Bailey et al., 2018).

³<https://mrc.ukri.org/publications/browse/review-of-pet-within-the-medical-imaging-landsca>

Chapter 3

Review of Amyloid PET

Quantification Theory

Literature

This chapter covers the theory of quantification techniques which have been applied to amyloid PET data in the literature. It also reviews the comparative literature to assess the accuracy and suitability of the various techniques, as well as their limitations, to determine the appropriate gold standard approach.

3.1 Introduction to Amyloid PET Tracers

The first tracer successfully developed for imaging amyloid- β plaques in humans was ^{11}C Pittsburgh compound B (PiB), which was developed by modifying a thioflavin based histological dye used for *post mortem* staining (Klunk et al., 2004). Unlike FDG which is incorporated into a cell through glucose transporters where it remains trapped, the PiB radioligand binds reversibly to amyloid- β protein. Due to the short half life of ^{11}C (approximately 20 minutes), such tracers are not considered to be clinically viable, so this led to the development of three ^{18}F radiotracers which have been approved for clinical use; Florbetapir (Amyvid, Eli Lilly and Company), Florbetaben (Neuraceq, Piramal Imaging), and Flutemetamol (Vizamyl, GE Healthcare), see Morris et al. (2016) for details. Since it was developed first, most of the validation of

quantification methods has been carried out using PiB data, as demonstrated in the following review. More recent work using ^{18}F -Florbetapir is also included in the review as this tracer is used for the experimental parts of this thesis.

For amyloid tracers, once injected into the blood stream a proportion of the tracer is carried in the blood plasma, and may be transferred across the blood brain barrier (BBB). Initially the tracer is free within the tissue i.e. it is unbound. However it will then undergo either specific or non-specific binding. Specific binding occurs when the tracer molecule interacts with its target, which in this case means that the radioligand binds to the amyloid- β protein which forms the plaques in the brain. For amyloid tracers this is a reversible process such that the tracer will become unbound and eventually washed out of the tissue (Price et al., 2005; Choi et al., 2009). Non-specific binding, which is an undesirable but often unavoidable process where the tracer binds with non-target materials, also occurs across the brain.

The tracer may also be metabolised in the body and converted into another compound, referred to as a metabolite. These metabolites are carried in the blood and could contribute to the PET signal, confounding the results. In general it is assumed that metabolites will not cross the BBB (Klunk et al., 2004; Choi et al., 2009), thus their contribution to the signal in the brain is negligible. However, a significant proportion of the activity in the vasculature will contain metabolites which must be considered within the analysis.

3.2 Quantification of Dynamic PET data

3.2.1 Background theory

Pharmacokinetic Modelling Nomenclature. For amyloid PET quantification, the aim is to determine the amount of $\text{a}\beta$ deposition across the brain. Consequently the quantification is analogous to neuro-receptor studies as it is the density of target binding sites which is of interest. Pharmacokinetic (PK) models describe the processes of tracer uptake, binding and washout mathe-

matically, and are used to generate physiologically meaningful metrics such as the density of the tracer target. Generally pharmacokinetic modelling is applied to reconstructed PET images, though it is also possible to perform direct reconstruction of kinetic parameters from dynamic PET data. This approach can reduce noise in the parameter estimates and even provide motion information (Jiao et al., 2017), though this is currently rarely used in practice. The notation used in this thesis follows that recommended by Innis et al. (2007) for *in vivo* imaging of reversibly binding radioligands.

In general the density of the tracer target is expressed either as the volume of distribution in target tissue (V_T), or the binding potential (BP). V_T represents the ratio of the unmetabolised tracer concentration in target tissue (C_T) compared to that in plasma (C_P) at equilibrium, as described by equation (3.1), and has units of $\text{mL} \cdot \text{cm}^{-3}$. The volume of distribution in tissue can be divided by the non-displaceable (ND) volume of distribution to give the DVR (equation(3.2)). Here, non-displaceable is used to mean tracer which is not bound to the target, as this tracer cannot be displaced by adding more unlabelled tracer which binds to the same target, and includes both free and non-specifically bound tracer.

$$V_T = \frac{C_T}{C_P} \quad (3.1)$$

$$DVR = \frac{V_T}{V_{ND}} = \frac{C_T}{C_{ND}} \quad (3.2)$$

Binding potential represents the receptor density multiplied by the tracer affinity to the target, and is often defined as the ratio of specifically bound tracer to non-displaceable tracer (BP_{ND}), equation (3.3), which is unitless. Since the affinity is an intrinsic property of the tracer, and is therefore constant across subjects, differences in BP_{ND} reflect differences in target density. Equation (3.3) demonstrates how, by definition, BP_{ND} and DVR are closely related. The way in which these parameters (V_T , DVR , BP_{ND}) are calculated

depends on the PK model used.

$$BP_{ND} = \frac{V_T - V_{ND}}{V_{ND}} = DVR - 1 \quad (3.3)$$

Compartmental Modelling Nomenclature. Compartmental models are based on the assumptions that the system can be broken down into distinct compartments where each compartment consists of a chemical species in a physical space (e.g. unmetabolised amyloid tracer within brain tissue), the tracer is uniformly distributed within each compartment, and the tracer exchanges between the compartments at a fixed rate. The tracer concentration within each compartment is denoted here by $C_{compartment}$ and the rate constants which describe the exchange of tracer between the compartments by k_{number} . Here, models are defined by the number of tissue compartments (TC) and the number of k parameters (P), and the concentration of tracer in plasma is not considered to be a compartment. Depending on the complexity of the tracer kinetics *in vivo*, a different number of compartments and rate constants may be included in the model in different configurations.

One of the simplest kinetic models is the 1 tissue compartment 2 parameter (1T2P or 1TC) model, which is set up as in figure 3.1. The model is driven by the arterial input function (AIF) which is the concentration of the parent (unmetabolised) tracer in blood plasma ($C_P(t)$) which is the tracer input into the system. Generally arterial blood plasma is used instead of whole blood as only the tracer held in plasma is available to exchange with tissue. This is normally measured via invasive arterial blood sampling using *in vitro* analysis. The tracer concentration in tissue, $C_T(t)$, is taken from the PET imaging data acquired dynamically over time to cover the tracer delivery to tissue and subsequent wash out. This dynamic data can also be referred to as a time activity curve (TAC) which can either be defined for a single voxel, or by region using a spatial average of voxels based on anatomical segmentation.

There are 2 rate constants associated with the 1TC model; K_1 which is the transfer rate constant from plasma to tissue ($\text{mL} \cdot \text{cm}^{-3} \cdot \text{min}^{-1}$), and k_2

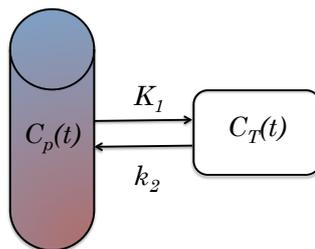


Figure 3.1: 1 tissue compartment model, where C_p and C_T are the tracer concentrations in blood plasma and tissue respectively, K_1 is the transfer rate constant from blood to tissue ($\text{mL} \cdot \text{cm}^{-3} \cdot \text{min}^{-1}$), and k_2 is the transfer rate constant from tissue to blood (min^{-1}).

(min^{-1}) which is the rate constant in the opposite direction. These parameters are assumed to remain constant over the PET acquisition and are estimated by fitting the model to the measured data.

Equation 3.4 is the differential equation derived from the 1T2P model. This can be solved for C_T by applying the inverse Laplace transform to yield equation 3.5. More complex kinetic models, with additional tissue compartments and associated rate constants, may be set up and solved in a similar way.

$$\frac{dC_T(t)}{dt} = C_p(t)K_1 - C_T(t)k_2 \quad (3.4)$$

$$C_T(t) = \int_0^t C_p(t)K_1 - C_T(t)k_2 = C_p(t)K_1 \otimes e^{-k_2 t} \quad (3.5)$$

In general, PK models are formulated such that the concentration of tracer in tissue (C_T) at a given time (t) is the result of a tracer input function and a tissue response function, (R), equation 3.6. The parameters of the model can be estimated by solving the differential equations for exchange between compartments. This is generally performed by applying non-linear least squares fit to minimise the difference between the fitted model and the data.

$$C_T(t) = C_p(t) \otimes R(t) \quad (3.6)$$

3.2.2 Pharmacokinetic modelling using arterial blood sampling

Compartmental PK modelling. The first successful human trials to show that a PET tracer can be used to image amyloid plaques were carried out by Klunk et al. (2004) using ^{11}C -PiB. Following this semi-quantitative first in human study, Price et al. (2005) carried out a full kinetic analysis of 15 subjects (5 healthy controls, 5 mild cognitive impairment and 5 clinically diagnosed AD) to assess the suitability of various quantitative measures.

Price et al. (2005) used the DVR as the outcome measure of interest to quantify the amyloid burden, see equation (3.2). Here, the non-displaceable volume of distribution (V_{ND}) was determined by fitting the models to the reference regions, in which it is considered that there is no specific binding and is therefore non-displaceable.

Three compartmental models of different complexities were used by Price et al. (2005), the most complex being the three-tissue six-parameters (3T6P) model which was based on the work by Koeppe et al. (1994), and can be described as in figure 3.2a. Here, the first of the three tissue compartments represents free tracer within the tissue which exchanges with the blood plasma (C_F). The tracer can then either be taken into specific binding sites (C_S) in the presence of amyloid, or non specific binding sites (C_{NS}). Due to the large number of parameters and the difficulty in distinguishing between the specific and non-specific binding compartments this model is unstable and difficult to fit.

The 3T6P model can be simplified by assuming that the free and non-specific binding compartments reach rapid equilibrium and are thus indistinguishable, which reduces the model to two tissues and 4 parameters (2T4P), figure 3.2b. This makes the model easier to fit, and was used by Price et al. (2005) as the primary model configuration. The differential equations can be set up as follows:

$$\begin{aligned}\frac{dC_{F+NS}(t)}{dt} &= C_p(t)K_1 - C_{F+NS}(t)k_2 + C_S(t)k_4 - C_{F+NS}(t)k_3 \\ &= C_p(t)K_1 - C_{F+NS}(t)(k_2 + k_3) + C_S(t)k_4\end{aligned}\quad (3.7)$$

$$\frac{dC_S(t)}{dt} = C_{F+NS}(t)k_3 - C_S(t)k_4 \quad (3.8)$$

$$C_T(t) = C_{F+NS}(t) + C_S(t) \quad (3.9)$$

The volume of distribution is defined at equilibrium where $\frac{dC_{F+NS}}{dt} = \frac{dC_S}{dt} = 0$. C_{F+NS} and C_S can then be found by combining and rearranging equation (3.7) and equation (3.8) and then substituted into equation (3.9) to give C_T . Using equation (3.1), this result can then be divided by C_P to give the volume of distribution for the 2T4P model as given in equation (3.10). Compartmental modelling in this way allows the volume of distribution to be estimated through the fitted rate constants, even if the tracer never reaches true equilibrium.

$$V_T(t) = \frac{K_1}{k_2} \left(1 + \frac{k_3}{k_4}\right) \quad (3.10)$$

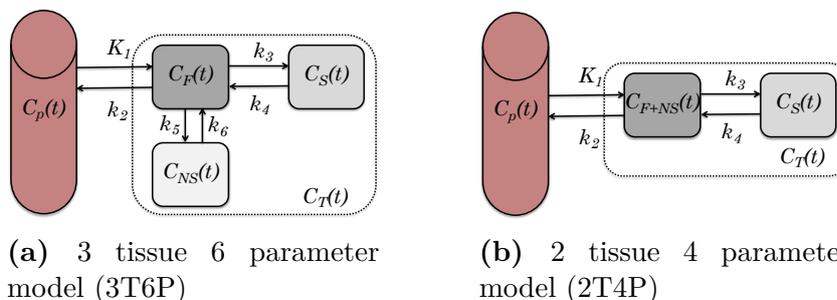


Figure 3.2: Schematics showing compartmental models used in Price et al. (2005)

Graphical PK models. As an alternative to compartmental models, which make assumptions about the behaviour of the tracer, data driven techniques may be used. Such techniques generalise the function in equation 3.6 to an arbitrary number of compartments. Graphical analysis techniques are one such example where tissue concentration and arterial plasma curves are combined and transformed into a single plot which approaches linearity under certain

conditions. This allows simple fitting of the linear data where the gradient is used to estimate the volume of distribution. There are many formulations of graphical analysis, however for the case of reversible binding, such as for amyloid tracers, the most commonly used are the Logan (Logan et al., 1990) and Ichise (Ichise et al., 2002) techniques. While graphical analysis methods are independent of any particular model structure, the slope can be interpreted in terms of a combination of model parameters for some model structure.

The Logan plot takes the general form of equation (3.11) (Logan et al., 1990). This form can be interpreted using the 2T4P model such that the volume of distribution is defined as in equation (3.10). To derive the volume of distribution the linear portion of the curve must be fitted, where linearity occurs at some time $t > t'$ where $t = 0$ at tracer injection. In general t' is defined as the point at which the tissue compartments and blood plasma reach equilibrium such that the intercept (int) becomes constant. For many tracers linearity is often reached before equilibrium allowing the length of the dynamic acquisition to be reduced. For the Ichise technique, the equations for Logan analysis are solved using multi-linear regression.

$$\frac{\int_0^T C_T(t)dt}{C_T(T)} = V_T \frac{\int_0^T C_p(t)dt}{C_T(T)} + int \quad (3.11)$$

Comparison of quantification methods using arterial sampling. In Price et al. (2005), the 3T6P, 2T4P, 1T2P, Logan and Ichise analysis methods were applied to 90 minute dynamic PET imaging, all using metabolite corrected arterial samples to provide the input function. Analysis showed that the most complex model, 3T6P, often failed, giving negative parameter values as the data did not support a model with this number of parameters. The 2T4k model yielded the lowest Akaike information criteria (AIC) and increased stability under test-retest conditions, and was taken to be the gold standard for comparison with other techniques. The authors concluded that the two-tissue Logan analysis was a viable alternative as it minimised the fitting error effectively, and was robust, as demonstrated by the low test-retest variability.

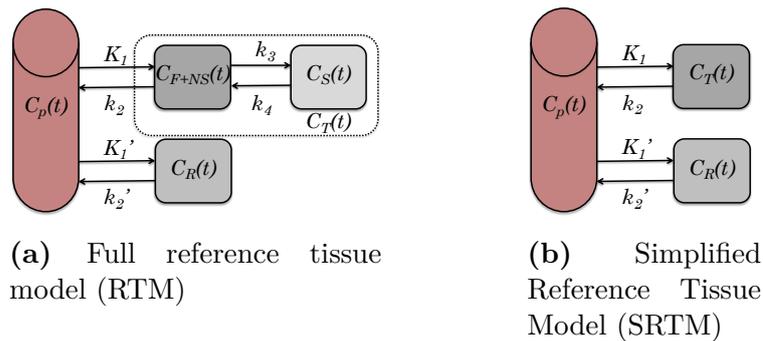


Figure 3.3: Schematics showing the compartments and exchange constants for reference tissue models

More recently, in a similar study for [^{18}F]-florbetapir data, Golla et al. (2018) compared performance of the 3T6P, 2T4P, 1T2P models on data acquired from eight participants with probable AD. Again the 2T4P models had the lowest AIC and was considered to be stable under test-retest conditions provided 90 minutes of data was used. This was consistent with the findings of Price et al. (2005) for ^{11}C -PiB data, supporting the use of similar analysis techniques for the two tracers.

3.2.3 Simplified PK models (without arterial sampling)

The methodology of Price et al. (2005) required arterial sampling to provide the plasma concentration for the analysis. However, these data are often not available, as this is a highly invasive procedure which not only introduces additional risk to the patient, but also requires specialist personnel and equipment for analysis. Therefore, simpler techniques for kinetic analysis were explored by Lopresti et al. (2005).

There are two ways to derive the input information from the images alone; either a region of interest (ROI) can be drawn directly around the carotid arteries in the image or a reference tissue region, in which specific binding is negligible, is used as a surrogate for the input function. Due to the small diameter of the carotid arteries, the arterial ROI method is highly susceptible to PVE caused by blurring of the signal due to the limited resolution of PET data, and generally requires manual segmentation of the vessels. The PET

signal must then be corrected for metabolites as there is no way to distinguish metabolite from parent tracer signal, and metabolites don't bind to the tracer target. Lopresti et al. (2005) used a population based metabolite correction on the image derived carotid input function with Logan analysis. When compared against the gold standard using arterial samples the method was found to be accurate but not reproducible due to difficulties in defining the artery region. A methodology to overcome these problems which combines high resolution MRI to define the arteries and the simultaneous estimation method (Wong et al., 2001) has been proposed (Sari et al., 2018), however this has not been validated for amyloid tracers.

The reference tissue model (RTM). The use of a reference tissue within the analysis of amyloid data was suggested by Price et al. (2005) as it was used to represent non-displaceable binding for DVR calculation. The cerebellum was selected as there was no significant difference in uptake in this region between healthy controls and AD subjects (Price et al., 2005; Klunk et al., 2004). Furthermore, *post mortem* histological studies show very low levels of fibrillar amyloid in this region (Joachim et al., 1989). When using a reference region it is assumed that the non-displaceable volume of distribution (V_{ND}), i.e. the contribution of free and non-specifically bound tracer, is uniform across the brain, equation (3.12). Here, a prime is used to denote the rate constants in the reference tissue.

$$\frac{K'_1}{k'_2} = \frac{K_1}{k_2} = V_{ND} \quad (3.12)$$

In the reference tissue model (RTM), the reference region is used in lieu of an arterial input function and the model is set up as in figure 3.3a, where the target tissue contains 2 compartments, and the reference tissue is considered to be a single tissue compartment containing free and non-specifically bound tracer only (Blomqvist et al., 1989; Lammertsma et al., 1996). Since the reference region is a 1TC model, the differential equation can simply be set up as in equation (3.13), which can be re-arranged for $C_p(t)$ negating the need to

measure an arterial input function.

$$\frac{dC_R(t)}{dt} = K'_1 C_p(t) - k'_2 C_R(t) \quad (3.13)$$

In the RTM non-displaceable binding potential (BP_{ND}) is the parameter of interest to determine amyloid burden, which was previously defined in equation (3.3). The volume of distribution for a 2TC model is given in equation (3.10), and can be combined with the non-displaceable volume of distribution in equation (3.12) to give equation (3.14).

$$BP_{ND} = \frac{V_T}{V_{ND}} - 1 = \frac{\frac{K_1}{k_2} (1 + \frac{k_3}{k_4})}{\frac{K'_1}{k'_2}} - 1 = \frac{k_3}{k_4} \quad (3.14)$$

The differential equations for the target tissue can be set up as for the 2T4P model, see equations (3.7-3.9). By combining these equations with (3.13) it is possible to derive a relationship between C_T and C_R that does not include C_{F+NS} , C_S or C_P but which contains 6 parameters.

To reduce the total number of free parameters the relative tracer delivery parameter R_1 is defined as in equation (3.15) as K_1 and K'_1 only appear as a ratio. Furthermore, k'_2 can be redefined by rearranging equation (3.12) and substituting in equation (3.15), as in equation (3.16). This leaves 4 parameters to fit in the model: R_1 , k_2 , k_3 , and BP_{ND} .

$$R_1 = K_1/K'_1 \quad (3.15)$$

$$k'_2 = \frac{k_2}{R_1} \quad (3.16)$$

The simplified reference tissue model (SRTM). The RTM can be further simplified by assuming that the free and nonspecific binding compartment quickly reaches equilibrium with the specific binding compartment, making them indistinguishable. This leads to the simplified reference tissue model (SRTM), as described by figure 3.3b (Lammertsma and Hume, 1996). In this

case, equation (3.13) remains unchanged, but the equations with the target tissue can be reduced to equation (3.17), where k_{2a} is the apparent rate constant for transfer from the specific compartment to plasma.

$$\frac{dC_T(t)}{dt} = K_1 C_P(t) - k_{2a} C_T(t) \quad (3.17)$$

The total tracer volume of distribution using the SRTM should be equal to that found using the RTM, as shown by equation (3.18). This can be rearranged for k_{2a} , where K_1 cancels, and substituted into equation (3.17) along with equations (3.13) and (3.16) to give equation (3.19).

$$V_T = \frac{K_1}{k_{2a}} = \frac{K_1}{k_2} (1 + BP_{ND}) \quad (3.18)$$

$$\frac{dC_T(t)}{dt} = R_1 \frac{dC_R(t)}{dt} + k_2 C_R(t) - \frac{k_2}{(1 + BP_{ND})} C_T(t) \quad (3.19)$$

As previously the inverse Laplace transform can then be used to derive the operational equation for the SRTM, given by equation (3.20). The model now contains only three parameters: R_1 , k_2 , and BP_{ND} . This model is more robust and converges faster than the RTM making it better suited for fitting noisy PET data, as demonstrated by its use in the comparison paper of Lopresti et al. (2005).

$$C_T(t) = R_1 C_R(t) + \left(k_2 - R_1 \frac{k_2}{1 + BP_{ND}} \right) C_R(t) \otimes e^{-\frac{k_2}{1 + BP_{ND}} t} \quad (3.20)$$

Reference tissue input for graphical analysis. The reference region can also be incorporated into graphical analysis (Logan et al., 1996) to derive the binding potential. This technique can be implemented using a population average value of backflux rate constant from the reference region to vascular space (k'_2), which must be derived from a compartmental model with arterial sampling. Lopresti et al. (2005) used the average k'_2 derived from the 2T4P

model to fix this parameter.

Comparison of simplified quantification methods. In the comparative study of simplified methods carried out by Lopresti et al. (2005) Logan analysis with an arterial input function was used as the gold standard (Price et al., 2005), which was compared with Logan analysis using carotid ROI input, reference tissue Logan analysis and the SRTM. It was found that the carotid ROI method gave the closest result to the gold standard method, although the difficulty in defining the region meant that it was not very reproducible. The reference tissue Logan approach resulted in negatively biased DVR estimates, particularly for patients with a high amyloid burden. The SRTM was found to be more reproducible than using the carotid ROI input method and less biased than the reference Logan technique, however the authors concluded that the suitable choice of technique depends on the kind of study to be performed.

Following on from this study, various implementations of the SRTM have been applied to ^{11}C -PiB data. Yaqub et al. (2008) compared several versions, including the SRTM2 (Wu and Carson, 2002) in which the value of k'_2 is fixed using a subject-specific average value, since it should be the same across all non-reference region voxels. This is performed by re-writing the operational equation (3.20) in terms of k'_2 then running the fit in two stages; first the equation is fitted for each brain voxel and a global value of k'_2 is found by averaging across valid voxels, then the model is re-fitted using the fixed k'_2 to derive R_1 and BP_{ND} . Yaqub et al. (2008) compared the SRTM, SRTM2, reference tissue Logan and Ichise graphical analysis and found that the SRTM2 using fixed k'_2 was the most accurate and consequently this methodology was used for PiB data by Zhou et al. (2007). For this implementation Zhou et al. (2007) used simultaneous fitting of all ROIs to a single value of k'_2 rather than a 2-step process.

In Yaqub et al. (2008), the implementation of the SRTM used the basis function pursuit method proposed by Gunn et al. (1997) to solve the non-linear term, in equation(3.20). For this method the operational equation is re-written

in terms of R_1 , ϕ and θ , as in equation (3.21). Here a set of basis functions are generated using a physiologically plausible range for k_2 and BP_{ND} to pre-calculate the non-linear term. The optimisation is performed by finding the best basis function based on the sum of squared differences (SSD) as described by algorithm (3.1). This allows the problem to be linearised, which facilitates faster and more robust optimisation compared to non-linear fitting.

$$C_T(t) = R_1 C_R(t) + \phi C_R(t) \otimes e^{-\theta t} \quad \text{where} \quad (3.21)$$

$$\phi = k_2 - R_1 k_2 / (1 + BP_{ND}), \quad \theta = k_2 / (1 + BP_{ND})$$

Following the conclusions of Yaqub et al. (2008), the SRTM proposed by Zhou et al. (2007) was used for the first human ^{18}F -Florbetapir studies (Wong et al., 2010), and many subsequent studies (Hsiao et al., 2012; Chen et al., 2015; Sojkova et al., 2015b). Validation of the SRTM against the gold standard 2T4P model for ^{18}F -florbetapir data was recently carried out by Golla et al. (2018). Evaluation on 8 controls and 8 AD subjects showed that the SRTM provided a reliable estimate of BP_{ND} for ^{18}F -florbetapir. This methodology has also successfully been applied to ‘coffee-break’ data, where the patient takes a break from scanning in the middle of the acquisition and the missing data is interpolated (Heeman et al., 2019; Bullich et al., 2018).

3.3 Quantification of static PET data

For the quantification techniques described in section 3.2, dynamic PET data must be acquired from injection of the tracer for between 60 and 90 minutes, depending on the kinetics of the tracer. Once this data has been acquired and reconstructed into time frames modelling software must be applied. This requires scanner availability and patient compliance to facilitate the long acquisition time, plus the availability of kinetic modelling software and expertise to apply it. In general, radiopharmaceutical manufacturers aim to develop tracers in which a static acquisition and simple analysis is enough to quantify

Algorithm 3.1 SRTM with basis functions

```

1: For each subject:
2:  $C_R(t)$ : measured reference region concentration
3:  $n_{basis}$ : number of basis functions
4:  $B_i(t) = C_R \otimes e^{-\theta t}$ : basis functions ( $i = 1 : n_{basis}$ )
5:  $w = \frac{\text{frame duration}}{e^{-\lambda t}}$ : weights
6:  $W = \text{diag}(\sqrt{w})$ : weighting matrix

7: for  $i = 1 : n_{basis}$  do
8:    $A = [C_R \ B_i]$ 
9:    $[Q, R] = \text{qr}(W \cdot A)$ 
10:   $M(2i - 1 : 2i, :) = R \setminus Q^T$ 
11: end for

12: For each region/voxel:
13:  $C_T(t_j)$ : measured region/voxel concentration
14:  $n_f$ : number of frames ( $j = 1 : n_f$ )

15: for  $i = 1 : n_{basis}$  do
16:    $\begin{bmatrix} R_1 \\ \phi \end{bmatrix} = M(2i - 1 : 2i, :) \cdot W \cdot C_T(t_j)$ 
17:    $\hat{C}_T(t_j) = R_1 C_R(t_j) + \phi B_i(t_j)$ 
18:    $SSD_i = \sum_{j=1}^{n_f} W \cdot (C_T(t_j) - \hat{C}_T(t_j))^2$ 
19: end for

20:  $i_{min} = \text{find}(\text{min}(SSD))$ 
21:  $\begin{bmatrix} R_1 \\ \phi \end{bmatrix} = M(2i_{min} - 1 : 2i_{min}, :) \cdot W \cdot C_T(t_j)$ 
22:  $k_2 = \phi + R_1 \cdot \theta$ 
23:  $BP_{ND} = \frac{k_2}{\theta} - 1$ 

```

the target density.

3.3.1 Standardised uptake value ratio (SUVR)

SUVR is a semi-quantitative measure of uptake, which requires approximately 10 minutes of static PET data. The name SUVR is derived from the standardised uptake value (SUV) commonly used in FDG PET quantification, where the measured activity concentration is normalised to the concentration if the tracer were uniformly distributed in the body. However since the SUVR is a ratio, the unit which converts tracer concentrations into SUV's cancel out such that it is sometimes referred to as relative uptake instead.

SUVR is intended to approximate estimates of DVR derived from the full dynamic PET data, by calculating the ratio of C_T to C_R at steady-state. The definition of DVR is given in equation (3.2), where C_T and C_{ND} are measured at equilibrium such that $\frac{dC_t}{dt} = \frac{dC_{ND}}{dt} = 0$. Assuming the same set up as the SRTM with 2TC in the target tissue and 1 in the reference region, equations (3.13) and (3.17) can be re-arranged as in equations (3.22,3.23) such that their ratio is equal to DVR, as given by equation (3.24).

$$C_R(t) = \frac{K_1' C_p(t)}{k_2'} = V_{ND} C_p(t) \quad (3.22)$$

$$C_T(t) = \frac{K_1 C_p(t)}{k_2 a} = V_T C_p(t) \quad (3.23)$$

$$SUVR(t) = \frac{C_T(t)}{C_R(t)} = \frac{V_T}{V_R} = DVR \quad (3.24)$$

3.3.2 Limitations of SUVR Quantification

By definition equation (3.24) is only valid at equilibrium, however many tracers will not reach equilibrium *in vivo*. Instead, often SUVR will be calculated when the tracer has reached a steady-state where $\frac{C_T}{C_R} = const$, also known as transient or pseudo equilibrium. This occurs approximately 50 minutes after injection for amyloid tracers (Cselényi and Farde, 2015). However, since the tracer concentration in the target tissue will always be higher than than in the reference tissue, due to higher binding and slower washout $C_T = \beta C_R$ where $\beta > 1$. This means that $\frac{dC_T}{dt} = \beta \frac{dC_R}{dt}$. As a result SUVR will overestimate DVR, particularly at higher DVR values. This phenomenon has been discussed at length in the literature (Carson et al., 1993; Slifstein, 2008; van Berckel et al., 2013) and was demonstrated for [^{18}F]-florbetapir data through simulation (Ottoy et al., 2017a) and by comparison to the 2T4P model (Ottoy et al., 2017b). Based on the SRTM, Alves et al. (2017) derived an expression to relate DVR to SUVR which was dependent on k_2' , $\frac{dC_R}{dt}$, k_2 , and $\frac{dC_T}{dt}$. This demonstrates that the error in SUVR is determined by both tracer binding and blood

flow.

In some cases the fact that SUVR can overestimate the true DVR at high DVR values could be considered advantageous. Lopresti et al. (2005) found that SUVR had a higher effect size when discriminating between AD subjects and control groups compared to modelling approaches using dynamic data. However this is an issue for longitudinal studies where changes in blood flow affect the delivery of the tracer to tissue and consequently alter the tracer concentration in the tissue when a static image is acquired. This has been highlighted in longitudinal studies, where changes in blood flow have caused spurious changes in SUVR values which do not reflect imaging target abundance (van Berckel et al., 2013; Cselényi and Farde, 2015).

There are a large number of factors which can affect CBF in healthy controls (Clement et al., 2017). Parkes et al. (2004) showed that blood flow varies both over the course of a week (67-80 ml/min/100g in grey matter), and even through a single day (64-84 ml/min/100g in grey matter). Parkes et al. (2004) also demonstrated that CBF changes with normal ageing, reducing by approximately 0.45% per year. Furthermore, a significant reduction in CBF is seen in patients with AD, where groups differences are in the order of 20% (Steketee et al., 2016).

Cselényi and Farde (2015) showed how CBF changes influence SUVR, where they demonstrated a significantly higher blood flow component in the AD group compared to healthy controls. Cselényi and Farde (2015) went on to show that for each 0.1 unit decrease in R_1 , there is a 0.04 unit increase in SUVR. van Berckel et al. (2013) found significant longitudinal changes in R_1 in patients with AD, and through simulation using parameters derived from clinical studies suggested that longitudinal changes measured using SUVR may in fact be due to changes in blood flow, rather than amyloid deposition. This bias from CBF changes is related to the lack of tracer equilibrium, consequently these limitations apply to any such tracer, including amyloid tracers (van Berckel et al., 2013) and tau (Hahn et al., 2017).

The estimates derived by kinetic modelling are not biased by blood flow, as the full dynamic curve contains blood flow information, which is parameterised within the model. Furthermore there is not requirement for the tracer to reach equilibrium as amyloid burden estimates can be derived from the fitted rate constants. However, the long scan duration is prohibitive for routine clinical use, where time is limited, and data integrity is risked by the increased chance of subject motion, therefore a choice must be made between accuracy and feasibility in a clinical setting.

3.4 Selection of a Gold Standard for Amyloid Quantification

Both Price et al. (2005) and Golla et al. (2018) found that the 2T4P model with arterial blood sampling to derive the AIF was the best model for ^{11}C -PiB and [^{18}F]-florbetapir quantification respectively, as it accurately reflected the kinetics of the tracer *in vivo* whilst being stable enough to generate reproducible results. However, Golla et al. (2018) also found that difficulties in measuring the tracer parent fraction from the blood samples affected the estimation of V_T and thus recommended using a reference tissue input if possible.

When accurate arterial blood sampling data are not available, Yaqub et al. (2008) showed that, for ^{11}C -PiB data, the SRTM2 was the most accurate and stable model, and as such it has been used in several ^{11}C -PiB and [^{18}F]-florbetapir studies (Zhou et al., 2007; Wong et al., 2010; Hsiao et al., 2012; Chen et al., 2015; Sojkova et al., 2015a). Golla et al. (2018) showed that the SRTM (not SRTM2) using the basis function approach is stable and accurate enough for use on [^{18}F]-florbetapir data¹ and was selected as the gold standard here.

Limitations of the SRTM as a gold standard. The theoretical assumptions of the SRTM and their effect on quantification were explored by Salinas

¹Details of SRTM implementations tested are not included in Golla et al. (2018) but were given in a personal communication with the lead author. The authors found that the SRTM2 did not give better results than the SRTM, so the SRTM is recommended

et al. (2014). The four key assumptions which were tested were 1) there is no displaceable component in the reference region, 2) the target and reference region can be described by one tissue compartment kinetics, 3) the blood volume contribution to the tissues is negligible, 4) the V_{ND} is the same in the target region as in the reference region.

Assumption 1) can be justified by the findings in Ottoy et al. (2017b) and Golla et al. (2018) where no significant difference in cerebellar grey matter binding was found between healthy controls and subjects with Alzheimer’s disease. Salinas et al. (2014) concluded from their simulations that assumption 2) could only produce a small bias provided the target and reference region contain the same number of compartments. Golla et al. (2018) (and Price et al. (2005) for ^{11}C -PiB) showed that both the target and reference tissue could be described by 2 compartment kinetics, thus validating the assumption. It is worth noting that for the SRTM2 formulation higher biases are anticipated than for the SRTM when this assumption is violated (Wu and Carson, 2002).

Currently there are no data to determine whether assumptions 3) and 4) are satisfied. Salinas et al. (2014) found that the violation of assumption 3), where it is assumed that the blood volume is negligible, will produce a bias that is more or less constant across a wide range of BP_{ND} values and therefore will have little impact unless blood volume constitutes a significant portion of the total tissue signal. Golla et al. (2018) incorporate the blood volume (V_B) into their 2T4P model however the authors do not discuss the proportion of the signal which originates from blood so it is difficult to determine the magnitude of the bias. Equally the non-displaceable volume of distribution is estimated by Golla et al. (2018) but not discussed, presumably due to the high correlation between the SRTM and the 2T4P model.

One other limitation of the SRTM is its instability when fitting regions which have a low BP_{ND} . This leads to k_2 being numerically unidentifiable which can propagate to the estimates of BP_{ND} and R_1 , however this can be minimised by careful selection of the range of basis functions generated (Gunn

et al., 1997). This effect is shown in the study by Golla et al. (2018) where the correlation between the 2T4P model is higher for the AD group than the healthy controls (who generally had a very low amyloid burden). However, the authors considered the correlation for both groups to be good ($r^2 = 0.93$ for AD, $r^2 = 0.73$ for control, and $r^2 = 0.83$ overall).

3.5 Conclusion

A wide variety of quantification approaches have been applied to amyloid PET data, and these vary greatly in complexity. In general it has been shown that the best model for the data is the 2T4P model, as this can accurately describe the tracer kinetics *in vivo*.

However, since the 2T4P model requires arterial sampling to generate the Arterial Input Function (AIF) which is invasive and can be unreliable, it is not suitable for routine use. The SRTM has been shown to offer a good compromise between quantitative accuracy and stability compared to the 2T4P model and is therefore taken as the non-invasive gold standard. However this still requires a long dynamic acquisition lasting approximately 60 minutes for [^{18}F]-florbetapir, which is challenging for many patients and reduces scanner throughput. Consequently SUVR is the most common metric used clinically, as this only requires a static 10 minutes acquisition and is simple to calculate. However it has been shown that the estimates of amyloid burden derived from SUVR contain significant bias which is related to blood flow and target density. As a result there is a pressing clinical need for a method which can produce accurate estimates of amyloid burden within a clinically feasible time frame which could potentially be achieved if blood flow can be estimated during a shortened PET acquisition.

Chapter 4

Blood flow modelling and quantitative ASL

This chapter covers the background and theory behind blood flow modelling, including an extension to the Renkin-Crone model, which can be applied to extraction-limited tracers to estimate blood flow. Arterial spin labelling (ASL) MRI is also introduced as a non-invasive method of measuring blood flow. A brief description of the application of blood flow models to ASL data is also covered, along with an overview of the considerations and challenges associated with ASL based quantification of cerebral blood flow.

4.1 Introduction

Perfusion can broadly be defined as the steady state delivery of nutrients to an organ. Perfusion can be described using several parameters, the most common being specific blood flow and blood volume. When considering the brain, specific blood flow is referred to as cerebral blood flow (CBF) and is a rate defined as the quantity of blood reaching a unit tissue volume per unit time. It is normally measured in ml/min/100g assuming a mean brain density of 1g/ml. Conversely, blood volume (or cerebral blood volume in the brain) is the proportion of blood in a mixed blood/tissue volume, and is normally expressed as a percentage of the volume or in ml of blood per 100g of tissue. This chapter looks at how CBF can be quantified *in vivo* using physiological

models.

4.2 Blood flow modelling

Blood flow modelling to derive CBF from imaging data was formalised for the analysis of H_2O^{15} -PET (radio-water) data, where the tracer is injected intravenously and behaves like blood water. H_2O^{15} -PET is generally considered to be the gold standard measurement of CBF *in vivo* (Fan et al., 2016).

4.2.1 Indicator dilution theory

The PET models are based on indicator dilution theory (Meier and Zierler, 1954), which describes a system which has one input artery delivering blood, one vein removing the blood, and a network of capillaries in between which perfuse the tissue, as shown in figure 4.1. In this system, the different paths the blood can take through the capillary bed have different transit times, defined as the time taken to get from A to V. However, the model assumes that these transit times are constant, this is known as the stationarity of flow. This model also assumes that the indicator (or tracer) introduced into the system, behaves in the same way as the fluid already in the system and that the tracer is not recirculated.

The Fick principle states that when the system is in a steady state, i.e. the tracer does not accumulate within the capillary, then the flux in is equal to the flux out. In this case the flux in is the arterial input (ϕ_a), and the flux out is the combination of the flux into the tissue (ϕ_T), and the tracer removed through the venous outflow (ϕ_v).

$$\phi_T = \phi_a - \phi_v = F(C_A(t) - C_V(t)) = \frac{dC_T(t)}{dt} \quad (4.1)$$

By definition $\phi = F \times C$, where F is the blood flow and C is the tracer concentration. A model for changing concentration of tracer in tissue with time ($\frac{dC_T}{dt}$) can be derived by combining indicator dilution theory with the Fick principle, as expressed in equation (4.1).

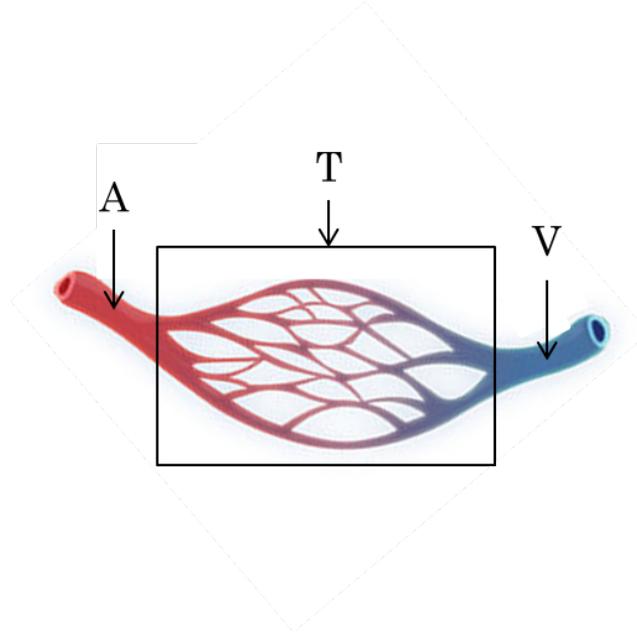


Figure 4.1: schematic of capillary system

$$\lambda = \frac{C_T}{C_v} \quad (4.2)$$

It may be assumed that the H_2O^{15} can freely diffuse into the tissue, and that it becomes well mixed within the brain. The latter assumption means that the tissue concentration C_T is equal to the outflowing venous concentration C_v if the difference in water density between these two regions is corrected for, equation (4.2). The factor λ is known as the blood-brain partition coefficient, and generally a brain averaged value is used despite it being tissue dependent, as it is non-trivial to measure (Alsop et al., 2015). This can be combined with equation (4.1) to give the Kety-Schmidt model in equation (4.3).

$$\frac{dC_T(t)}{dt} = FC_A(t) - \frac{F}{\lambda}C_T(t) \quad (4.3)$$

Equation (4.3) then takes the form of the 1TC model described in section 3.2, where $K_1 = F$ and $k_2 = \frac{F}{\lambda}$. CBF can then be estimated either through non-linear optimisation, through a basis function approach, through graphical analysis, or through autoradiography, where static PET images are combined

with dynamic blood sampling (Fan et al., 2016).

4.2.2 The Renkin Crone model

The Renkin-Crone capillary model, a development of the indicator dilution theory, is commonly used in PET and MRI to describe the exchange of substances between blood and tissue where the tracer is not assumed to be freely diffusible. It was developed independently to model extraction of potassium from the blood (Renkin, 1959) and to measure vessel permeability to inulin and sucrose (Crone, 1963).

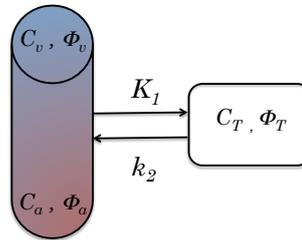


Figure 4.2: Illustration of the 1 tissue compartment Renkin-Crone model

The Renkin-Crone model is set up in the same way as for indicator dilution theory, with one tissue compartment, and two blood compartments (venous and arterial), but here two rate constants (K_1 and k_2) are also included, see figure 4.2. The rate constants reflect the transfer of tracer between the blood and tissue compartment in ml/min/100g.

$$E_{net} = \frac{C_A(t) - C_V(t)}{C_A(t)} \quad (4.4)$$

The net extraction fraction, E_{net} which is the fraction of the tracer which is extracted, can be defined as in equation (4.4). However, E_{net} is not necessarily a useful parameter, as C_v will contain both tracer which has passed back from tissue to blood, as well as tracer which was not taken up in tissue. To get the unidirectional extraction fraction E_u , which is the extraction of tracer from blood to tissue, consider the case at $t = 0$. Here, $C_T = 0$, therefore there is no flow of the tracer from tissue to blood, and $E_{net} = E_u$. Under these conditions, the 1 tissue compartment model (equation (3.4)), the Kety-Schmidt model

(equation (4.1)), and equation (4.4) can be combined to give a relationship between the rate constant K_1 and the cerebral blood flow, equation (4.5).

$$\frac{dC_T(t=0)}{dt} \Big|_{C_T=0} = K_1 C_A(t) = F(C_A(t) - C_V(t)) = F E_u C_A(t) \implies K_1 = F E_u \quad (4.5)$$

The Renkin-Crone model seeks to express the extraction fraction in terms of other physiological parameters. This is achieved by modelling the capillary as a rigid tube, and expressing tracer uptake into tissue as a diffusive process described by Fick's first law of diffusion. Under these assumptions, the Renkin-Crone model is given as equation (4.6), where PS is the vessel permeability surface area product.

$$E_u = 1 - e^{-\frac{PS}{F}} \implies K_1 = F E_u = F(1 - e^{-\frac{PS}{F}}) \quad (4.6)$$

The Renkin-Crone model demonstrates that the relationship between K_1 and flow is non-linear, and depends on the permeability of the vessel to the tracer, and on the surface area over which it can exchange. This model implies that if PS is low compared to F , then K_1 becomes independent of flow, as the delivery of the tracer is limited by extraction rather than flow. Conversely, as $\frac{PS}{F}$ exceeds 3, the exponential term tends to zero, and K_1 is approximately proportional to F , reducing back to the one compartment model in equation (4.3).

4.3 Magnetic Resonance Imaging based measurement of Blood Flow

There are several methods for measuring CBF using MRI, a notable example being dynamic susceptibility contrast (DSC) MRI. Here a contrast agent is injected intravenously to shorten the T2 and T2* of the blood, allowing the tracer concentration to be measured over time from dynamic T2 or T2* weighted imaging, such that the CBF can be estimated across the brain. However, techniques which don't require an exogenous contrast agent are beneficial

for PET-MR imaging, as the injection of the contrast agent, in addition to the PET tracer injection, can make the clinical workflow more complex.

Alternatively, phase contrast (PC) MRI can be used to measure the flow flux in the arteries which feed the brain to measure whole brain CBF completely non-invasively. This has been used to scale H_2O^{15} -PET data to remove the need for arterial sampling during simultaneous PET-MR scans (Ssali et al., 2018). However, PC-MRI on its own only gives whole brain CBF values.

ASL is currently the only technique which can measure CBF on a voxel-wise basis non-invasively.

4.3.1 Principles of Arterial Spin Labelling

ASL-MRI uses magnetically tagged blood as an endogenous contrast agent to measure cerebral blood flow non-invasively. This requires the acquisition of control images, with no tagging of the blood, and of label images in which the arterial blood has been magnetically tagged, see figure 4.3.

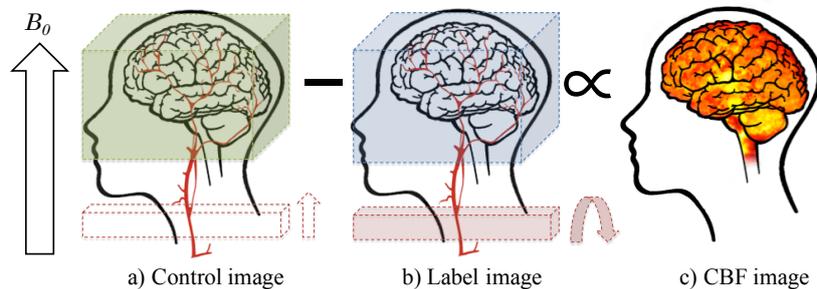


Figure 4.3: Illustration of ASL acquisition for cerebral blood flow mapping, showing a) the control image where the arterial blood net magnetisation (red arrow) remains aligned with the external B_0 field, then b) the label image where the arterial blood is tagged using an RF pulse, and finally c) showing that the difference between the control and label images is proportional to perfusion.

For the control image, with no tagging of the arterial blood, the spins are aligned to the scanner’s static magnetic field (B_0), therefore there is maximum signal available for the image acquisition. The label image is then acquired by first applying a RF pulse to a region in the neck, to locally alter the orientation of the spins, giving them a magnetic ‘tag’. The labelled arterial blood which passes through this region then flows into the brain and exchanges with the

tissue, and the label image is acquired.

In the label image there will be a loss of signal caused by the tagged blood, as the orientation of its spins is different to surrounding static tissue. The signal difference between the control and label image is proportional to the CBF, as the higher the blood flow in a region, the more tagged blood is present, and the greater the signal loss.

Since blood constitutes about 2-4% of the volume in the brain, and about 1% of the blood water is replaced each second, the signal to noise ratio (SNR) in ASL is very low (Golay, 2013). Consequently, multiple pairs of control and labelled images are acquired to average over several cardiac outputs and to increase SNR. A variety of labelling and readout schemes are available for ASL, as well as different quantification models, with different levels of accuracy, complexity, and acquisition time. The optimal selection of the acquisition and processing protocol depends on scanner hardware and the available acquisition time, and recommendations for clinical implementation have been published in a consensus paper (Alsop et al., 2015).

4.3.2 Models for Arterial Spin Labelling

Kinetic modelling is required to convert the perfusion maps created from the difference between the control-label pairs into an absolute measure of CBF. The model can be formulated to account for the delivery of the labelled spins to tissue, signal decay and washout, and the acquisition process. The following derivation is adapted from Golay (2013) and Chappell et al. (2018).

$$C_T(t) = \int_0^t \frac{dC_T(t)}{dt} = \int_0^t FC_A(t) - \frac{F}{\lambda} C_T(t) \quad (4.7)$$

Equation (4.3) models the change in tracer (or label) concentration over time. This can be integrated over time as in equation(4.7) to solve for C_T . Just as for the 1TC model described in chapter 3.2, this can be reformulated using the Laplace transform to give equation (4.8).

$$C_T(t) = FC_A(t) \otimes e^{-\frac{F}{\lambda}t} \quad (4.8)$$

In ASL, C_T is not directly measured but instead it is the difference in signal between the control ($SI_{control}$) and label (SI_{label}) images, as described in equation(4.9). To equate ΔM directly to flow, the system can be modelled as in equation (4.10), where C_a is the arterial input function, and $R(t)$ is the tissue response function.

$$\Delta M = SI_{control} - SI_{label} \propto C_T \quad (4.9)$$

For ASL, the arterial input function C_A can normally be modelled as a top-hat function due to the nature of the labelling, which is applied in close proximity to the tissue to minimise dispersion. The total signal is twice the arterial magnetisation of the blood ($M_{0,a}$) as the signal has been inverted, making the difference between the control and label magnetisation equal to $2\alpha M_{0,a}$. Here α is the labelling efficiency ($\alpha \leq 1$), which is the proportion of arterial blood which is tagged. This signal decays over time with the T1 of blood and is modelled differently for pulsed ASL (PASL) and pseudo-continuous ASL (pCASL), see section 4.3.3.

$$\Delta M = FC_A(t) \otimes R(t - \tau)dt = FC_A(t) \otimes r(t - \tau)m(t - \tau)dt \quad (4.10)$$

The tissue response function, $R(t)$, can be subdivided into two terms, namely $r(t) = e^{-\frac{ft}{\lambda}}$ which accounts for the outflow of spins from the tissue as calculated in equation (4.8), and $m(t) = e^{-\frac{t}{T_1}}$ which models the T1 decay of the magnetisation of the spins within the tissue.

Equation (4.10) is the general kinetic model for ASL proposed by Buxton et al. (1998), however a simplified version of this model is more commonly applied, as recommended by Alsop et al. (2015). In this case the following three simplifying assumptions are made: the whole labelled bolus is delivered to the target tissue (delay time between labelling and acquisition must be greater than the time taken for the bolus to reach the tissue), there is no

outflow of labelled blood water, as the tissue water pool is much larger than the blood water pool (allowing $r(t)$ to be ignored), and the decay of the label signal is determined by the T1 of blood (negating the need to model signal decay separately in blood and in tissue, but errors can become significant if the label spends extended periods in the tissue prior to readout). Under these assumptions the model can be simplified to give equation (4.11) (for pCASL data), where the 6000 is simply for conversion into units of ml/min/100g, τ is the labelling duration and PLD is the post labelling delay, i.e. the time between blood labelling and image acquisition.

$$\text{CBF} = \frac{6000\lambda}{2\alpha} \frac{\Delta M}{S_0} \frac{e^{\text{PLD}/T1_{\text{blood}}}}{T1_{\text{blood}}(1 - e^{-\tau/T1_{\text{blood}}})} [\text{ml}/100\text{g}/\text{min}] \quad (4.11)$$

4.3.3 ASL quantification considerations

Recommendations on how to acquire and quantify ASL data were published by Alsop et al. (2015) and are referenced extensively below.

Labelling scheme. ASL labelling techniques are generally divided into three categories; continuous ASL (CASL), pseudo continuous ASL (pCASL), and pulsed ASL (pASL).

For CASL, the labelling pulse occurs continuously in a thin labelling plane positioned near the parotid gland. This allows a bolus of blood to be labelled as the blood flows through the neck. The main drawback of CASL is that long RF pulses are demanding on hardware and would require extra coils; this is a barrier to implementation (Borogovac and Asllani, 2012), therefore CASL is not used on clinical scanners.

As a solution to this, pCASL uses a train of short, shaped RF pulses in combination with gradients to create the effect of a continuous pulse, without requiring specialised hardware. This technique can also be adapted to label individual blood vessels rather than all of the vessels which pass through the labelling plane, potentially yielding additional diagnostic information (Boro-

govac and Asllani, 2012).

The third implementation of labelling is pASL, which uses short pulses to label spins within an inversion slab (volume) to create the bolus, instead of using a longer labelling pulse applied in a plane as in CASL. Due to the use of a labelling slab, where all spins are labelled at the same time, the T1 decay of the signal across the labelled bolus must be taken into account in the modelling. The most inferior part of the bolus will take longer to reach the tissue than the superior part, as it is physically further from the tissue, and hence will have decayed more.

In general, pCASL labelling is the preferred labelling scheme over pASL as a higher signal can be achieved. This is due to a combination of T1 decay across the labelling slab in pASL and the longer labelling durations achievable in pCASL.

Labelling delay and duration. To allow the labelled bolus to reach the tissue there must be a delay between labelling and image acquisition, often referred to as the post labelling delay (PLD). For pCASL this is well defined, as the labelling pulse has a known duration, τ , and the PLD is the time between the end of the labelling and the beginning of the acquisition. For PASL the label duration is short, making the PLD difficult to estimate, which affects quantification. If a QUIPSS-II (Quantitative imaging of perfusion using a single subtraction, second version) pulse is applied to saturate the tail of the bolus and give a defined end time, then the PLD is defined making quantification possible.

Buxton et al. (1998) originally proposed that images could be acquired at multiple PLDs within a single experiment. This allows the estimation of arterial transit time (ATT) as well as CBF, which means that differences in transit time across the brain and between patients can be accounted for. However, as this approach requires many more images to be acquired and more complex modelling this technique was not recommended by Alsop et al. (2015), although it continues to be used in research. In an attempt to minimise the

effects of ATT on CBF estimation Alsop et al. (2015) recommended different PLDs depending on the demographics of the patient being scanned, to ensure that the bolus has reached the tissue before imaging while minimising signal loss due to T1 decay.

Background suppression and crusher gradients. In a typical ASL experiment the difference between label and control images constitutes $<1\%$ of the signal in each voxel. This results in a low SNR and makes the technique highly sensitive to motion of the static tissue. Background suppression can be used to increase the SNR and reduce the sensitivity of the measurement to motion by applying 180° inversion pulses across the brain. The 180° pulses are timed such that image acquisition occurs as the signal from static tissue passes through zero, minimising its contribution to the image.

Another unwanted signal which can affect quantification comes from vascular artefacts, caused by labelled spins remaining in large arteries during readout. This signal can be removed using crusher gradients which dephase it, but this can reduce SNR. As a general rule this approach is not recommended by Alsop et al. (2015), as it results in a decreased SNR, and the PLD can be selected to minimise the amount of label left in the blood. However, Alsop et al. (2015) suggest crusher gradients may be useful in specific cases, such as when the artefacts will obscure other pathologies.

Readout schemes. The ideal image readout scheme would be a 3D acquisition, which would maximise the efficiency of background suppression and SNR, which is insensitive to field inhomogeneities, and fast enough to allow full coverage of the brain. Alsop et al. (2015) recommend segmented (multi-shot) 3D readout schemes such as 3D gradient and spin echo (GRASE) or Rapid Acquisition with Refocused Echoes (RARE stack of spirals, also known as turbo or fast SE, see chapter 2.3.2), as they have high SNR, low sensitivity to $T2^*$ effects, and single excitation per TR, which maximises the efficiency of background suppression. Single shot 3D sequences were not considered by Alsop et al. (2015) due to the lack of evidence available at the time of the review to

evaluate them, and to the presence of image blurring due to signal decay over the acquisition. 2D readouts can be used, but the background suppression will not work as effectively as it relies on acquiring the image when the tissue signal passes through zero, which will not be the case for all 2D slices. EPI based readouts have been used extensively in the literature and in clinical practice, but it is important to note that they are sensitive to differences in $T2^*$, which causes distortion and signal loss near air/tissue interfaces.

Estimation of other model parameters. For absolute quantification the equilibrium magnetisation of arterial blood ($M_{0,a}$) is needed as this is included in the arterial input function. Instead of measuring this directly, as finding a pure arterial voxel at the resolution used in ASL is impossible, this is estimated by measuring the equilibrium magnetisation in tissue (S_0) and dividing it by the blood-brain partition coefficient (λ). This is generally done on a voxel by voxel basis, either by acquiring a proton density image with a long TR, or by fitting a series of saturation recovery images acquired at different saturation times for both T1 and S_0 .

For simplicity, Alsop et al. (2015) recommend set values for the remaining parameters in equation (4.11), namely labelling efficiency, blood-brain partition coefficient, and T1 of blood.

4.4 Conclusion

This chapter demonstrates that a simple model can be used to estimate absolute CBF from ASL data, however when doing so it is important to consider the many assumptions that have been made and whether they are valid. Furthermore, for ASL there are many different acquisition and processing methodologies that can be used which will affect the resulting CBF estimation. Alsop et al. (2015) made recommendations on how to select a suitable combination to ensure that reasonable CBF maps are produced. However it is important to note that Alsop et al. (2015) considered availability and ease of use, as well as accuracy, when making the recommendations and as such following the

recommendations may not produce optimal CBF estimations.

Chapter 5

Simplified Reference Tissue Model with reduced acquisition time

Chapter 3 reviewed the available techniques for PET quantification, and demonstrated the trade-off between quantitative accuracy and acquisition time, due to the importance of accounting for tracer delivery and washout. In this chapter, a novel approach for combining blood flow information from simultaneously acquired MRI into PET pharmacokinetic modelling to significantly reduce the acquisition time required for quantification is described. This method is referred to as the reduced acquisition time simplified reference tissue model (RT-SRTM). For this preliminary study the framework was evaluated for [^{18}F]-florbetapir on a small cohort of healthy older subject and was initially published in Scott et al. (2016).

5.1 Introduction

As described in section 3.2.3, the simplified reference tissue model (SRTM) is commonly applied to the kinetic modelling of neurological PET data, as it is a robust model which doesn't require invasive blood samples. The model contains three parameters: R_1 which is the delivery of tracer to the target tissue relative to the reference region, k_2 which is the rate constant from target tissue

to blood, and the parameter of interest BP_{ND} which represents the amyloid burden when analysing amyloid PET data.

The SRTM is generally fitted to 60 minutes or more of data, which may approximately be divided into two phases. In the early phase the signal is dominated by the delivery of the tracer. This data contains information about blood flow (Hsiao et al., 2012), and is therefore important for the estimation of R_1 . The later part of the signal, i.e. 20:60 minutes post injection, contains more information about the binding of the tracer to the target and its subsequent washout, which is essential for the estimation of k_2 and BP_{ND} .

Consequently, if the blood flow information can be estimated independently from the PET data, then the acquisition time can be reduced, as scanning during the early phase can be omitted. The estimation of BP_{ND} then requires three steps: the estimation of R_1 from another source, the extrapolation of the reference region curve, C_R , as the model contains a convolution term which requires the full time series from injection to compute, and the fitting of the modified SRTM using ASL derived R_1 and the extrapolated reference region.

ASL-MRI, introduced in chapter 4.3, is a non-invasive imaging technique which can be used to estimate cerebral blood flow (CBF), as validated by comparison with the gold standard radiolabelled water PET (Fan et al., 2016). Since PET- R_1 is related to CBF, the ASL-CBF maps can be used in the first step of the framework: deriving R_1 from another source. The introduction of simultaneous PET-MRI scanners means that ASL and late phase PET data can be acquired concurrently. By incorporating CBF information from ASL into the PET pharmacokinetic modelling to provide early phase delivery information, the total acquisition time can be significantly reduced, increasing patient comfort and throughput, without sacrificing quantitative accuracy.

In this work the framework was implemented using PET data acquired for 30 minutes (from $t=30:60$ minutes post injection) with concurrently acquired ASL-MRI. The $t= 30:60$ minute time window was selected to represent the

late phase of the data. For comparison, the standardised uptake value ratio (SUVR), see section 3.3, was also calculated as this is the current clinically used metric for amyloid burden estimation.

5.2 Methods

5.2.1 Gold standard PET kinetic modelling

The SRTM (Lammertsma and Hume, 1996) was used as the gold standard for PET pharmacokinetic modelling, as it is commonly applied in amyloid studies (Wong et al., 2010; Hsiao et al., 2012; Chen et al., 2015; Sojkova et al., 2015b) and has been validated against pharmacokinetic modelling with arterial sampling for [¹⁸F]-florbetapir (Golla et al., 2018). The SRTM employs a reference region, which is considered to be devoid of the imaging target, to replace the plasma input function. Cerebellar grey matter was used as it is assumed to be devoid of amyloid- β Klunk et al. (2004).

The operational equation between the tracer concentration in the target tissue $C_T(t)$ and the reference region $C_R(t)$ can be formulated as in equation (5.1) for the implementation using basis functions Gunn et al. (1997). Here t denotes time with tracer injection at $t = 0$, and \otimes represents the convolution operator.

$$C_T(t) = R_1 C_R(t) + \phi C_R(t) \otimes e^{-\theta t} \quad \text{where} \quad (5.1)$$

$$\phi = k_2 - R_1 k_2 / (1 + BP_{ND}), \quad \theta = k_2 / (1 + BP_{ND})$$

The SRTM contains three parameters: R_1 , which is the rate constant of tracer delivery to the target tissue relative to the reference tissue; k_2 , which is the rate constant from target tissue to blood; and the non-displaceable binding potential BP_{ND} , which is proportional to target density i.e. the density of amyloid- β , see supplementary materials.

An in-house implementation of the SRTM using the basis function method was fitted to the PET data to derive regional gold standard parameter esti-

mates, denoted by an asterisk, of BP_{ND}^* , k_2^* and R_1^* from dynamic PET data conventionally acquired for $t \in [0, 60]$ minutes, denoted here as $t = 0, 60$ minutes.

5.2.2 RT-SRTM framework

A summary of the framework is shown in figure 5.1.

Step 1: Estimating relative tracer delivery (R_1) from ASL-CBF. R_1 is defined as $R_1 = K_1/K'_1$ where K_1 is the transfer rate constant from blood to target tissue and K'_1 is the transfer rate constant from blood to reference tissue. According to the Renkin-Crone capillary model (Renkin, 1959; Crone, 1963), introduced in chapter 4.2, the relationship between tracer delivery, K_1 , and blood flow, F , can be described as:

$$K_1 = EF = (1 - e^{-\frac{PS}{F}})F \quad (5.2)$$

The Renkin-Crone model includes a term for the net extraction of the tracer from the capillaries, E , which is dependent on the vessel permeability surface area product, PS . Under common physiological flow conditions, where PS/F is high (≥ 3), the relationship between K_1 and flow F is approximately linear. If we assume that PS is sufficiently high, the relationship between K_1 and F , and in turn the relationship between R_1 and F , can be approximated as a linear function Chen et al. (2015), as expressed in equation (5.3).

$$R_1 = \frac{K_1}{K'_1} = \beta_0 + \beta_1 \frac{F}{F'} \quad (5.3)$$

ASL can be used to measure the CBF, F , and may be converted into an R_1 estimate to use in the RT-SRTM, using the relationship from equation (5.3), where F' indicates the CBF measured in the reference region.

In this study, linear regression between R_1 and CBF was performed in a group of subjects to determine whether this approximation is valid. The slope and intercept of the linear regression, β_0 and β_1 , can account for systematic differences between the modalities and an extraction fraction of $E < 100\%$.

The derived β_0 and β_1 can then be applied to a different group of subjects to convert ASL-CBF to a derived R_1 value.

Step 2: Extrapolating reference region curve (C_R). To compute the convolution term in equation (5.1), C_R must be known from injection, at $t = 0$, to the end of the scan, $t = 60$. However, when reducing the acquisition time, C_R is only measured between t_s and t_e . Here the *scaled mean C_R* method to estimate the whole reference input $C_R(t)$ for $t \in [0, 60]$ is used. For this method a mean population reference region curve, $\overline{C}_R^{pop}(t)$, is scaled to the subject's measured $C_R(t)$. The scaling factor, α , is determined by a least squares fit of $\overline{C}_R^{pop}(t), t \in [t_s, t_e]$ to $C_R(t), t \in [t_s, t_e]$. The subject's $C_R(t), t \in [0, 60]$ is then calculated by multiplying $\overline{C}_R^{pop}(t), t \in [0, 60]$ by α , as in equation (5.4). The extrapolated values are only used where measured data is not available, i.e for $C_R(t) \ t \notin [t_s, t_e]$.

$$C_R(t) \approx \alpha \overline{C}_R^{pop}(t) = K'_1 e^{-k'_2 t} \otimes \gamma \overline{C}_p^{pop}(t) \quad (5.4)$$

The *scaled mean C_R* method models the reference tissue as a single tissue compartment, as given by equation (3.5), which can then be re-written as in the right hand side of equation (5.4). Looking at the underlying assumptions of this approach, this method uses a population based arterial input function, $C_p^{pop}(t)$, with a subject specific scaling factor, γ , and a population based washout rate, k'_2 , whilst allowing for a subject specific tracer delivery rate, K'_1 . This method provides a simple way to extrapolate the measured C_R and the inclusion of scaled population based arterial input functions and reference tracer washout rates is supported by their use in the literature (Logan et al., 1996; Schain et al., 2016).

Step 3: Fitting the RT-SRTM. To apply the modifications from steps 1 and 2 to the SRTM with reduced acquisition time, the operational equation in (5.1) is re-written as in equation (5.5). This groups the measured parameter $C_T(t)$, with the derived $C_R(t)$ and derived R_1 into a dummy variable, $C_T^\ddagger(t)$, as they are determined prior to fitting.

$$C_T^\dagger(t) = C_T(t) - R_1 C_R(t) = \phi C_R(t) \otimes e^{-\theta t} \quad \text{where} \quad (5.5)$$

$$\phi = k_2 - R_1 \frac{k_2}{(1 + BP_{ND})}, \quad \theta = \frac{k_2}{(1 + BP_{ND})}$$

To solve equation (5.5) for a reduced acquisition time where $t \in [t_s, t_e]$, the basis function approach (Gunn et al., 1997) is used to pre-calculate the convolution term using the extrapolated $C_R(t)$, $t \in [0, 60]$ with a range of biologically plausible values for θ . A least squares fit to the pre-determined data, $C_T^\dagger(t = t_s, t_e)$, is performed for each θ to estimate ϕ , and the instance of θ which yields the lowest sum of squares difference is selected. BP_{ND} and k_2 are then derived from ϕ , θ and the CBF-derived R_1 . The pseudo code for this is shown in algorithm 5.2.

5.2.3 Data acquisition and pre-processing

PET acquisition and image reconstruction. List mode PET data were acquired for 60 minutes following intravenous injection of [^{18}F]-florbetapir, which targets amyloid- β . For PET image reconstruction, simultaneously acquired structural T1- and T2-weighted MR images were used to synthesise CT data and calculate the attenuation map (μ -map) (Burgos et al., 2015), as validated by Ladefoged et al. (2017). The μ -map was propagated into PET space by registering the T1-weighted image to a static PET frame reconstructed at $t=55:60$ minutes (reconstructed using the scanner standard UTE derived μ -map).

Dynamic PET data were binned into 31 time frames ($15\text{s} \times 4$, $30\text{s} \times 8$, $60\text{s} \times 9$, $180\text{s} \times 2$, $300\text{s} \times 8$), and reconstructed into $2 \times 2 \times 2\text{mm}$ voxels using the manufacturer's software. An ordinary Poisson ordered subset expectation maximisation (OP-OSEM) algorithm was used with 3 iterations, 21 subsets, and a 3.5mm Gaussian filter, accounting for dead-time, attenuation, scatter, randoms and normalisation.

Algorithm 5.2 RT-SRTM with basis functions

```

1: For each subject:
2:  $C_R(t)$ : extrapolated reference region concentration
3:  $R_1$ : derived from ASL-CBF
4:  $n_{basis}$ : number of basis functions
5:  $B_i(t) = C_R \otimes e^{-\theta t}$ : basis functions ( $i = 1 : n_{basis}$ )
6:  $B_i(t_k)$ : basis functions for measured time frames ( $k = t_s : t_e$ )
7:  $w = \frac{\text{frame duration}}{e^{-\lambda t}}$ : weights
8:  $W = \text{diag}(\sqrt{w})$ : weighting matrix
9:  $W(t_k) = W(t_s : t_e)$ : weighting matrix for measured time frames

10: for  $i = 1 : n_{basis}$  do
11:    $A = [B_i(t_k)]$ 
12:    $[Q, R] = \text{qr}(W(t_k) \cdot A)$ 
13:    $M(i, :) = R \setminus Q^T$ 
14: end for

15: For each region/voxel:
16:  $C_T(t_k)$ : measured region/voxel concentration
17:  $C_T^\ddagger(t_k) = C_T(t_k) - R_1 C_R(t_k)$ 

18: for  $i = 1 : n_{basis}$  do
19:    $[\phi] = M(i, :) \cdot W(t_k) \cdot C_T^\ddagger(t_k)$ 
20:    $\hat{C}_T(t_k) = B_i(t_k) \cdot \phi$ 
21:    $SSD_i = \sum_{k=t_s}^{t_e} W(t_k) \cdot (C_T(t_k) - \hat{C}_T(t_k))^2$ 
22: end for

23:  $i_{min} = \text{find}(\text{min}(SSD))$ 
24:  $\phi = M(i_{min}, :) \cdot W(t_k) \cdot C_T^\ddagger(t_k)$ 
25:  $k_2 = \phi + R_1 \cdot \theta$ 
26:  $BP_{ND} = \frac{k_2}{\theta} - 1$ 

```

ASL acquisition and CBF quantification. CBF was estimated from a pseudo-continuous ASL (PCASL) acquisition with the following parameters: 3D GRASE readout (Günther et al., 2005) with 36 partitions and a reconstructed voxel size of $1.88 \times 1.88 \times 4$ mm, TE/TR=20.3/4000ms, 4-shot with turbo-factor/EPI-factor=14/28, bandwidth 2298Hz/pixel; 10 control-label pairs were acquired with a pulse duration (τ) and post labelling delay (PLD) both equal to 1800ms. The acquisition duration was 5min 20s ($t = 55, 60$). CBF maps were computed with equation (5.6) (Buxton et al., 1998)

$$\text{CBF} = \frac{6000\lambda \Delta S}{2\alpha S_0} \frac{e^{\text{PLD}/T1_{\text{blood}}}}{T1_{\text{blood}}(1 - e^{-\tau/T1_{\text{blood}}})} [\text{ml}/100\text{g}/\text{min}] \quad (5.6)$$

with 0.9 ml/g for the plasma/tissue partition coefficient (λ), a blood T1 of 1650 ms ($T1_{\text{blood}}$), and a labelling efficiency of 0.85 (α) as recommended in the ASL consensus paper (Alsop et al., 2015). ΔS is the signal difference between the control and label images, S_0 maps were estimated by fitting saturation recovery images acquired with the same sequence at three different saturation times (1,2,4s) using NiftyFit (Melbourne et al., 2016).

Regional analysis. Figure 5.1 shows the data pre-processing stages where T1-weighted images were parcellated (Cardoso et al., 2015) into 17 regions: amygdala, brainstem, caudate, cerebellum (white and grey separately), cerebral white matter, hippocampus, pallidum, putamen, thalamus and 6 cortical grey matter regions, with left and right hemispheres combined. The T1-weighted image was rigidly registered to both ASL and PET space (Modat et al., 2014a), and the transformation was propagated to the parcellation. Regional average CBF values were calculated (negative CBF values were excluded from averaging), and the PET time activity curves were averaged across the region prior to kinetic modelling. PET data acquired during 30:60 minutes were used to evaluate the proposed method.

For comparison, regional SUVR values were calculated from a static image for $t = 50 : 60$ minutes created by summing the last two reconstructed dynamic PET frames ($t = 50 : 55$ and $t = 55 : 60$ minutes). Regional averages were calculated using the registered parcellation and divided by the mean cerebellar grey matter tracer concentration. As for the kinetic modelling, cerebellar grey matter was used as the reference region. To compare SUVR with BP_{ND} , $\text{SUVR}-1$ was used, as SUVR approximates DVR which is equal to $BP_{\text{ND}} + 1$.

5.2.4 Data

The data used here consisted of the first 20 subjects scanned for Insight46, a neuroimaging sub-study of the Medical Research Council National Survey of

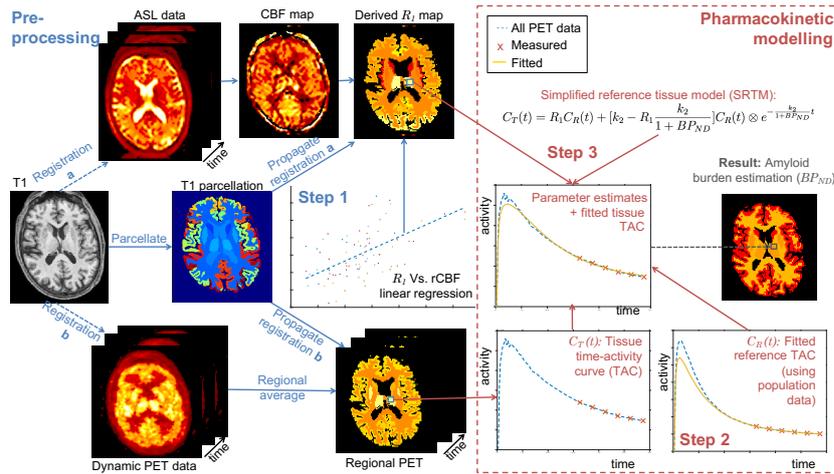


Figure 5.1: Overview- parcellation is registered to PET and ASL to calculate regional average values. ASL data is converted into CBF-derived R_1 values using the linear regression relationship. A population reference tissue time activity curve of 0:60 mins combined with the measured reference tissue data (30:60 mins) is used with the CBF-derived R_1 and the measured PET tissue data (30:60 mins), to apply the modified simplified reference tissue model to estimate BP_{ND} .

Health and Development Lane et al. (2017). These 20 subjects had a mean age of 69.5 years (range 69.3-69.8 years) with an MMSE score of ≥ 28 . A description of the full cohort can be found in appendix A. The 20 subjects were randomly divided into a training set and a testing set, see figure 5.2. The training set was used to estimate the relationship between ASL-CBF and PET- R_1 ($n=5$) and to generate a population mean C_R for the reference region extrapolation ($n=14$), and the testing set was used to apply the complete framework ($n=6$).

Figure 5.2 also shows the proportion of subjects found to be amyloid positive ($a\beta^+$) as defined using whole cortical grey matter SUVR with whole cerebellum reference region and a cut-off of 1.077. Whilst the group assignment of subjects was performed randomly (positivity was determined subsequent to this analysis), the $a\beta^+$ subjects are evenly split between groups and comprise 15% of the dataset which is comparable to the positivity rate across the whole Insight46 study (18.6%).

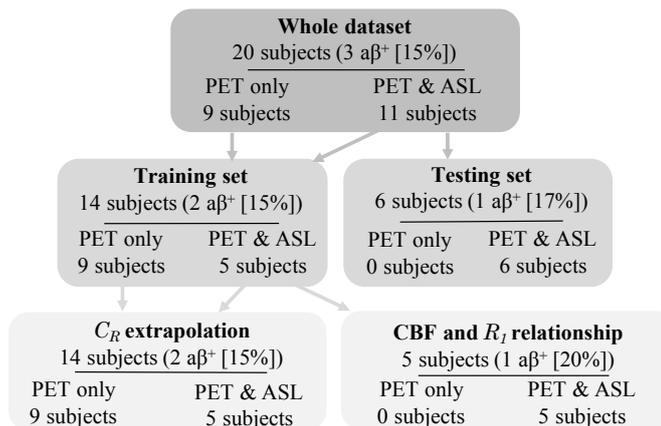


Figure 5.2: Division of data for training and testing

5.2.5 Statistical Analysis

To compare different techniques with the gold standard, mean square error ($\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - Y_i^*)^2$) and mean error ($\text{ME} = \frac{1}{n} \sum_{i=1}^n (Y_i - Y_i^*)$) were used, where n is the number of estimates, Y_i^* is the gold standard estimate or measured value and Y_i is the estimate being evaluated. To compare different techniques, statistical significance was tested using paired, 2-tailed Wilcoxon signed-rank test for MSE (as the data are not normally distributed), and paired 2-tailed t-tests for ME.

5.3 Results

5.3.1 Comparison of proposed method with gold standard

Figure 5.3a shows BP_{ND} estimated using the proposed RT-SRTM with ASL-CBF and 30:60 minutes of PET data plotted against the gold standard using the SRTM and the full 60-minutes of dynamic PET data. Linear regression of all subjects and regions shows that the proposed method offers a good approximation of the gold standard as it closely follows the line of identity (blue dashed line), which is within the 95% confidence interval (CI) of the regression (shaded area). Furthermore, the Pearson correlation coefficient shows a high linear correlation ($r=0.9480$, $p < 0.001$, 95% CI [0.9228, 0.9650]) and a paired

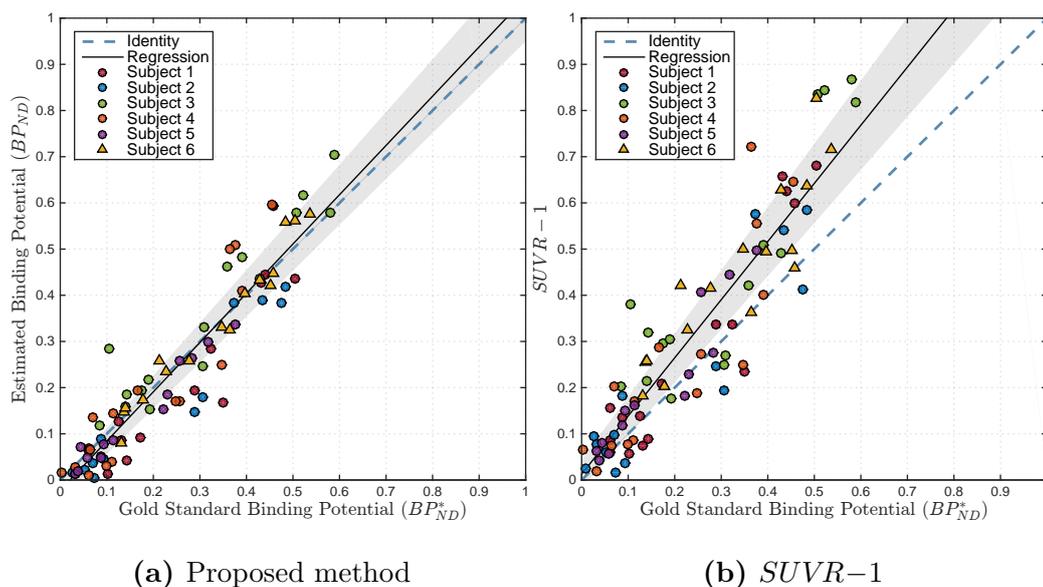


Figure 5.3: Estimated amyloid burden against the gold standard value calculated using full PET time series.

t-test shows that the estimate of BP_{ND} using the RT-SRTM is not significantly different from the gold standard ($p=0.2674$).

The alternative measure used in clinical practice, $SUVR-1$, was calculated from PET data over 50:60 mins for comparison, Figure 5.3b. Whilst the overall correlation is still high ($r=0.9209$), a clear bias is shown as $SUVR-1$ overestimates the binding potential at higher values. The mean error (ME) quantifies the bias between the estimates and the gold standard which is significantly higher for $SUVR-1$ (ME=0.0721, $p < 0.001$), indicative of the systematic overestimation, compared to -0.0071 for the proposed method. The proposed method also has a lower mean square error (0.0039 compared to 0.0153 for $SUVR-1$, $p = 0.0313$), and variance (0.0039 compared to 0.0103 for $SUVR-1$).

Figures 5.4 and 5.5 show BP_{ND} maps for an amyloid positive and amyloid negative subject respectively, comparing the gold standard with the proposed method and $SUVR-1$ regionally. The proposed RT-SRTM with ASL-CBF shows good agreement with the gold standard, with slight overestimation of the cortical white matter in both cases. This is reduced when using the gold

standard R_1^* in the RT-SRTM rather than deriving R_1 from ASL. For the $SUVR-1$ estimation, amyloid burden is greatly overestimated within both grey and white matter structures, particularly in the amyloid positive subject. The difference map shows that the errors in the proposed method are far lower than for $SUVR-1$.

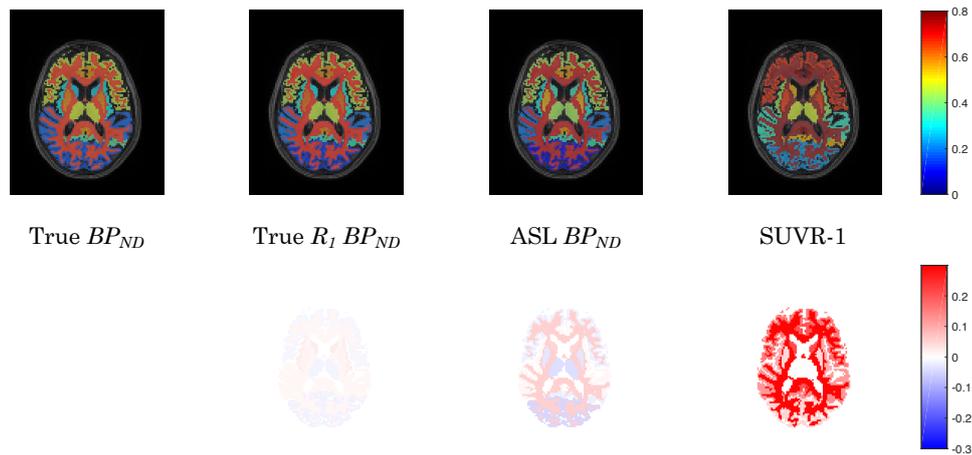


Figure 5.4: Example amyloid positive ($A\beta^+$) subject (subject 6). **Top:** Regional average binding potential maps for (left to right) gold standard, proposed method using gold standard R_1 , proposed method using ASL and $SUVR-1$. **Bottom:** difference maps compared to gold standard.

5.3.2 Influence of CBF-derived R_1 estimation on amyloid quantification

Whilst Figure 5.3a demonstrates a high similarity between binding potential estimation using the gold standard and the proposed method, there is a noise component which introduces variation around the line of identity. This is due to noise in the PET data, noise in the CBF-derived R_1 estimate from the ASL data, and inaccuracies in the estimation of the reference tissue input.

To demonstrate the influence of the CBF-derived R_1 estimate using ASL data, the proposed method was applied using the R_1^* estimated using the gold standard technique instead of the CBF-derived R_1 . The population input function and 30:60 mins PET data were used as before. This represents the optimal case in which R_1 can be determined exactly from the ASL data. Fig-

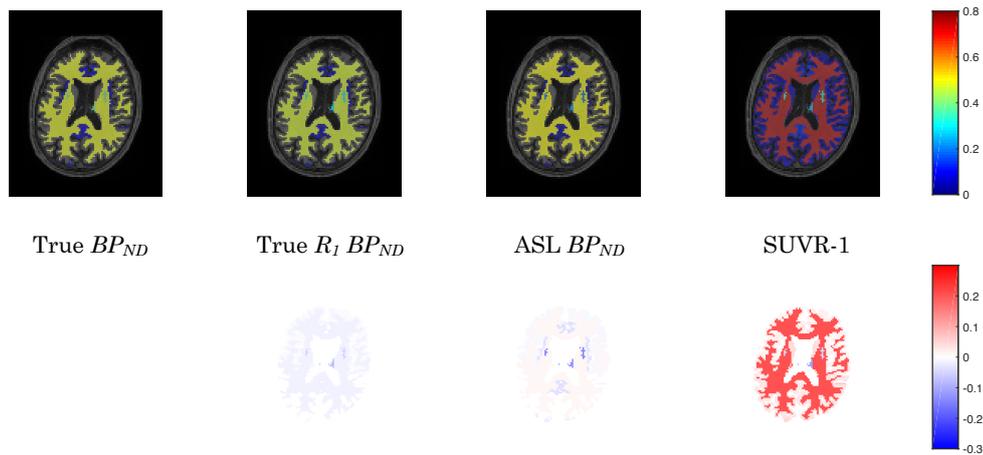
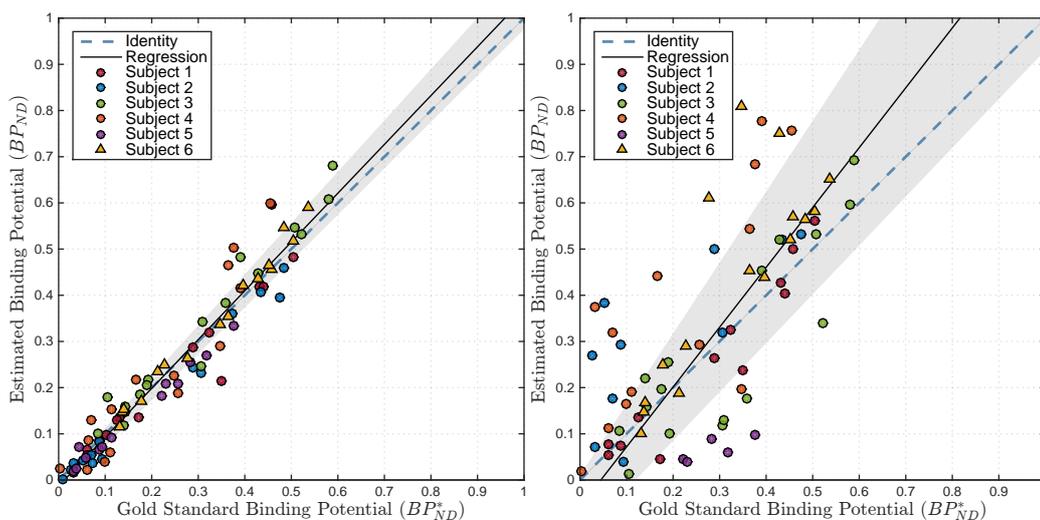


Figure 5.5: Example amyloid negative ($A\beta^-$) subject (subject 2). **Top:** Regional average binding potential maps for (left to right) gold standard, proposed method using gold standard R_1 , proposed method using ASL and $SUVR-1$. **Bottom:** difference maps compared to gold standard.

Figure 5.6a shows that the variance in the binding potential estimate has been reduced (from 0.0039 to 0.0020), a narrower 95% CI in the regression. This is expected since the CBF map from the ASL is noisy, and linear regression performed to determine the relationship between CBF and R_1 was performed with only 5 subjects, and therefore may not be generalisable. Furthermore, the differences shown in the white matter regions in figures 5.4 and 5.5 when using CBF-derived R_1 or R_1^* suggest that the relationship may vary regionally. Nevertheless, for the data used in this study the estimation of R_1 from CBF is sufficiently accurate that the BP_{ND} estimates between the proposed method using CBF-derived R_1 and gold standard R_1 are comparable and there is a reduced bias in the estimates compared to $SUVR-1$.

Figure 5.6b compares BP_{ND} estimation using 30:60 mins PET data only (with extrapolated C_R) to the gold standard to demonstrate the need of a CBF-derived R_1 . Due to the lack of data to support the kinetic modelling, the results are noisy and extreme parameter estimates occurred for some regions. These points have skewed the linear regression such that it no longer follows the identity line, with a wide 95% CI.



(a) 30:60 mins PET + gold standard R_1 (b) 30:60 mins PET data only

Figure 5.6: Estimated binding potential plotted against the gold standard value calculated using full PET time series

5.3.3 Influence of C_R extrapolation on amyloid quantification

The remaining errors in BP_{ND} estimation in figure 5.6a are due either to errors in the extrapolation of the reference region or the uncertainty in fitting the model to fewer datapoints. Figure 5.7 shows the extrapolated reference region curves for the six subjects. The *scaled mean C_R* method works well for most subjects, however it greatly underestimates the peak for subject 4. This is likely to be due to the strict assumptions that a population arterial input function and washout rate can be used to describe an individual subject. Figure 5.6a shows that this propagates into the estimate of BP_{ND} as this subject consistently falls outside the 95% confidence interval in the regression across all subjects.

5.4 Discussion and Conclusion

This work demonstrates that the proposed RT-SRTM using just 30:60 minutes of PET data together with blood flow information from ASL produces estimates of amyloid burden which are comparable to full pharmacokinetic mod-

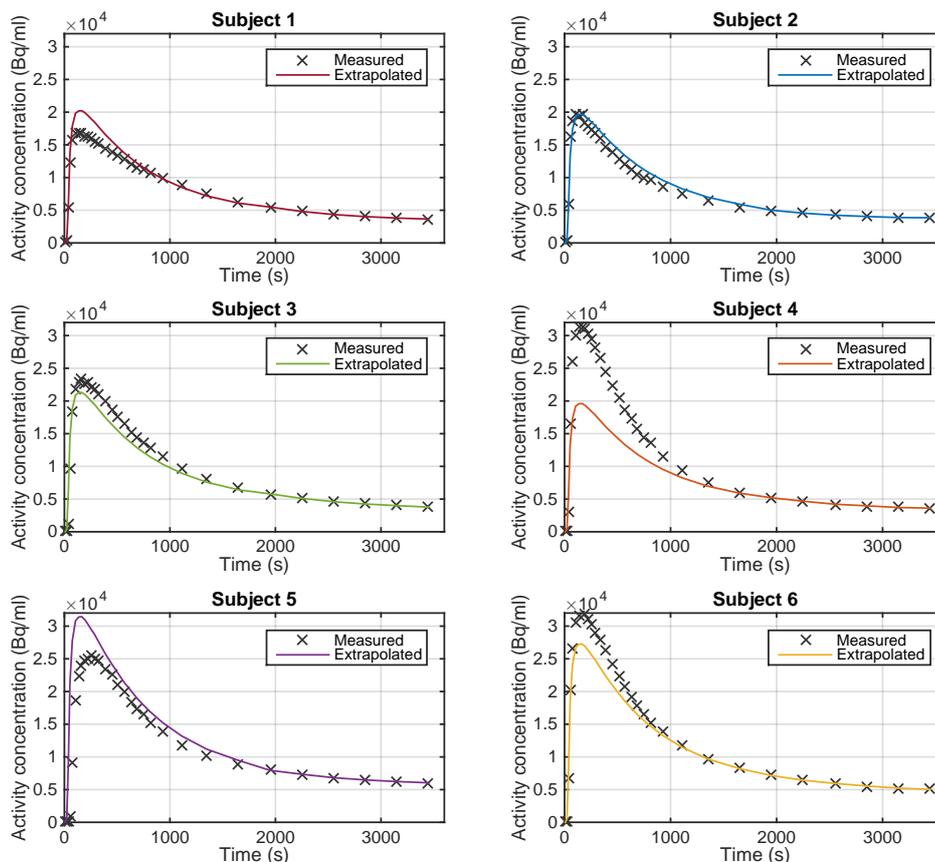


Figure 5.7: Extrapolated reference regions (C_R)

elling of 0:60 minutes of [¹⁸F]-florbetapir PET data. The proposed method is more accurate than the simplified estimate of amyloid burden, $SUVR-1$, which showed a positive bias especially at higher binding potential values.

The results of the proposed technique depend on the CBF-derived R_1 estimate from the ASL data. The ASL data used here were acquired for only 5 minutes without motion correction, and thus susceptible to artefacts and noise. Linear regression between CBF and R_1 using just 5 subjects could produce errors which may propagate to the binding potential estimation and cannot account for potential regional differences in the relationship.

Furthermore, the *scaled mean* C_R method for extrapolating C_R works well for most subjects but the strict assumption that population values can be scaled to fit a subject was violated in one case, suggesting a more flexible

approach may be required in the future. However, the technique shows promise and has the potential to be applied to other PET tracers which bind to other biological targets of interest.

Chapter 6

Correlation of PET tracer delivery and ASL cerebral blood flow

Chapter 5 introduces a framework for reduced acquisition time PET quantification (RT-SRTM) which requires the ASL-CBF maps to be converted into the relative PET tracer delivery rate constant, R_1 . This chapter reviews the existing literature which evaluates the correlation between PET based CBF measurement, PET- R_1 , and ASL-CBF. Then the relationship between PET- R_1 and ASL-CBF is tested using [^{18}F]-florbetapir PET data with simultaneously acquired ASL-MRI. This work has been published in Scott et al. (2018c).

6.1 Introduction

The first step of the RT-SRTM method proposed in chapter 5 is to convert ASL-CBF maps into PET- R_1 maps. This was achieved using the theoretical Renkin-Crone model, described in chapter 4.2, which offers a way to equate the tracer delivery rate constant, K_1 , to CBF. This model can also be adapted for the parameter of interest, R_1 , by normalising to the reference region. When applying this model in chapter 5, it was assumed that the extraction of the tracer from the blood plasma into the tissue was sufficiently high that the Renkin-Crone model could be reduced to a linear regression (equation (5.3)),

however there was limited data available to validate this assumption.

There is currently no published experimental data directly comparing amyloid PET K_1 or R_1 with ASL-CBF. However, the gold standard CBF measurement H_2O^{15} -PET (also known as water-PET) has been compared to amyloid PET K_1 , R_1 and ASL-CBF, so it is used here as an intermediate step to determine which factors influence the correlation between these parameters.

6.1.1 Comparison of ASL-CBF with H_2O^{15} -PET

Fan et al. (2016) carried out an extensive review of literature which compares the CBF estimated using H_2O^{15} -PET with ASL. This review highlights the difficulty in drawing definite conclusions from multiple studies due to the wide range of study designs, ASL acquisition implementations, and analysis techniques used. Whilst recommendations have been issued in an attempt to standardise ASL acquisition and analysis (Alsop et al., 2015), this has yet to have a significant impact on comparative ASL to H_2O^{15} -PET literature. Furthermore, since the recommendations were designed to be easily implementable across many sites, it has been suggested that the recommended methodology may not be optimal for quantifying CBF using ASL (van Osch et al., 2017).

Overall, Fan et al. (2016) found a high correlation between whole brain CBF estimated by the two modalities (e.g $r^2 = 0.6$, Heijtel et al. (2014)), however regional correlation tended to be much lower (e.g $r^2 = 0.34$, van Golen et al. (2014)). Physiological, acquisition, and processing effects all influence the correlation between the two modalities. Here it is assumed that most of the variability comes from ASL, as H_2O^{15} -PET is the gold standard of CBF measurement and different H_2O^{15} -PET acquisition and processing techniques generally produce consistent results (Boellaard et al., 2006)

Physiological effects. There are two primary physiological factors which affect the accuracy of ASL-derived CBF measurements; differences in flow rates, and differences in ATT.

In areas of the brain with lower flow rates there will be fewer labelled spins delivered to the tissue, and therefore the signal difference between the

control and label images is smaller, decreasing the already low SNR. In healthy subjects this is observed in the 2-3.6 times lower flow rates in white matter tissue compared to grey matter (Van Osch et al., 2009). Zhang et al. (2014) showed that simultaneous ASL (pCASL) and H_2O^{15} -PET yield similar mean white matter CBF values, but the ASL gives a much higher coefficient of variation (56% for ASL compared to 29% for PET). This indicates that, in this case, lower SNR leads to increased variability.

Furthermore, the ATT is longer in white matter compared to grey matter, which means that for single inversion time techniques the labelled bolus may not have reached in the tissue before image readout. This was demonstrated by Ye et al. (2000) who found no significant difference in CBF estimation between ASL (using CASL) and H_2O^{15} -PET in a cortical grey matter region, but a statistically significant 30% difference in cortical white matter.

Increased ATT is also a hallmark of ageing and disease and, as for the white matter regions, this can lead to images acquired before the bolus has reached some brain regions (Fan et al., 2017), or to the signal decaying before it reaches the tissue. This physiological effect will influence the choice of parameters selected for the acquisition.

Acquisition effects. Overall, Fan et al. (2016) found that global ASL-CBF literature values were generally within 15% of the PET-CBF values, however the systematic errors depended on the choice of labelling scheme. It was found that pCASL tends to overestimate CBF, whereas PASL underestimates. This may be due to intrinsic differences in the labelling schemes, or could be related to differences in the delay between labelling and readout between the two techniques.

In general longer delay times between labelling and readout are desirable, as this ensures that the bolus has had enough time to reach the tissue (Fan et al., 2017). Otherwise signal can remain in the arteries, causing an overestimation of the CBF in the surrounding area compared to H_2O^{15} -PET. This was noted in the study by Heijtel et al. (2014), which showed that ASL over-

estimated CBF in deep cortical tissues which are highly vascularised. This effect was reduced when applying crusher gradients which remove the signal in the blood vessels, however this approach was not recommended in the ASL consensus paper (Alsop et al., 2015), as it greatly reduces the overall SNR.

Whilst an increased delay between labelling and readout ensures that the bolus has reached the tissue, it is important to consider that the longer the delay, the lower the measured SNR. This is because the signal of the tagged spins is decaying at a rate determined by the T1 of the surrounding blood or tissue. It also means that any blood which arrives to the tissue early will decay with the T1 of tissue rather than blood, where the T1 of tissue is much shorter than that of blood. This is known as the *T1 effect* and could introduce regional errors depending on ATT, unless more complex, multi compartmental modelling is applied (Parkes and Tofts, 2002).

The readout method used also has a large influence on the regional ASL-CBF values. Alsop et al. (2015) recommend segmented 3D sequences such as 3D GRASE to maximise SNR and the efficiency of background suppression, whilst minimising the influence of local field inhomogeneities which affect echo planar imaging (EPI) sequences. Kilroy et al. (2014) performed a head-to-head comparison between pCASL with 3D GRASE and 2D EPI readouts. This showed that 3D GRASE had better repeatability over time and better consistency between patient groups compared to 2D EPI. However, when comparing both methods to H₂O¹⁵-PET, the results suggested a better correlation between H₂O¹⁵-PET and the 2D EPI readout than 3D GRASE, although this difference did not reach statistical significance.

Processing effects. The literature also shows that there are methods to reduce the errors between H₂O¹⁵-PET and ASL-CBF. The simplest way to account for global errors is to normalise the CBF values to a reference region, as performed in Goetti et al. (2014). Here the cerebellum was used as the reference region, and the correlation between PET and ASL-CBF rose from $r = 0.27$ for absolute values to $r = 0.67$ for relative CBF. The main challenge with this

approach is defining an appropriate reference region, however for comparison with R_1 , which includes a reference region by definition, this approach is ideal.

PVEs also have an influence on both PET and ASL estimates of CBF. Here, the limited spatial resolution of the PET scanner, combined with the finite voxel size used to represent both modalities will introduce a dependence of CBF estimates on region size and the surrounding voxel intensities. In comparative literature this is generally addressed by smoothing the ASL-CBF maps to match the resolution of the PET data (van Golen et al., 2014).

Finally, it is trivial to see that the larger the temporal gap between the H_2O^{15} -PET scan and the ASL-MRI, the higher the chance that the CBF of the patient has changed, affecting the correlation between the two modalities. When performing comparison between H_2O^{15} -PET and ASL-CBF, it is not always possible to conduct the studies at the same time or even on the same day. Figure 6.1 shows how the correlation coefficient between H_2O^{15} -PET and ASL decreases as the time elapsed between scans increases. This makes it difficult to separate intrinsic differences between the techniques from physiological CBF fluctuations over time, particularly as there are so many uncontrolled factors, such as caffeine, which can influence CBF (Clement et al., 2017). This demonstrates the importance of simultaneous acquisitions for confounder-free comparisons.

6.1.2 Comparison of amyloid PET- R_1 with H_2O^{15} -PET

The only direct comparison of H_2O^{15} -PET CBF with R_1 generated from [^{18}F]-florbetapir PET data was performed by Ottoy et al. (2017a), who analysed 39 subjects (10 healthy controls, 19 patients with mild cognitive impairment (MCI), and 10 patients with AD), with an unspecified delay between PET scans. Ottoy et al. (2017a) calculated R_1 from dynamic data using a 2T4P model, with the arterial input function derived from arterial blood samples. Ottoy et al. (2017a) found statistically significant Pearson correlations between PET- R_1 and H_2O^{15} -CBF for all 8 regions analysed.

The findings of Ottoy et al. (2017a) are supported by multiple similar

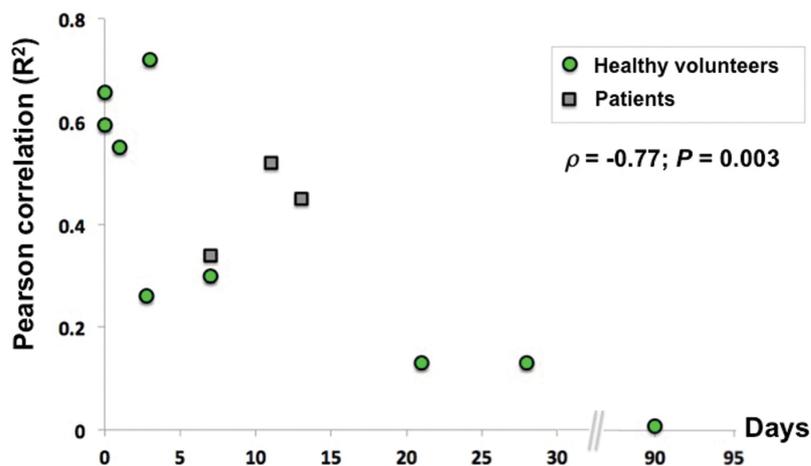


Figure 6.1: Pearson correlation coefficient between H_2O^{15} -PET and ASL reported in comparison studies, as a function of the time elapsed between scans. Reproduced from Fan et al. (2016) with permission.

studies which used ^{11}C -PiB data to estimate R_1 (Bilgel et al., 2019; Chen et al., 2015; Rodell et al., 2013; Ponto et al., 2019). The largest study was conducted by Bilgel et al. (2019). In this study, which included 149 subjects (143 healthy controls, 4 patients with MCI and 2 patients with AD), the SRTM was used to calculate ^{11}C -PiB R_1 , and H_2O^{15} and ^{11}C -PiB acquisitions were performed consecutively in the same session. The median Pearson correlation across all subjects was 0.67, although there was no separate analysis to determine whether this correlation varied between the large healthy control group and the MCI and AD subjects. This study also went on to show that R_1 can be used as a marker of change in CBF over time, although the correlation between the longitudinal change in ^{11}C -PiB R_1 and H_2O^{15} reduced to 0.42.

Chen et al. (2015) also conducted both ^{11}C -PiB and H_2O^{15} imaging in a single session for 19 subjects (6 healthy controls, 8 patients with MCI and 5 patients with AD). Various methodologies were employed to quantify the amyloid PET data, including the 2T4P model with arterial sampling, recommended in Price et al. (2005) for K_1 estimation, and the SRTM for R_1 . The standard single tissue model was applied to the H_2O^{15} -PET using K_1 as a

direct estimate of CBF.

Comparison was performed between ^{11}C -PiB and H_2O^{15} absolute quantification estimates (K_1), and relative quantification estimates (R_1) across 13 regions and one combined cortical region using Spearman's correlation coefficient (ρ). Direct comparison of ^{11}C -PiB and H_2O^{15} -PET K_1 s showed a low correlation across most regions, with only the subcortical white matter achieving a statistically significant correlation ($\rho = 0.699$).

However, when comparing the relative K_1 for H_2O^{15} -PET with R_1 from the SRTM fit of ^{11}C -PiB, all regions achieved significance with $\rho \geq 0.839$. This is partly due to the reduction of bias achieved when normalising to a reference tissue, and to the increased stability of the SRTM compared to more complex kinetic models.

Finally, Rodell et al. (2013) and Ponto et al. (2019) carried out similar comparison studies including 13 subjects (8 healthy controls and 6 patients with AD) and 24 subjects (an unspecified mix of healthy controls and patients with MCI), respectively. Both studies used the SRTM to generate R_1 from ^{11}C -PiB data and again found a good correlation between H_2O^{15} CBF and ^{11}C -PiB R_1 , with Ponto et al. (2019) reporting a mean R^2 value of 0.817. Rodell et al. (2013) also reported that group differences in estimated CBF between healthy controls and patients with AD, reduced when normalising to a reference region. This could be an issue if using R_1 to distinguish between groups but it is not an issue in this case, as the correlation of ^{11}C -PiB R_1 with H_2O^{15} CBF is the parameter of interest.

6.1.3 Correlation of R_1 and ASL for amyloid PET/MR data

The available literature shows that, at least on a regional basis, there is evidence of correlations between ASL-CBF and H_2O^{15} -PET CBF, as well as between H_2O^{15} -PET CBF and amyloid PET R_1 . This suggests that there should also be a correlation between ASL-CBF and amyloid PET R_1 . However, the comparative literature on ASL-CBF and H_2O^{15} -PET CBF suggests

that the relationship may vary by region and is strongly influenced by how the ASL data is acquired and processed. This relationship is explored here for simultaneously acquired [^{18}F]-florbetapir PET R_1 and pCASL ASL-CBF.

6.2 Methods

6.2.1 Data

The data used within the current chapter and chapter 7 is drawn from two studies; imaging data were collected from 60 cognitively normal subjects participating in Insight 46 (1946), a neuroimaging sub-study of the Medical Research Council National Survey of Health and Development (Lane et al., 2017), and 4 subjects from a study of young onset Alzheimer’s Disease (YOAD) with an intermediate or high certainty diagnosis (McKhann et al., 2011), see appendix A for more details.

All subjects underwent 60 minutes of simultaneous amyloid PET and multi-modal MR imaging on a Siemens Biograph mMR PET/MR scanner. Of the 64 subjects used for analysis, (mean age 69.6 years, range 61.7-70.5 years), 45 had both PET and ASL data, and for 19 the ASL data were missing, either due to repetition of other scans (4) or imaging artefacts (15).

The subjects were divided into 2 sets; an optimisation set containing 39 subjects, and a testing set containing the remaining 25 subjects, see figure 6.2. The clinically diagnosed YOAD subjects were evenly split between the two sets, as were amyloid positive ($a\beta^+$) subjects from Insight 46. Amyloid positivity was defined using mean cortical grey matter SUVR with a whole cerebellum reference region, see appendix A.3 for details.

Within the optimisation set the 20 subjects with PET and ASL data were used in this chapter to derive the relationship between PET- R_1 and ASL-CBF.

6.2.2 CBF estimation with ASL-MRI

ASL data were acquired and processed as described in section 5.2.3, before being resampled into a $2\times 2\times 2$ mm³ grid to improve the parcellation into brain regions and to match the PET data. ASL data were all checked before inclu-

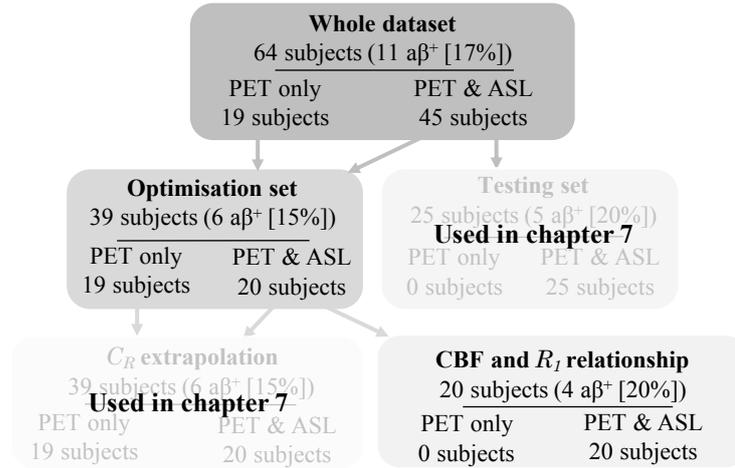


Figure 6.2: Division of data for optimisation

sion in analysis, see appendix B.2 for details. All subjects used were visually assessed to ensure that motion was minimal, that the fitting of the M_0 maps was successful, and that there were no major artefacts. However, some common artefacts were accepted due to their high prevalence, namely fat shift and flow, as described in section 6.3.2.

6.2.3 Dynamic PET acquisition and reconstruction

List mode PET data were acquired for 60 minutes following intravenous injection of [^{18}F]-florbetapir, which targets amyloid- β . For PET image reconstruction, simultaneously acquired structural T1- and T2-weighted MR images were used to synthesise CT data and calculate the attenuation map (μ -map) (Burgos et al., 2015), as validated in Ladefoged et al. (2017). The μ -map was propagated into PET space by registering the T1-weighted images to a full 60-minute non-attenuation corrected reconstructed PET image.

Dynamic PET data were binned into 31 time frames ($15\text{s} \times 4$, $30\text{s} \times 8$, $60\text{s} \times 9$, $180\text{s} \times 2$, $300\text{s} \times 8$), and reconstructed into $2 \times 2 \times 2\text{mm}^3$ voxels using the open source NiftyPET package (Markiewicz et al., 2017), see Appendix B. Spatial smoothing was applied to the images to reduce noise in the time activity curves. A 2mm Gaussian filter was found empirically to be sufficient to reduce noise in the regional time activity curves.

6.2.4 Regional analysis framework

Analysis was performed as described in section 5.2.3, using 16 regions for analysis with cerebellar grey matter as a separate reference region. Gold standard PET- R_1^* was then correlated against relative ASL-CBF (rCBF).

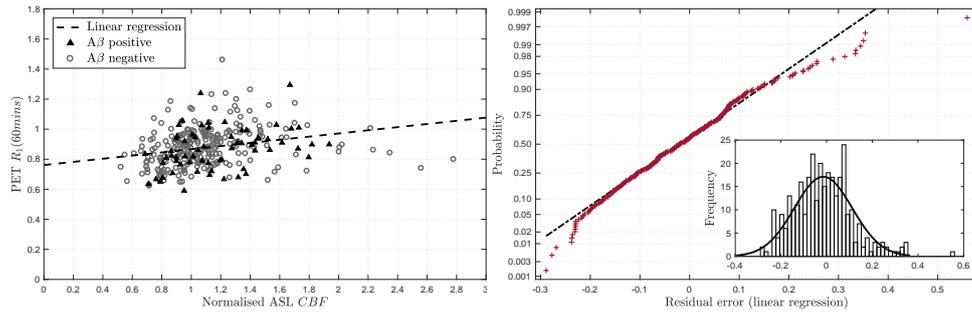
6.2.5 Linear regression between ASL-CBF and PET- R_1

Linear regression between R_1 and CBF was performed to determine whether the linear approximation in equation (5.3) is valid. The slope and intercept of the linear regression, β_0 and β_1 , can account for systematic differences between the modalities and an extraction fraction of $E < 100\%$, see section 4.2.2. The derived β_0 and β_1 can then be applied to a different group of subjects to convert ASL-CBF to a derived R_1 value.

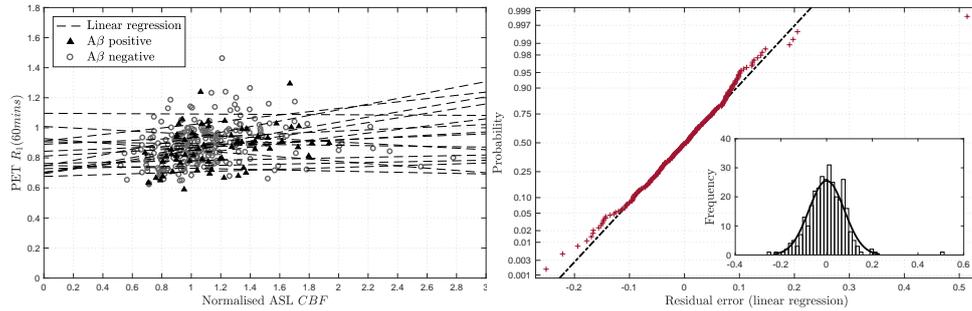
Due to systematic errors in ASL-CBF estimates in certain regions of the brain, multi-linear analysis was also performed to determine whether β_0 and β_1 may be region dependent. Multi-linear regression has the capacity to model interaction terms between ASL-CBF and the region, and can be described as in equation 6.1.

$$\begin{aligned}
 R_1 = \beta_0 + \beta_1 \frac{F}{F'} + \sum_{n=2}^{regions} \beta_n I[region_n] \\
 + \sum_{m=regions+1}^{2 \times regions} \beta_m \times \frac{F}{F'} \times I[region_m]
 \end{aligned} \tag{6.1}$$

Here $I[region_n]$ is equal to 1 when $region_n$ is being considered, and 0 otherwise. Two further multi-linear analyses were also performed i) using the subject as a covariate, and ii) using both subject and regions as covariates, to determine their relative influence on R_1 estimation.



(a) PET- R_1 Vs ASL-CBF- single linear regression (black dashed line). (b) Residual normality plot for (a). Inset: histogram of residuals



(c) PET- R_1 Vs ASL-CBF- multi-linear regression (black dashed lines). (d) Residual normality plot for (b). Inset: histogram of residuals

Figure 6.3: Correlation of PET- R_1 with ASL-CBF for 20 optimisation set subjects. Linear regression was calculated and applied to the ASL-CBF data to show the residual error in the fit.

6.3 Results

6.3.1 Global relationship between ASL-CBF and PET- R_1

Figure 6.3a shows the relationship between PET- R_1 and ASL-CBF across the 16 regions for the 20 optimisation set subjects. Linear regression shows a statistically significant correlation between the two parameters ($\rho = 0.349, p < 0.001$), however there is some variability which is not explained by the linear model. Noise and artefacts in the ASL data are considered to be the main causes of variability, however violations of the model assumptions may also contribute.

Residual analysis was performed to determine whether a non-linear model could be fitted to the data, as suggested by equation (5.2). The normality plot

for the residual error in figure 6.3a is shown in figure 6.3b. This demonstrates that the residuals follow an approximately normal distribution, which supports the use of a linear model. However, there is some deviation from normality at the extremes, which is also illustrated by the histogram inset in figure 6.3b. This shows that there are some outlying positive residuals which skew the distribution.

6.3.2 Regional relationship between ASL-CBF and PET- R_1

Using tissue type as a covariate. The literature reviewed in section 6.1.1 suggests that the bias within the ASL estimate of CBF may be region dependent. To investigate this, first the tissue type was used as a covariate in the linear regression. Figure 6.4 shows the same data as figure 6.3a, but this time a single linear regression has been applied for each tissue type. Figure 6.4 demonstrates similar relationships between PET- R_1 and ASL-CBF for grey matter, deep grey matter and white matter. The overestimation due to vascular signal in deep grey matter seen by Heijtel et al. (2014) is not seen here. This can be explained by the longer delay between labelling and acquisition, which was 1800ms in this work compared to 1525ms in Heijtel et al. (2014). However, when looking at the correlations it is clear that the strongest relationship between PET- R_1 and ASL-CBF is seen for grey matter regions. This is likely to be due to the higher blood flow in these regions which increases the SNR.

Figure 6.4 also shows that the correlation between PET- R_1 and ASL-CBF in the brainstem is very weak. The brainstem as a region was not reported separately in the comparative literature reviewed in section 6.1.1, however Warnert et al. (2014) performed some optimisation work specifically for ASL of the brainstem. Warnert et al. (2014) showed that the brainstem is a challenging region in which to perform ASL based CBF measurements due to its high vascularity, shorter transit time (approximately 100ms shorter than grey matter), and lower perfusion relative to grey matter. Warnert et al. (2014) recommend

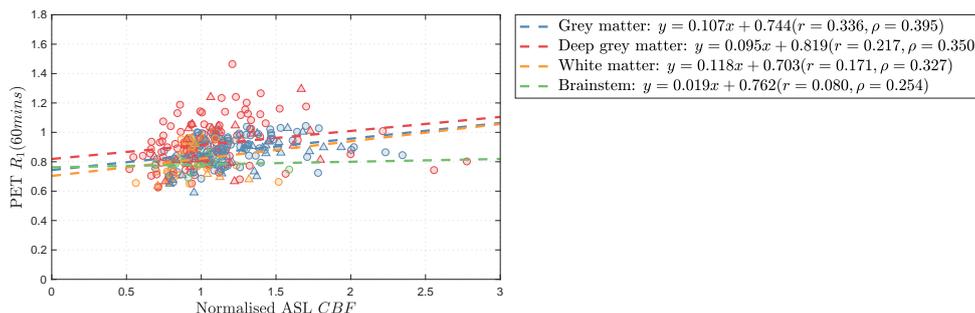


Figure 6.4: Correlation of PET- R_1 with ASL-CBF for 20 optimisation set subjects with a linear regression performed for each tissue type. \triangle indicates A β -positive subjects and \circ A β -negative

using a multi-TI ASL acquisition with a 2 compartment model to account for the shorter transit time and *T1 effect*. Consequently, a single TI measurement analysed using the single compartment method results in a noisy measurement.

Whilst figure 6.4 shows reasonable correlations for the individual tissue types, multi-linear regression using the tissue type as a covariate doesn't explain a significant amount of the variation in the data. This multi-linear regression resulted in an R^2 value of 0.163 (adjusted $R^2 = 0.144$), which is higher than that of the single linear regression (0.062, adjusted $R^2 = 0.063$), but is still relatively small. Consequently other covariates were explored to understand the relationship between PET- R_1 and ASL-CBF.

Using brain region as a covariate. To quantify the regional variation in the relationship between ASL-CBF and PET- R_1 , multi-linear analysis was also performed using the region name as a covariate. This was found to explain much more of the variation than the single linear regression and gave an R^2 value of 0.650 (adjusted $R^2 = 0.613$), compared to 0.062 (adjusted $R^2 = 0.063$) for single linear regression. Figure 6.3d shows that the residual error using multi-linear regression using brain region is lower and more normally distributed than single linear regression.

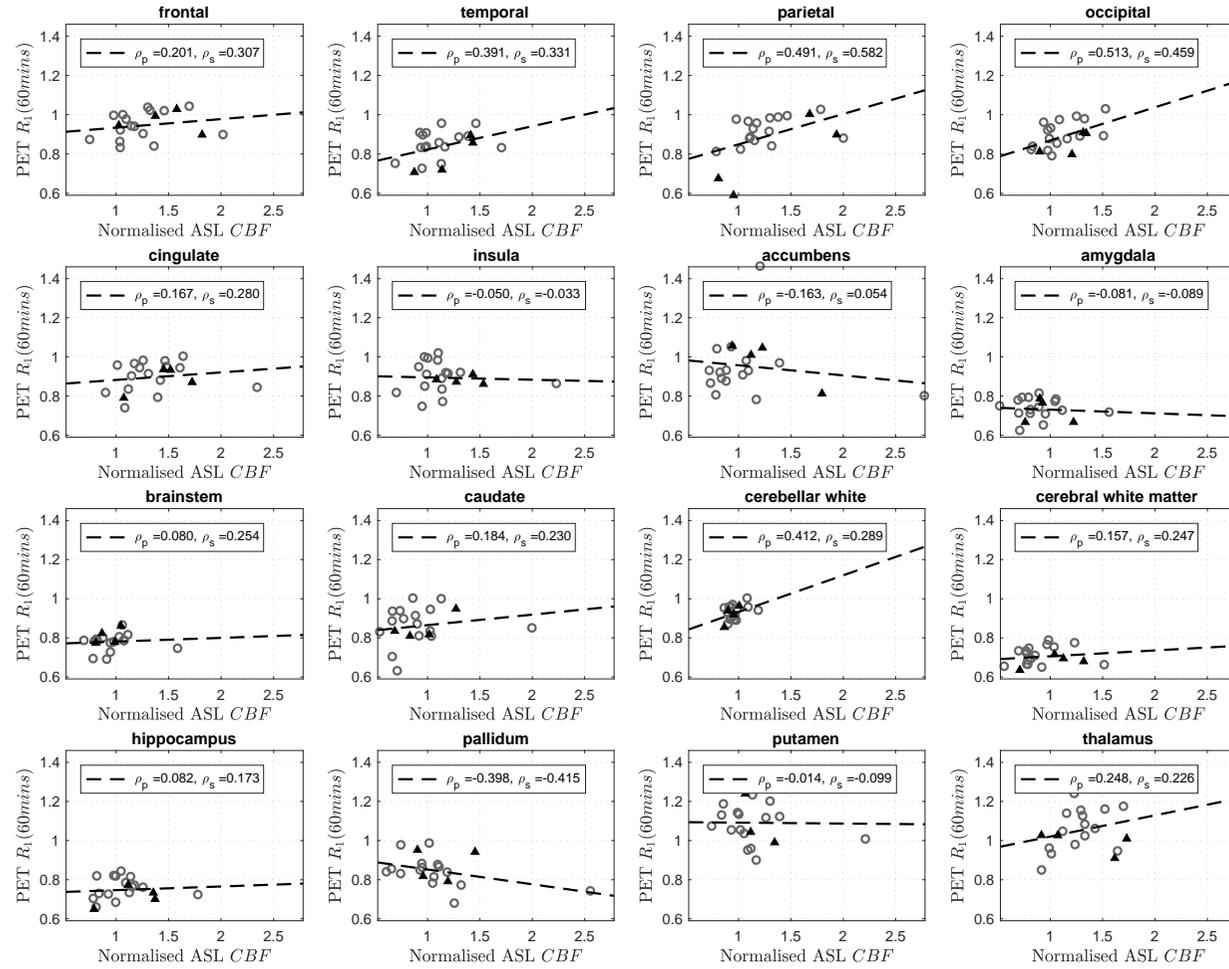
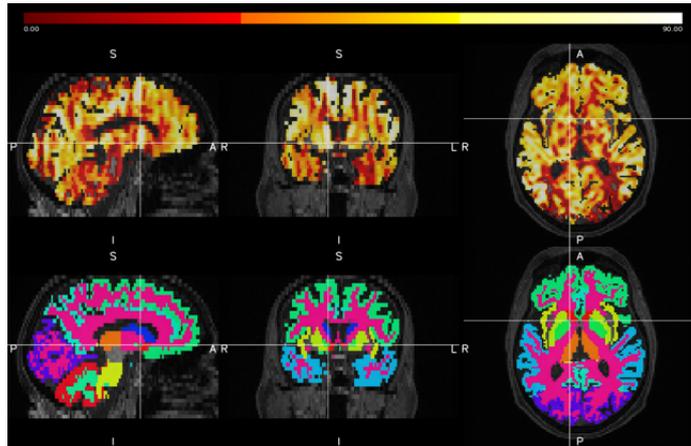


Figure 6.5: Regional plots of ASL-CBF normalised to the reference region against PET- R_1

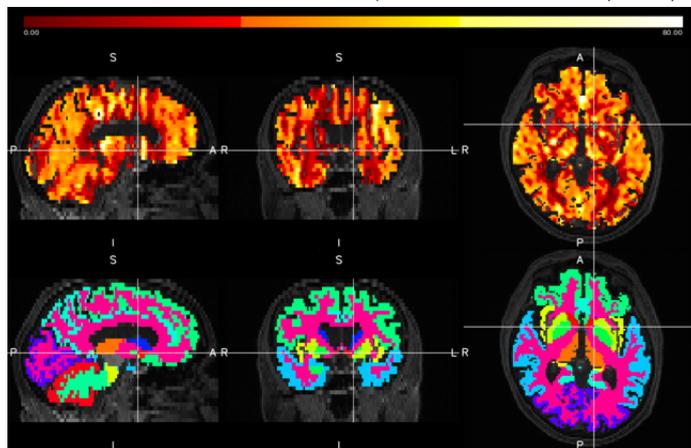
Figure 6.3c shows the multi-linear regression by region, demonstrating the variability in slope and intercept between regions. These differences can mainly be attributed to regional differences in bolus transit times, meaning that the ASL label image is acquired before the bolus exchanges with the tissue, but could also be related to partial volume effects in small structures. Acquiring ASL data with multiple post-labelling delay times can be used to reduce some of this variability and would be particularly helpful when transit time changes are caused by pathology which cannot be modelled.

Figure 6.5 shows the relationship between ASL-CBF and PET- R_1 by region if linear regression is performed on each region individually. This shows that some regions, such as the cortical grey matter regions on the top row, have a strong correlation between the two metrics. The high correlation in cerebellar white matter seen on the third row of figure 6.5 demonstrates that it is possible to generate robust estimations of CBF in white matter using this acquisition protocol. However, for the cerebral white matter the correlation is much lower, indicating that there may be some other influence on this region.

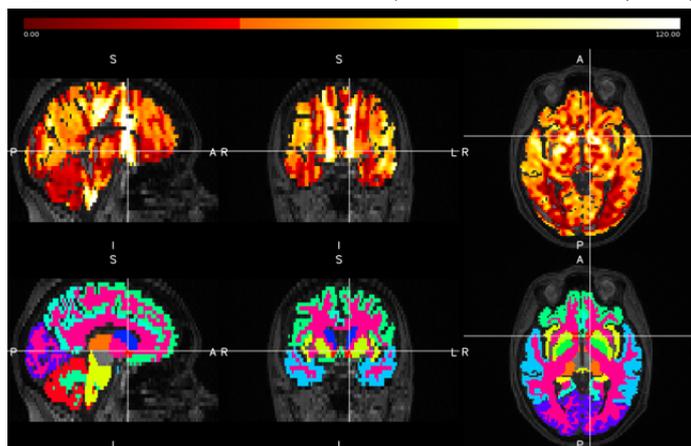
Figure 6.6 demonstrates how the flow artefact (seen in 85% of ASL images reviewed) manifests. This artefact is due to blood flow during an MR acquisition which causes artefacts through 2 different processes; inflow and velocity induced phased effects. Inflow effects refer to the blood flowing through the imaging slice, and therefore for SE sequences a bolus of blood may not receive both the 90° and 180° pulses leading to signal loss. Alternatively, for GE sequences a fresh fully relaxed bolus can enter the imaging slice and receive the 90° pulse, and even if it moves out of the slice it will still contribute to the signal, as the gradients are not slice selective, leading to high intensity signals. Velocity induced phase effects arise due to the movement of the blood spins while the gradients are being applied, leading to continuously changing resonant frequencies, and inducing signal loss and ghosting of the vessels in the image. These effects cause both bright and dark flow artefacts in the control and label images, which propagate through to the CBF estimation and



(a) Typical bright flow artefact (CBF scale 0-90ml/min/100g)



(b) Bright and dark flow artefacts (CBF scale 0-80ml/min/100g)



(c) Outlier subject showing very bright flow artefact (CBF scale 0-120ml/min/100g)

Figure 6.6: Examples of subjects with flow artefacts affecting the CBF estimation. This shows that the affected regions include the accumbens (red), pallidum (green), putamen (yellow), and cerebral white matter (pink).

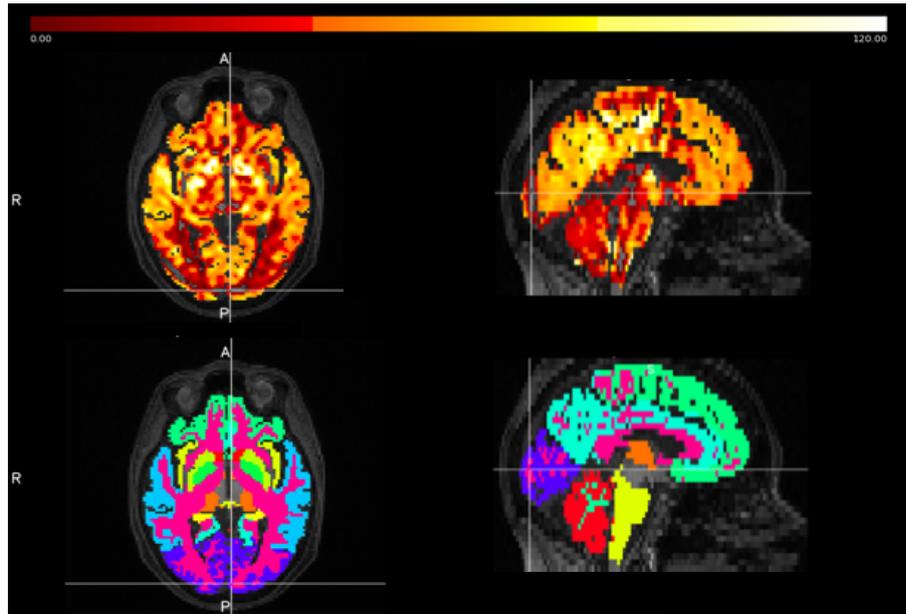


Figure 6.7: Top: CBF map overlaid on T1 showing fat shift artefact, Bottom: parcellation over T1 showing that the artefact appears in the occipital region (purple)

are compounded by patient motion. Figure 6.6a, is a typical example of this artefact with bright streaks cutting vertically through the accumbens (red), pallidum (green), putamen (yellow), and cerebral white matter (pink) on both sides of the brain. An example of a dark flow artefact is shown in figure 6.6b, however this can more easily be masked out in regional CBF estimates as it results in negative values.

The images of the subject shown in figure 6.6c demonstrate a large bright flow artefact, which is causing high CBF estimates. This subject is the outlier point seen in many brain regions in figure 6.6c.

The other recurrent artefact seen in many of the images is fat shift as shown in figure 6.7, which is a form of chemical shift artefact. This type of artefact is caused by atomic electrons shielding the nucleus from the applied magnetic field altering the resonant frequency of the nucleus. The magnitude of the shift depends on the chemical environment of the nucleus and is a particular issue in fat, which is a large molecule causing a large frequency shift compared to water. Whilst the acquisition does include fat suppression, in some cases this is imperfect, and the artefact manifests as a shift of the fat

from the back of the head into the posterior of the brain. This is because the 3D GRASE readout is implemented such that there is an EPI readout in the x-y plane, with the phase direction in y. Due to the small bandwidth per pixel in the phase direction, small shifts in frequency result in large spatial shifts in image space. However, the high correlation in the occipital region shows that the fat shift artefact seen in figure 6.7 does not affect the regional average values. This is because the artefact causes a negative CBF value to be generated which can easily be excluded from regional averaging as a negative flow rate is not physiologically possible.

Whilst some regions show low correlations due to artefacts, the linear model is able to describe most of the variation in the data, and demonstrates strong linear relationships in some tissue regions. Where the correlation between PET- R_1 and ASL-CBF is low, estimates of PET- R_1 derived from ASL-CBF will tend towards an average R_1 value. Based on this the multi-linear regression using region as a covariate was considered to be the best model available to describe the data and to predict PET- R_1 values from ASL-CBF.

6.3.3 Additional covariates

To quantify the subject specific component of the relationship between normalised ASL-CBF and R_1 , which cannot be modelled in a new set of subjects, multi-linear regression with the subject as a covariate was performed. This gave an R^2 value of 0.436 (adjusted $R^2 = 0.358$), indicating that there is some variation between subjects, but that this accounts for less of the variation than the variation between regions. Finally, multi-linear regression using both region and subject as covariates yielded an R^2 value of 0.862 (adjusted $R^2 = 0.824$), showing that most of the variation can be explained by these two parameters. Including amyloid status as an additional covariate did not give any additional information.

6.4 Conclusions

This comparison between ASL-CBF and PET- R_1 demonstrates that the relationship between blood flow and tracer delivery for [^{18}F]-florbetapir can be adequately approximated using a multi-linear regression using brain region as a covariate. This conclusion is supported by the theoretical relationship between blood flow and tracer delivery which was established using the Renkin-Crone model, and supporting literature which compares ASL-CBF and amyloid PET R_1 . When applied to the available clinical data across all regions and subjects, multi-linear regression provided a model which was better able to describe the relationship between the two parameters than a single linear regression.

However, when the relationship was scrutinised on a regional basis artefacts, noise, and partial volume effects lead to errors in CBF estimation in some regions. This could potentially lead to errors in R_1 estimation, and consequently errors in the fitted BP_{ND} when used in the RT-SRTM. However, the model was considered to be suitable for deriving PET- R_1 from ASL-CBF since the regions affected had low gradients, and as such R_1 values derived from these regions would tend towards the population average. The benefit of being able to describe most of the variation in the data ($R^2=0.650$) was considered to outweigh the drawbacks of including low correlation regions. The influence of derived R_1 errors on the estimate of binding potential are investigated in chapter 7.

There is not sufficient evidence that the data used here could support a more complex non-linear model between PET- R_1 and ASL-CBF, as the residual errors of all the linear models follow an approximately normal distribution. Since multi-linear regression using region as a covariate was best able to describe the relationship between PET- R_1 and ASL-CBF, this model was selected for use when applying the RT-SRTM in the following chapter.

Chapter 7

Optimisation of Reduced Acquisition Time PET implementation

In this chapter the reduced time SRTM (RT-SRTM) framework outlined in chapter 5 is optimised; two techniques for extrapolating the reference region curve, C_R , are compared (*scaled mean C_R* method and *PCA C_R* method), and the optimal timing window ($t_s:t_e$) for the acquisition of the PET data is assessed. Finally the optimised RT-SRTM is applied, using multi-linear regression evaluated in chapter 6 to estimate R_1 from ASL.

7.1 Introduction

Chapter 5 described the three stages required for the RT-SRTM; 1) estimating relative tracer delivery (R_1) from ASL-CBF, 2) extrapolating reference region curve (C_R), and 3) fitting the SRTM with CBF-derived R_1 and extrapolated $C_R(t)$. Chapter 6 looked at the first step in detail, where it was determined that multi-linear regression using brain region as a covariate was the most accurate method for estimating PET- R_1 from ASL-CBF. However the propagation of errors in R_1 estimation to BP_{ND} has not been explored, nor have the other two steps have been optimised. In this chapter I perform sensitivity analysis by simulating data to see how accurate the estimate of R_1 needs to be to ensure

a reliable BP_{ND} estimate.

In chapter 5, step 2 was carried out by scaling a population mean reference region curve to the available subject data to extrapolate the missing time points. This approach assumes that the subject's reference region curve, $C_R(t)$ can be described using a population based arterial input function with a subject specific scaling factor, and a population based washout rate, k'_2 . However, this may not be the case, particularly in diseased populations who may have pathological changes in blood flow which will affect the arterial input function and the tracer washout. Consequently, a more flexible method which can represent a wider range of reference region curve shapes would be required to apply the technique to a heterogeneous patient population.

For step 3, where the RT-SRTM is fitted to the measured PET data, $t = 30, 60$ was selected as the later part of the signal contains information about the binding and washout of the tracer, which are the parameters which need to be estimated. However, the choice of acquisition timing window also determines which part of the reference region curve is measured and will consequently affect its extrapolation in step 2). Therefore in this chapter I also look at which part of the PET data to acquire.

7.2 Methods

7.2.1 Data

The imaging dataset is described in section 6.2.1, and the split between optimisation and testing of the technique is shown in figure 7.1. The optimisation of the relationship between ASL-CBF and PET- R_1 was derived in chapter 6 and the results of this optimisation, using the multi-linear regression method, is used here. Analysis was performed on a regional basis using the tissue regions defined in section 5.2.

7.2.2 Extrapolation of PET reference region, C_R

As described in section 5.2.2, C_R must be known from injection, at $t = 0$, to the end of the scan, $t = t_e$, where $t_s, t_e = 0, 60$ minutes for the gold standard.

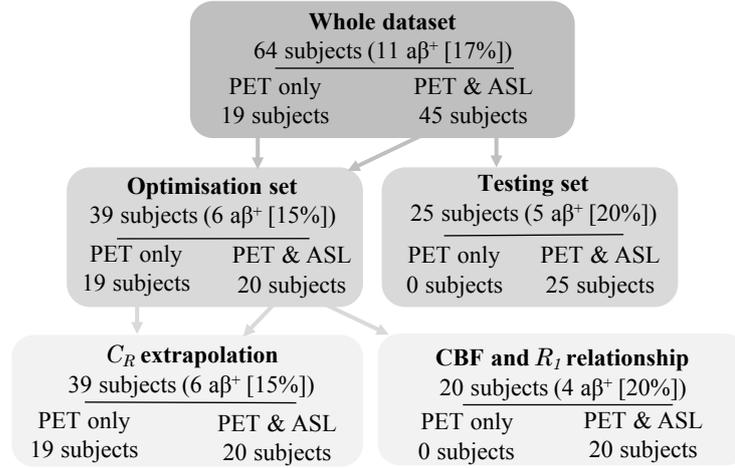


Figure 7.1: Division of data between testing and optimisation

However, when reducing the acquisition time, C_R is only measured between t_s and t_e where $t_s \neq 0$, therefore a strategy is required to extrapolate the missing data. Here we define $C_R(t)$ as a vector containing the reference region tracer concentration over time, $C_R^{pop}(t)$ as a matrix containing the reference region concentration for a population of subjects, and $\bar{C}_R^{pop}(t)$ as a vector containing the mean population tracer concentration. For clarity, the acquisition time, t , is expressed as a discrete variable, as the dynamic data are binned into frames.

In this work two different approaches are evaluated to estimate the whole reference input vector $C_{R_{[t=0,t_e]}}$ for an unseen subject's $C_{R_{[t=t_s,t_e]}}$. Both techniques make use of a population of subjects for which the full $C_{R_{[t=0,t_e]}}$ was measured.

The first method, referred to as the *scaled mean C_R* method and introduced in chapter 5 (Scott et al., 2016), scales the population average reference region curve, $\bar{C}_{R_{[t=0,t_e]}}^{pop}$, as:

$$C_{R_{[t=0,t_e]}} \approx \alpha \bar{C}_{R_{[t=0,t_e]}}^{pop} \quad , \quad (7.1)$$

where α is a subject specific scaling factor determined through a least squares fit of $\bar{C}_{R_{[t=t_s,t_e]}}^{pop}$ to $C_{R_{[t=t_s,t_e]}}$. This results in an individual estimate of $C_{R_{[t=0,t_e]}}$ to be used in the RT-SRTM.

The second method employs statistical shape modelling to build a model of the variation in C_R within the population of subjects (Cootes et al., 1995). This requires principal component analysis (PCA) of a set of subjects to determine the M components, $\mathbf{U} = [\vec{u}_1, \dots, \vec{u}_M]'$ where $\vec{u}_i = [u_1, \dots, u_{frames}]$, and is therefore referred to as the *PCA C_R* method.

The $C_{R[t=0,t_e]}$ of each subject in the set can be expressed as the mean population reference curve, $\bar{C}_{R[t=0,t_e]}^{pop}$, plus a linear combination of the weighted principal components, where the weight of the i^{th} mode, \vec{u}_i , is w_i . A subset of the components L , where $L < M$, which describe the majority of the variation are selected. An unseen reference region curve $C_{R[t=t_s,t_e]}$ can then be fitted by adjusting the weights, as per equation (7.2). The same weights and modes can then be used with $C_{R[t=0,t_e]}^{pop}$ to generate $C_{R[t=0,t_e]}$.

$$C_{R[t=t_s,t_e]} \approx \bar{C}_{R[t=t_s,t_e]}^{pop} + \sum_{i=1}^L w_i \vec{u}_i \quad (7.2)$$

7.2.3 Optimisation of C_R extrapolation and PET acquisition window

The *scaled mean C_R* and *PCA C_R* methods were tested using a leave-one-out strategy, whereby $C_{R[t=0,t_e]}^{pop}$ contained all subjects except the one being fitted.

For the *scaled mean C_R* method, $\bar{C}_{R[t=0,t_e]}^{pop}$ was calculated by scaling each reference curve to match the peak value, then calculating the mean for each frame. $\bar{C}_{R[t=0,t_e]}^{pop}$ was then scaled using a least squares fit to the subject data between $t = t_s, t_e$.

For the *PCA C_R* method the weight for each mode of variation, w_i , must be optimised. Equation 7.2 can be solved through linear optimisation of the weights; however, preliminary work showed that when this is applied to an incomplete dataset, i.e. when only the datapoints between t_s and t_e are available, the problem is insufficiently constrained and the fit is poor. Consequently, non-linear optimisation of the weights was employed.

Following PCA of $C_{R[t=0,t_e]}^{pop}$, each mode of variation has a weight for

each subject in the population set. The standard deviation for each mode is the square root of the variance given by the eigenvalues of the covariance matrix, which is computed within PCA. The standard deviation of the weights for each mode can be used to define upper and lower bounds for the weight when fitting an unseen $C_R(t = t_s, t_e)$. Assuming that C_R^{pop} is representative of the general population, it is unlikely that an unseen C_R would deviate by more than ± 3 times the standard deviation of the population, suggesting that this could be used to define reasonable limits (Cootes et al., 1995). However when fitting reduced time data a smaller range may help to constrain the optimisation.

To find the optimal values for the upper and lower bounds of the weights, 30 minutes of data were fitted multiple times with different bounds which ranged between ± 0.5 and 3 times the standard deviation for each mode of variation. This was estimated for a single time window ($t = 30, 60$ minutes), and the three best performing parameters were then compared across timing windows.

The number of components, L , was calculated by performing PCA on the reference region curves, $C_R^{pop}(t)$, and calculating the percentage of the covariance explained using different numbers of components. The maximum value of L , L_{max} , was selected based on the smallest number that could explain $\geq 99.9\%$ of the variation. Optimisation was performed using $L = 1, \dots, L_{max}$.

To determine the optimal PET acquisition window ($t = t_s, t_e$), four different 30 minute windows were tested, covering the full 60 minutes of available data. No ASL data were used in this section to ensure that errors introduced by the estimation of R_1 using CBF do not confound the results. To evaluate the influence of the timing window on estimated BP_{ND} , a fit was performed using the true subject C_R , and the gold standard R_1^* values. This was then compared to the fits using the derived reference region curves and gold standard R_1^* values.

$SUVR-1$ was calculated at five 10 minute acquisition windows, starting 30 minutes post injection, to cover the recommended time range. The $SUVR-1$

values estimated were then compared to the gold standard BP_{ND} , and the average mean square error (MSE) and bias (mean error) was calculated across the 25 testing set subjects.

7.2.4 Sensitivity analysis of RT-SRTM

Sensitivity analysis was performed with simulated dynamic PET data to assess the influence of errors in R_1 estimation from ASL-CBF. Regional parameters from 1 positive subject (from the young onset Alzheimer's disease study) were used to generate time activity curves for $t=0,60$ mins using the SRTM and a mean reference region curve. The testing set subjects were used to determine the range of noise and the error in R_1 estimation used in the simulation.

$$SD(t_i) = SF \sqrt{\frac{e^{\lambda t_i} C_T(t_i)}{\Delta t_i}} \quad (7.3)$$

A range of noise levels were added to the simulated curves using equation 7.3 (Ichise et al., 2003), based on Gaussian noise, to represent different region sizes and statistics. To determine the noise range to add to the simulated data, the standard deviation of the residual error in the gold standard fit for $t=40,60$ mins for all testing set subjects was calculated. 200 noise realisations were generated for each noise percentage, and 15 noise levels were used.

The percentage error in R_1 estimation using multi-linear regression on the ASL data was found for all testing set subjects. 15 error percentages were selected to cover the range of those found.

Finally, the simulated curves were fitted over the optimised 30 minute timing window using the RT-SRTM, where R_1 was fixed by the simulation. Estimates of BP_{ND} were compared to the ground truth simulated values and $BP_{ND}+1$ was used to calculate percentage errors for consistency with existing literature using SUVR.

7.3 Results

7.3.1 Extrapolation of PET reference region, C_R

7.3.1.1 Optimisation of *PCA* C_R method

Table 7.1 shows the percentage of the variation described as the number of components are increased for the population of 39 subjects, $C_R^{pop}(0,60)$. This demonstrates that six principal components are required to describe 99.9% of the variation within the data, therefore a maximum of six components are used in the optimisation ($L \leq 6$).

| Number of components (L) | Variation described (%) |
|------------------------------|-------------------------|
| 1 | 76.4 |
| 2 | 95.0 |
| 3 | 97.8 |
| 4 | 99.1 |
| 5 | 99.6 |
| 6 | 99.9 |

Table 7.1: The percentage of the variation explained using increasing number of principal components following PCA on the $C_R(t = 0,60)$ of 39 optimisation set subjects

In addition to the number of principal components, the upper and lower bounds of the weights must be defined for the optimisation. Figure 7.2 shows the MSE in the fit of $C_R(t = 30,60)$ using leave-one-out cross validation and averaged across subjects. This demonstrates that constraining the upper and lower bounds of the weights reduces the error when fitting to data with missing timepoints. Figure 7.2 shows that the error in the estimate of C_R is largest when both the number of components and the size of the weight bound limits are large (top right), showing that when the model is minimally constrained then it cannot accurately estimate the missing data. Figure 7.3 (created using the same data as figure 7.2) shows how the error varies by subject for a given number of principle components. This shows that using just one principal component the results are more stable across the range of weight bounds, although a lower error can be achieved using more principal components.

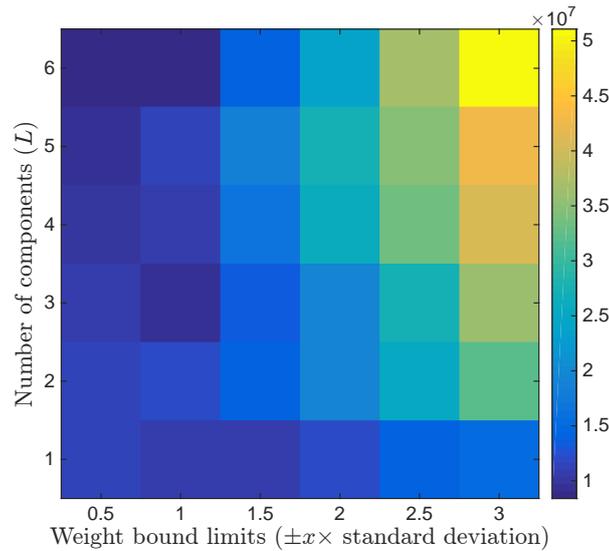


Figure 7.2: MSE in the fit of $C_R(t = 30, 60)$ using the *PCA* C_R method when optimising the number of components used and the upper and lower bounds for the weights using leave-one-out analysis on 39 optimisation set subjects

Overall the combinations of components and weight bounds with the lowest error are $L = 6$ with either ± 0.5 or 1 times the standard deviation for the bounds, or $L = 3$ with ± 1 times the standard deviation for the bounds. These three combinations give a similar MSE for $t = 30, 60$, however figure 7.4 shows that $L = 6$ with 1 times the standard deviation for the bounds performs consistently better across different timing windows, and therefore this was used for comparison with the *scaled mean* C_R method in the following sections. It is worth noting, however, that using three principal components with a weight bound of ± 1 standard deviation also performs well. This could be useful if fitting fewer PET frames as the number of frames determines the maximum number of principal components to avoid an underdetermined problem.

For the *PCA* C_R method, 6 modes of variation were required to explain 99.9% of the variation between subjects. For these 6 modes of variation, the upper and lower bounds of the weights must be defined. Figure 7.3 shows how the mean square error in the fit of $C_R(t = 30 : 60)$ changes as the weights can vary between ± 0.5 and 3 times the standard deviation for each mode of variation. As the range increases the model is better able to model variation,

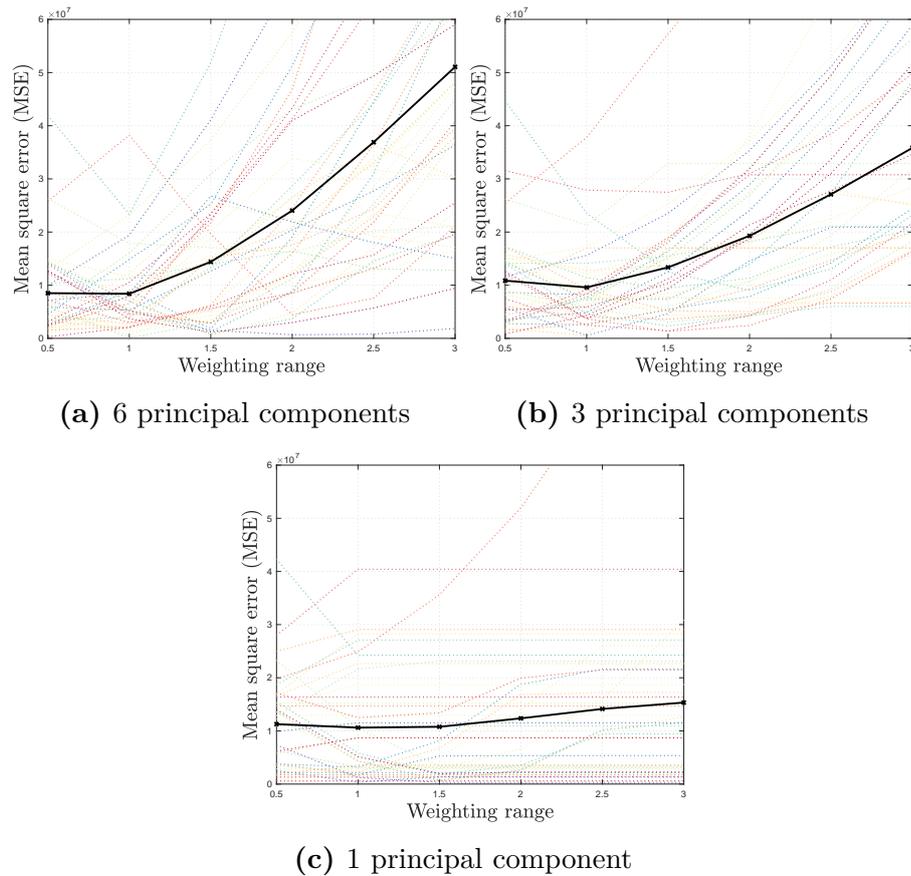


Figure 7.3: Plot of mean square error in the fit of C_R using the $PCA C_R$ method and different ranges of weighting factors for 24 subjects (30:60 mins PET data). Upper and lower bound of the weighting factor are \pm weighting range times standard deviation. Coloured lines represent individual subjects and the black line is the average across all subjects.

however due to the limited acquisition window, the fit is not sufficiently constrained leading to implausible estimates of C_R . From this experiment, the optimal range was taken to be ± 1 standard deviation for each mode, as this produced the lowest mean square error.

7.3.1.2 Comparison of $PCA C_R$ and *scaled mean C_R* methods:

The boxplot in figure 7.5, which summarises the ME across all subjects, shows that the $PCA C_R$ appears to perform better than the *scaled mean C_R* method for all time windows except $t=10,40$ with a small range of errors and a generally lower bias. However, neither the difference in MSE nor ME reached statistical significance (MSE: $p \geq 0.3218$, ME: $p \geq 0.2015$). The influence of this error on

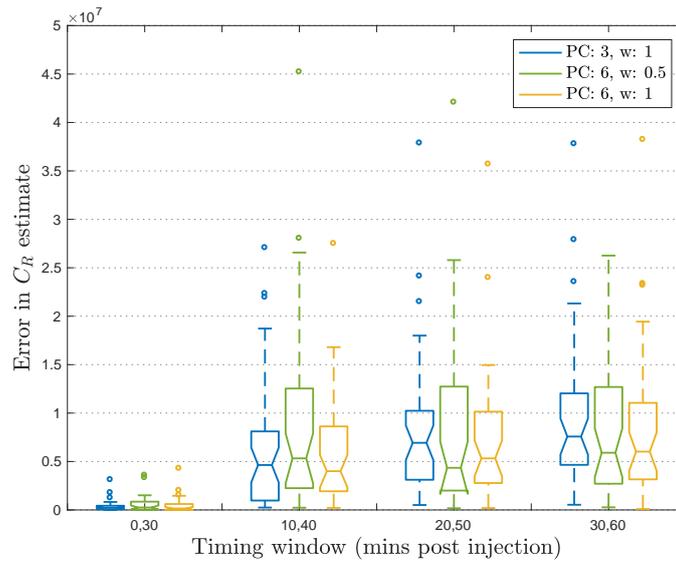


Figure 7.4: Mean error for in the fit of $C_R(t = 30, 60)$ using the *PCA* C_R method across timing windows with difference constrains on a leave-one-out analysis on 39 optimisation set subjects

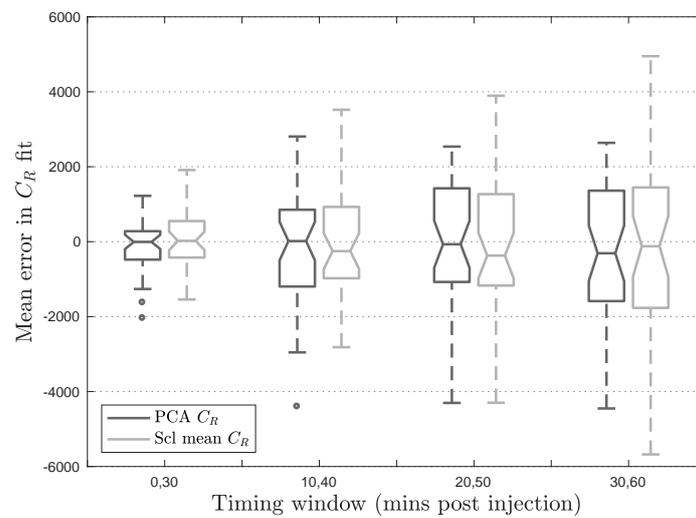


Figure 7.5: ME in estimated C_R using *PCA* C_R method and *scaled mean* C_R method compared to measured C_R calculated using leave-one-out cross validation in 39 optimisation set subjects.

the estimation of BP_{ND} at different acquisition windows is assessed in the next section.

7.3.2 Optimisation of acquisition time window

The MSE and ME in the estimation of BP_{ND} using different data acquisition windows are shown in table 7.2. Extrapolation of C_R is not strictly necessary for $t = t_e, 60$ as the basis functions can be generated provided that the information starts from injection at $t = 0$. However, estimating the $t = t_e, 60$ timepoints reduces the error in the estimate of BP_{ND} as it is closer to the gold standard. For example, for $t = 0, 30$ using only the measured data for C_R yields an MSE = 0.1960 and ME = -0.3704 which is much higher than that from either extrapolation method shown in table 7.2. Consequently, in all cases C_R was extrapolate to cover $t = 0, 60$.

In table 7.2 the true C_R column uses the full measured C_R ($t = 0, 60$), therefore errors are introduced purely due to the limited number of datapoints available. When the later frames are omitted and only $t = 0, 30$ minutes of data are acquired, large errors are introduced as there is little information about the late phase which contains the signal relating to target binding. As the time window is shifted later, the MSE and ME are reduced.

However, the results in table 7.2 also show that there is little to be gained by including data acquired more than 50 minutes post injection as the error increases. This is because the signal has plateaued by this point so, for a fixed 30-minute window, a better fit can be obtained by including some of the earlier data where the tracer concentration changes more rapidly over time. This is further illustrated in figure 7.6, which shows the smallest distribution of errors

| Time window ($t = t_s, t_e$) | True C_R | | PCA C_R | | Mean C_R | |
|-----------------------------------|---------------|---------------|---------------|---------------|---------------|----------------|
| | MSE | ME | MSE | ME | MSE | ME |
| 0,30 | 0.0089 | 0.0303 | 0.0085 | 0.0299 | 0.0085 | 0.0299 |
| 10,40 | 0.0035 | 0.0202 | 0.0036 | 0.0202 | 0.0035 | 0.0211 |
| 20,50 | 0.0008 | 0.0050 | 0.0010 | 0.0041 | 0.0012 | 0.0088 |
| 30,60 | 0.0030 | -0.0084 | 0.0032 | -0.0090 | 0.0034 | -0.0050 |

Table 7.2: MSE and ME between gold standard BP_{ND}^* and BP_{ND} at different 30-minute acquisition windows averaged across 16 regions and 39 optimisation set subjects. True C_R uses the true reference region curve, PCA C_R and mean C_R extrapolate the reference region curve using the *PCA C_R* and *scaled mean C_R* methods, respectively. All methods used the gold standard R_1^* .

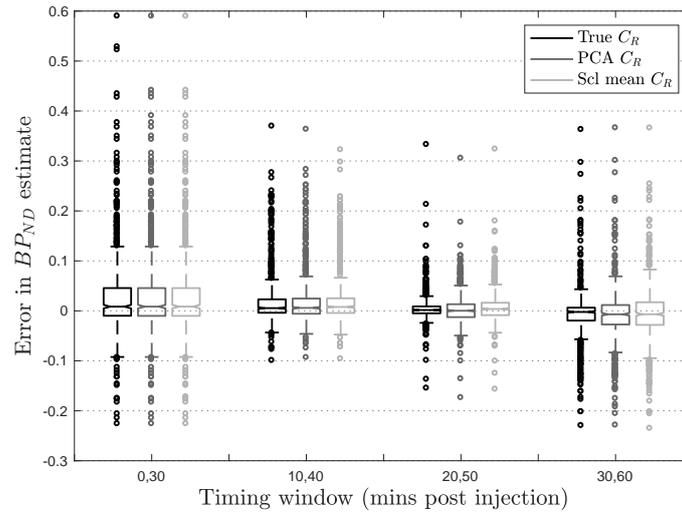


Figure 7.6: Error in BP_{ND} estimates across different timing windows using different estimates of C_R calculated using leave-one-out cross validation in 39 optimisation set subjects..

in BP_{ND} at $t = 20, 50$.

For the two techniques which extrapolate C_R from the available time window, additional variability is introduced into the BP_{ND} estimates due to errors in the extrapolation, which is reflected by the higher MSE in table 7.2. For both the $PCA C_R$ and the *scaled mean* C_R method, the early phase of the data becomes more important, as much of the variation between the C_R of subjects is contained in these frames. This can be seen in figure 7.5, as for $t = 0, 30$ the error in C_R estimation is at a minimum, and increases as the acquisition window shifts later. When the early frames are included, the error in BP_{ND} is very similar to using the true C_R , whereas when only the late frames are used, the error in BP_{ND} for the extrapolation methods increases relative to the true C_R , figure 7.6. However, this is balanced out by the fact that later data is required to get the best fit to C_T and hence BP_{ND} . Table 7.2 shows that the optimal timing window for this technique is $t = 20, 50$ minutes post injection as it yields the lowest MSE and ME.

The $PCA C_R$ method produces a consistently lower MSE which was statistically significantly lower than that obtained using the *scaled mean* C_R method

| Time window ($t_s:t_e$) | | | | | | $SUVR-1$ | |
|---------------------------|-------|-------|-------|-------|-------|---------------|---------------|
| 30:35 | 35:40 | 40:45 | 45:50 | 50:55 | 55:60 | MSE | Bias |
| ● | ● | ○ | ○ | ○ | ○ | 0.0269 | 0.1143 |
| ○ | ● | ● | ○ | ○ | ○ | 0.0301 | 0.1288 |
| ○ | ○ | ● | ● | ○ | ○ | 0.0312 | 0.1318 |
| ○ | ○ | ○ | ● | ● | ○ | 0.0301 | 0.1275 |
| ○ | ○ | ○ | ○ | ● | ● | 0.0253 | 0.1108 |

Table 7.3: Mean square error (MSE) and bias (mean error) between the gold standard BP_{ND}^* and $SUVR-1$ at different acquisition windows across 25 testing set subjects.

for $t = 20, 50$ ($p = 0.002$). This is due to the increased flexibility in this method as the model is built from the many population subjects, rather than using a population mean with the assumption of a population washout rate, k_2 . This allows the $PCA C_R$ method to better describe the variation in unseen C_R shapes. The $PCA C_R$ method also produces BP_{ND} estimates with lower bias at $t = 20, 50$, as demonstrated by the median lines in figure 7.6 and the ME in table 7.2 ($p < 0.001$).

Due to the lower MSE and ME the $PCA C_R$ method with the $t = 20, 50$ minute timing window was selected for the full implementation of the RT-SRTM in subsequent sections.

7.3.3 Optimisation of SUVR acquisition window

Table 7.3 shows the MSE and ME between the gold standard BP_{ND}^* and $SUVR-1$ at different acquisition windows across 25 testing set subjects. As anticipated, the optimal available timing window with the lowest bias and error was found to be $t = 50, 60$ minutes post injection, in concordance with Cselényi and Farde (2015). This is verified by figure 7.7, which demonstrates that SUVR reaches a plateau at this timepoint. This indicates that steady state has been reached, which is the condition under which SUVR can be used to approximate BP_{ND} . Therefore $SUVR-1$ was calculated at $t = 50, 60$ for comparison with the proposed method.

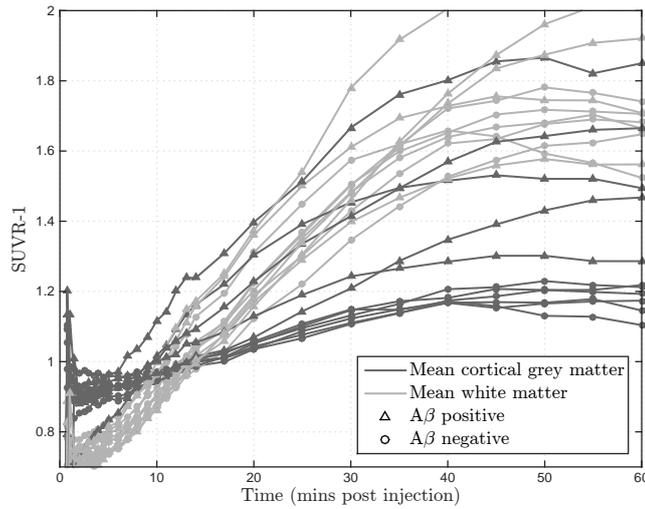


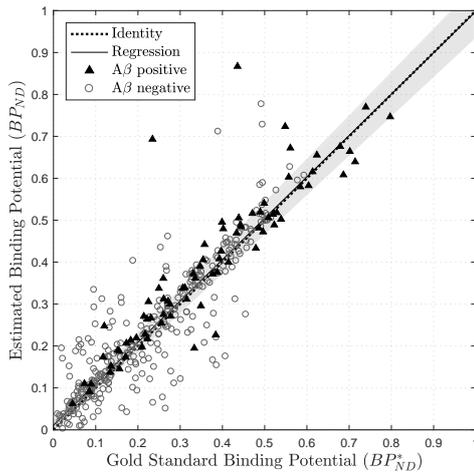
Figure 7.7: Plot of average SUVR against time post injection for a representative subset of 10 testing set subjects showing that a plateau is reached at approximately 50 minutes post injection, using cerebellar grey matter as a reference region.

7.3.4 Comparison of proposed RT-SRTM with gold standard

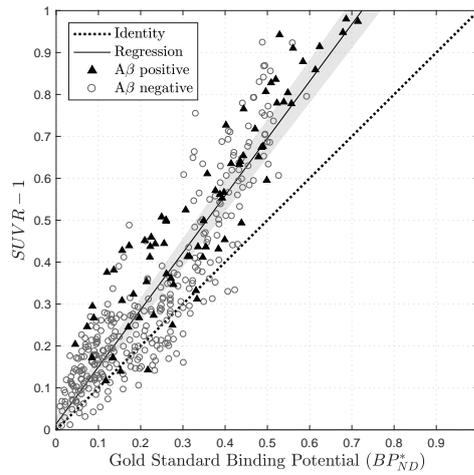
Figure 7.8a shows BP_{ND} , estimated using the RT-SRTM with $t = 20, 50$ minutes of data plotted against the gold standard BP_{ND}^* . Linear regression of the data shows that the RT-SRTM method offers a good approximation of the gold standard as it closely follows the line of identity (dashed), which is within the 95% confidence interval (CI) of the regression (shaded). The linear correlation between the two estimates was tested using the Pearson correlation coefficient which demonstrated a high, statistically significant result ($r = 0.896$, $p < 0.001$, 95% CI [0.875, 0.914]).

7.3.5 Comparison of proposed RT-SRTM with $SUVR-1$

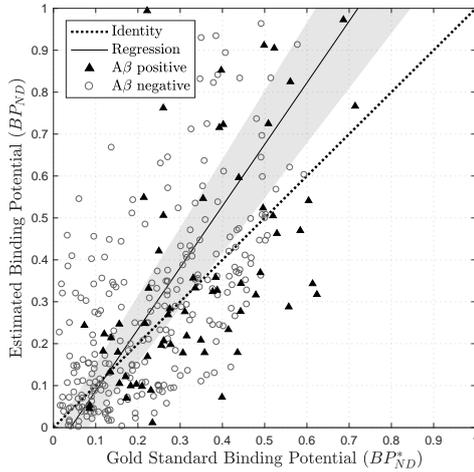
Figure 7.8b shows the amyloid burden estimates generated for $SUVR-1$ ($t = 50, 60$ minutes). Whilst the correlation between $SUVR-1$ and BP_{ND}^* is evident, a positive bias is shown as $SUVR-1$ overestimates the binding potential at higher values. This is due to the fact that the target and reference tissue concentrations reach equilibrium with blood plasma at different points depending on tracer binding, as has been explored in detail in the liter-



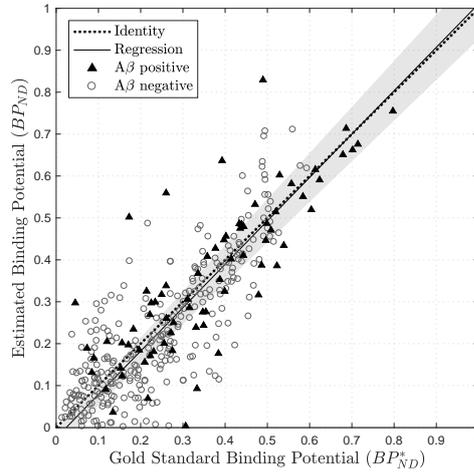
(a) RT-SRTM where $t = 20, 50$ minutes (ASL derived R_1)



(b) $SUVR-1$ where $t = 50, 60$ minutes (static scan)



(c) SRTM where $t = 0, 30$ minutes (PET data only, no C_R extrapolation)



(d) RT-SRTM where $t = 20, 50$ minutes (PET data only, extrapolated C_R)

Figure 7.8: Estimated amyloid burden against the gold standard value calculated using full PET time series for 25 testing set subjects. The grey shaded region covers the 95% confidence interval in the regression.

ature (Slifstein, 2008). The ME quantifies the bias between the estimates and the gold standard which is 0.1038 for $SUVR-1$, indicative of the systematic overestimation, compared to 0.0079 for the RT-SRTM method ($p < 0.001$). The RT-SRTM method also has a lower MSE (0.0066 compared to 0.0235 for $SUVR-1$, $p < 0.001$), showing that overall this technique estimates BP_{ND}^* more accurately than the simplified technique.

This overestimation in $SUVR-1$ is likely to be a combination of a linear systematic error in the estimation, as well as the influence of blood flow. Systematic error could result from estimating $SUVR-1$ when not at steady-state, and could potentially be accounted for using a population correction factor.

To determine the influence of blood flow on the estimation of target density, the correlation between estimated target density and tracer delivery was calculated, as in Cselényi and Farde (2015). Spearman’s correlation (ρ) between R_1^* and the estimates was calculated, as the relationship is in theory non-linear. For the gold standard BP_{ND}^* and BP_{ND} using RT-SRTM, there was no significant correlation with R_1^* ($p = 0.336$ and 0.106 respectively). However, for $SUVR-1$ there was a significant negative correlation ($\rho = -0.226$, $p < 0.001$), which suggests that the proposed RT-SRTM method may be more robust to changes in blood flow than $SUVR-1$.

7.3.6 Comparison of proposed RT-SRTM with short acquisition time PET

Pharmacokinetic modelling can be applied to reduced acquisition time PET data without incorporating ASL data in several ways, e.g.: i) fitting the first 30 minutes of PET data only using the standard SRTM in equation 3.20, ii) extrapolating C_R as in RT-SRTM, but R_1 is estimated from the PET data, iii) the ‘coffee break’ or ‘dual time-point’ protocol where early (from injection) and late PET data are acquired with a break in-between and the missing data is interpolated (Bullich et al., 2018; Heeman et al., 2019). In addition to the errors introduced by the interpolation, the ‘coffee break’ protocol requires the registration of the pre- and post-break data for pharmacokinetic modelling, scatter and attenuation correction, which is non-trivial. This approach is also generally applied to tracers which require scans of ≥ 90 minutes to justify the inconvenience of scanning the patient twice. As a result at the time of writing there are no published papers using this method with Florbetapir data, and since the result cannot be fully simulated from the data acquired here, it is not included in this comparison.

For the first method, i), figure 7.8c compares the estimation of BP_{ND} using PET data for $t = 0, 30$ minutes to the gold standard. It is evident that the absence of late-phase data leads to a high error in the estimate, which is significantly higher than using the RT-SRTM (MSE= 0.0813, $p < 0.001$; ME = 0.0569, $p < 0.001$).

For the second method, ii), where BP_{ND} is estimated using 30 minutes of PET data and $PCA C_R$ extrapolation, the timing window $t = 20, 50$ was selected by calculating the MSE for all the timing windows used in table 7.2 and $t = 20, 50$ yielded the lowest value. Figure 7.8d shows a significantly lower MSE for this method compared to using the first 30 minutes (MSE = 0.0121, $p < 0.001$). This method also outperforms $SUVR-1$ with a lower MSE ($p < 0.001$) and ME ($p < 0.001$), however it is significantly correlated with the gold standard R_1^* ($p = 0.004$).

Comparison of this PET only method in figure 7.8d with the proposed RT-SRTM including ASL derived R_1 estimates (figure 7.8a) shows that the additional CBF information improves the estimate of BP_{ND} , yielding a significantly lower MSE ($p = 0.028$), ME ($p < 0.001$) and variance (two-tailed F-test $p < 0.001$). This can be seen in figure 7.8a where the points are more tightly clustered around the line of identity when using the proposed RT-SRTM compared to figure 7.8d using PET data only.

7.3.7 Sensitivity analysis of RT-SRTM

Table 7.4 shows the percentage error in the estimated $BP_{ND}+1$ for a given data noise level and R_1 error. This shows that, as expected, the error in the estimate of $BP_{ND}+1$ increases as the error in R_1 increases. However, it also demonstrates that the noise in the time activity curve has a significant influence on the estimate.

When evaluating all regions across all testing set subjects, the mean R_1 error was found to be 6.0%, with 85% of the estimates below 10% R_1 error. The maximum noise in the regional time activity curves was found to be 3.4%, with an average of 1.7%. This shows that, for a typical region, the error

in BP_{ND+1} is less than 1.4%. For a noisy region with a 10% error in R_1 estimate this increases to around 2.5%. However, if the current methodology was extended to voxel-wise analysis where the noise may be up to 10%, and the error in R_1 could be higher, the error in BP_{ND+1} would increase dramatically, up to 10% or more.

In Cselényi and Farde (2015) the blood flow dependent component in SUVR for the Alzheimer's disease subjects was found to be 0.03. Table 7.5 shows that the mean absolute error for the RT-SRTM method is generally far below this value for regional analysis (noise < 3.4%, R_1 error < 10%), and would be much lower for a mean cortical grey matter region as used by Cselényi and Farde (2015). This supports the conclusion that, on a regional level, the RT-SRTM could potentially be more robust for longitudinal analysis than SUVR. However, if the methodology were to be extended to voxel-wise analysis, a lower R_1 error would be required due to the increased noise in the PET data which increases the uncertainty in the model fit.

| % R_1 error | % noise in time activity curve | | | | | | | | | | | | | | |
|------------------|--------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 0.5 | 0.6 | 0.8 | 1.0 | 1.2 | 1.5 | 1.8 | 2.2 | 2.8 | 3.4 | 4.2 | 5.3 | 6.5 | 8.1 | 10.0 |
| 0.0 | 0.174 | 0.208 | 0.278 | 0.360 | 0.463 | 0.617 | 0.855 | 1.108 | 1.357 | 1.665 | 1.936 | 2.511 | 3.166 | 3.876 | 4.715 |
| 0.2 | 0.177 | 0.211 | 0.279 | 0.361 | 0.461 | 0.615 | 0.854 | 1.108 | 1.356 | 1.668 | 1.936 | 2.512 | 3.170 | 3.877 | 4.715 |
| 0.3 | 0.180 | 0.213 | 0.280 | 0.363 | 0.465 | 0.616 | 0.855 | 1.107 | 1.353 | 1.668 | 1.936 | 2.514 | 3.171 | 3.876 | 4.716 |
| 0.4 | 0.183 | 0.216 | 0.282 | 0.365 | 0.466 | 0.617 | 0.856 | 1.105 | 1.352 | 1.669 | 1.937 | 2.515 | 3.172 | 3.877 | 4.719 |
| 0.6 | 0.192 | 0.224 | 0.289 | 0.370 | 0.471 | 0.620 | 0.859 | 1.108 | 1.354 | 1.673 | 1.937 | 2.516 | 3.174 | 3.877 | 4.718 |
| 1.0 | 0.221 | 0.251 | 0.310 | 0.390 | 0.484 | 0.633 | 0.867 | 1.118 | 1.362 | 1.677 | 1.944 | 2.514 | 3.170 | 3.883 | 4.721 |
| 1.6 | 0.286 | 0.309 | 0.362 | 0.432 | 0.521 | 0.662 | 0.889 | 1.141 | 1.377 | 1.687 | 1.952 | 2.525 | 3.178 | 3.901 | 4.729 |
| 2.6 | 0.425 | 0.442 | 0.479 | 0.537 | 0.611 | 0.748 | 0.946 | 1.193 | 1.419 | 1.724 | 1.977 | 2.544 | 3.195 | 3.908 | 4.744 |
| 4.1 | 0.660 | 0.671 | 0.702 | 0.739 | 0.799 | 0.919 | 1.085 | 1.309 | 1.519 | 1.801 | 2.047 | 2.603 | 3.235 | 3.945 | 4.772 |
| 6.5 | 1.054 | 1.055 | 1.079 | 1.109 | 1.142 | 1.248 | 1.389 | 1.565 | 1.746 | 1.992 | 2.220 | 2.736 | 3.346 | 4.042 | 4.859 |
| 10.3 | 1.679 | 1.678 | 1.684 | 1.714 | 1.753 | 1.837 | 1.940 | 2.089 | 2.225 | 2.419 | 2.589 | 3.051 | 3.617 | 4.271 | 5.063 |
| 16.4 | 2.657 | 2.651 | 2.657 | 2.688 | 2.723 | 2.801 | 2.891 | 2.991 | 3.110 | 3.261 | 3.385 | 3.743 | 4.231 | 4.865 | 5.594 |
| 26.1 | 4.187 | 4.188 | 4.209 | 4.238 | 4.274 | 4.350 | 4.421 | 4.509 | 4.615 | 4.733 | 4.833 | 5.160 | 5.566 | 6.153 | 6.756 |
| 41.5 | 6.732 | 6.727 | 6.735 | 6.771 | 6.812 | 6.871 | 6.951 | 7.038 | 7.117 | 7.307 | 7.438 | 7.810 | 8.167 | 8.653 | 9.091 |
| 66.0 | 11.40 | 11.40 | 11.41 | 11.47 | 11.48 | 11.56 | 11.61 | 11.59 | 11.61 | 11.63 | 11.71 | 11.72 | 11.87 | 11.99 | 12.12 |
| SRTM | 0.088 | 0.107 | 0.149 | 0.204 | 0.270 | 0.335 | 0.410 | 0.505 | 0.619 | 0.756 | 0.944 | 1.171 | 1.445 | 1.769 | 2.299 |

Table 7.4: Percentage absolute error in $BP_{ND}+1$ using the RT-SRTM on simulated curves compared to gold standard BP_{ND}^*+1 . Each value was averaged across the 200 noise realisations for simulated data. The dashed lines indicate the upper limits of R_1 % error and noise found for regional analysis.

| % R_1 error | % noise in time activity curve | | | | | | | | | | | | | | |
|------------------|--------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 0.5 | 0.6 | 0.8 | 1.0 | 1.2 | 1.5 | 1.8 | 2.2 | 2.8 | 3.4 | 4.2 | 5.3 | 6.5 | 8.1 | 10.0 |
| 0.0 | 0.002 | 0.003 | 0.004 | 0.005 | 0.006 | 0.008 | 0.011 | 0.015 | 0.018 | 0.022 | 0.025 | 0.033 | 0.041 | 0.050 | 0.061 |
| 0.2 | 0.002 | 0.003 | 0.004 | 0.005 | 0.006 | 0.008 | 0.011 | 0.015 | 0.018 | 0.022 | 0.025 | 0.033 | 0.041 | 0.050 | 0.061 |
| 0.3 | 0.002 | 0.003 | 0.004 | 0.005 | 0.006 | 0.008 | 0.011 | 0.015 | 0.018 | 0.022 | 0.025 | 0.033 | 0.041 | 0.050 | 0.061 |
| 0.4 | 0.002 | 0.003 | 0.004 | 0.005 | 0.006 | 0.008 | 0.011 | 0.015 | 0.018 | 0.022 | 0.025 | 0.033 | 0.041 | 0.050 | 0.061 |
| 0.6 | 0.003 | 0.003 | 0.004 | 0.005 | 0.006 | 0.008 | 0.011 | 0.015 | 0.018 | 0.022 | 0.025 | 0.033 | 0.041 | 0.050 | 0.061 |
| 1.0 | 0.003 | 0.003 | 0.004 | 0.005 | 0.006 | 0.008 | 0.011 | 0.015 | 0.018 | 0.022 | 0.026 | 0.033 | 0.041 | 0.050 | 0.061 |
| 1.6 | 0.004 | 0.004 | 0.005 | 0.006 | 0.007 | 0.009 | 0.012 | 0.015 | 0.018 | 0.022 | 0.026 | 0.033 | 0.041 | 0.051 | 0.061 |
| 2.6 | 0.006 | 0.006 | 0.006 | 0.007 | 0.008 | 0.010 | 0.012 | 0.016 | 0.019 | 0.023 | 0.026 | 0.033 | 0.042 | 0.051 | 0.061 |
| 4.1 | 0.009 | 0.009 | 0.009 | 0.010 | 0.010 | 0.012 | 0.014 | 0.017 | 0.020 | 0.024 | 0.027 | 0.034 | 0.042 | 0.051 | 0.062 |
| 6.5 | 0.014 | 0.014 | 0.014 | 0.014 | 0.015 | 0.016 | 0.018 | 0.021 | 0.023 | 0.026 | 0.029 | 0.036 | 0.044 | 0.052 | 0.063 |
| 10.3 | 0.022 | 0.022 | 0.022 | 0.022 | 0.023 | 0.024 | 0.025 | 0.027 | 0.029 | 0.032 | 0.034 | 0.040 | 0.047 | 0.055 | 0.066 |
| 16.4 | 0.034 | 0.034 | 0.034 | 0.035 | 0.035 | 0.036 | 0.038 | 0.039 | 0.041 | 0.043 | 0.044 | 0.049 | 0.055 | 0.063 | 0.072 |
| 26.1 | 0.054 | 0.054 | 0.055 | 0.055 | 0.056 | 0.057 | 0.058 | 0.059 | 0.060 | 0.062 | 0.063 | 0.067 | 0.072 | 0.080 | 0.088 |
| 41.5 | 0.088 | 0.088 | 0.088 | 0.088 | 0.089 | 0.090 | 0.091 | 0.092 | 0.093 | 0.095 | 0.097 | 0.102 | 0.106 | 0.113 | 0.118 |
| 66.0 | 0.149 | 0.149 | 0.149 | 0.149 | 0.149 | 0.150 | 0.151 | 0.151 | 0.151 | 0.151 | 0.152 | 0.152 | 0.154 | 0.155 | 0.157 |
| SRTM | 0.001 | 0.001 | 0.002 | 0.003 | 0.003 | 0.004 | 0.005 | 0.007 | 0.008 | 0.010 | 0.012 | 0.015 | 0.019 | 0.023 | 0.030 |

Table 7.5: Mean absolute error in $BP_{ND}+1$ using the RT-SRTM on simulated curves compared to gold standard BP_{ND}^*+1 . SRTM is the gold standard fit using the full SRTM for comparison. Each value was averaged across the 200 noise realisations for simulated data. The dashed lines indicate the upper limits of R_1 % error and noise found for regional analysis.

7.4 Discussion

In this chapter I have optimised the RT-SRTM framework for quantitative PET analysis with significantly reduced acquisition time, exploiting blood flow information from simultaneously acquired ASL MRI data.

I have evaluated a new technique for extrapolating the reference region time activity curve, C_R , using PCA which introduces a lower error than the method used in Scott et al. (2016) where the mean population is scaled. The timing of the PET acquisition was then optimised, and found to be $t = 20, 50$ minutes post injection.

When the RT-SRTM estimates of BP_{ND} using $t = 20, 50$ minutes of PET data and ASL derived R_1 were compared to the gold standard using the full 60 minutes of PET data a strong linear correlation was found. This demonstrates that the RT-SRTM with a 30-minute acquisition could potentially be used as a proxy for the full 60-minute acquisition for this tracer and subject group.

By comparison, the simplified measure, $SUVR-1$, using 10 minutes of data showed a strong positive bias in the target density estimation, and the results were correlated with the delivery of the tracer as determined by the gold standard R_1 estimates. This implies that, in addition to systematic error within the $SUVR-1$ estimates, there is also a bias introduced due to local differences in blood flow. This may confound longitudinal studies, as blood flow may change over time, over the progression of disease, or due to disease modifying interventions. Conversely, the RT-SRTM estimates of target density were not correlated with R_1 , suggesting that this technique may be robust to changes in blood flow and could be a suitable alternative for longitudinal studies. However this needs to be validated in a longitudinal dataset.

$SUVR-1$ estimation at different timing windows showed that $t = 50, 60$ minutes gave the best estimation of BP_{ND} . Since $SUVR-1$ appears to have plateaued by this point, it is unlikely that the estimation can be improved by acquiring data at a later timepoint.

It has also been demonstrated that, in the absence of ASL data to inform

R_1 , BP_{ND} can be estimated from PET data acquired at $t = 20, 50$ combined with $PCA C_R$ extrapolation. This method was shown to be significantly more accurate than $SUVR-1$. This suggests that a 30 minute PET acquisition could be used to estimate BP_{ND} , however it was found that when the RT-SRTM was applied with ASL-derived R_1 , the estimation of BP_{ND} was significantly improved. Consequently, ASL data should be used where available.

This chapter focused on the optimisation of a simultaneous 30-minute PET/MR acquisition. Concurrent acquisition ensures that the CBF measured by ASL represents the flow at tracer injection, avoiding errors introduced by physiological flow changes throughout the day (Parkes et al., 2004). This assumes negligible change in blood flow between tracer injection and the end of the scan. This can be controlled through measures used for routine clinical PET scans, such as keeping the patient lying down in an uptake room from injection to scan start. The influence of auditory stimulation, which would be present during the ASL scan but not during tracer injection, on CBF should also be considered, and if necessary the conditions in the scanner should be emulated in the uptake bay. If conditions could be adequately controlled the technique could be extended to separate PET and MRI acquisitions. However this would be heavily dependent on the scheduling of the scans and the acquisition time saved is reduced compared to the simultaneous PET/MRI method.

A 30-minute acquisition was selected to accommodate a typical MRI neuroimaging session, whilst still greatly increasing patient throughput and comfort. This could be further reduced depending on the MRI data acquired, where the minimum time is determined by the acquisition of the ASL data and the images required for attenuation correction of the PET data. In this case, using the $PCA C_R$ method with fewer principal components should be considered to avoid an under-determined problem where there are more parameters to fit than datapoints available.

The ASL data used in this study was acquired for just 5.5 minutes over 50

minutes into the scan with no motion correction and no patient restraint. For this reason approximately 30% of the ASL-CBF maps failed quality control checks, largely due to motion induced artefacts. This represents a challenging dataset which could be significantly improved by increasing the number of acquisitions and motion correction. However, the fact that the RT-SRTM worked so well on this dataset indicates that it could be a clinically useful tool. It is worth noting that the data acquired from the 4 clinically diagnosed YOAD subjects all passed the quality control checks.

The limited time available within the protocol for the ASL acquisition also meant that a single delay time between blood tagging and image acquisition was used. This yields errors in CBF estimation due to different bolus transit times for different brain regions, either due to normal physiology or pathological changes. A multi-delay time ASL acquisition would make the methodology more robust as the transit time is parameterised within the model.

The RT-SRTM was here applied to an amyloid- β tracer, however the methodology could potentially be used for any tracer which can be described by the SRTM, and which has a sufficiently high extraction fraction, such as [^{18}F]-florbetaben, another amyloid- β tracer, or tau (τ) tracers (Baker et al., 2016). The kinetics of these tracers are slower than those of [^{18}F]-florbetapir, and as such require longer dynamic acquisitions. Here, the RT-SRTM could potentially offer a greater reduction in acquisition time, and could be compared with the dual time-window protocol, another acquisition time reduction method used on such tracers where early and late PET data are acquired with a break in-between (Bullich et al., 2018). Furthermore, this approach could be broadened to other kinetic models which have a tracer delivery parameter that can be approximated using CBF from ASL. Reference region curve extrapolation could also be used in reference Logan analysis Logan et al. (1996), as an alternative to a previously proposed reduced acquisition time method which cannot account for blood flow changes Tantawy et al. (2009).

The main limitation of this study is that the optimisation of the RT-SRTM

has been performed on cross-sectional dynamic scans of mostly healthy volunteers. The introduction of subjects with disease may increase the variability in C_R between subjects. However, some variability already exists in the dataset used as the healthy subjects undergo normal ageing and we include 4 clinically diagnosed subjects, which the *PCA* C_R method can handle. Provided that the dataset used to build the model for the *PCA* C_R method includes diseased subjects, this variability can also be accounted for.

Another limitation within the proposed technique is that tracer delivery is estimated directly from ASL-CBF using a region dependent linear relationship. This works well for regional data as artefacts can be averaged out, however this is not possible for voxel-wise analysis where errors will propagate through to the R_1 estimate. A more robust methodology which propagates database R_1 values into the subject space based on local image similarity has been proposed (Scott et al., 2017) and will be investigated in the following chapter. Such approaches may also be better able to handle low quality ASL data which would have failed the QC performed here. This would facilitate voxel-wise analysis for quantitative parametric imaging within a clinically feasible time frame.

7.5 Conclusions

In this chapter I have demonstrated that the three stages of the RT-SRTM should be carried out as follows: 1) R_1 estimation from ASL-CBF should be carried out using regional multi-linear regression which sensitivity analysis indicates is sufficiently accurate for regional analysis, 2) the *PCA* C_R method should be used to extrapolate the subject's reference region and 3) PET data acquired over $t = 20, 50$ minutes should be used to fit the model. Using this methodology provides reasonable estimates of BP_{ND} for regional analysis of [^{18}F]-florbetapir data using ASL data which has passed QC. However, if voxelwise analysis is required or if subjects with poor quality ASL scans are to be included then more robust approaches may be required, as investigated in the

following chapters.

Chapter 8

Comparison of R_1 estimation methods: towards voxelwise analysis

The aim of this chapter is to test whether advanced image synthesis techniques can be used to improve PET- R_1 estimation from ASL-CBF such that the RT-SRTM framework could be extended to voxelwise analysis. Here, different image synthesis methods using both multi-atlas propagation with image fusion and deep learning based regression are proposed and compared, and both the accuracy of the PET- R_1 estimation and its influence on the resulting BP_{ND} are assessed. This is also compared to the linear regression methods introduced in chapters 5 and 7. Initial work on the voxelwise analysis was presented at MICCAI 2017 (Scott et al., 2017) and PSMR 2018 (Scott et al., 2018b) and is extended to a larger subject cohort here.

8.1 Introduction

The RT-SRTM, proposed and evaluated in the preceding chapters, requires the conversion of ASL-CBF into the PET tracer delivery parameter, R_1 , which is then fixed for pharmacokinetic modelling of the 30 minutes of dynamic PET data. This approach is based on a physiological model which assumes that the relationship between ASL-CBF and PET- R_1 can be approximated using

a linear regression (LR) derived from regional average values. However, image artefacts, noise inherent in the image acquisition and assumptions in the ASL model used to calculate CBF corrupt the ASL-CBF map. Applying linear regression to this corrupted data means that the errors are propagated to the estimate of R_1 and consequently to BP_{ND} .

In the previous chapters errors in the ASL-CBF maps were minimised as quantification was performed on a regional basis. Since only regional R_1 estimates were required, this allowed unrealistic values of CBF, such as negative values, to be discarded and the noise was reduced by averaging across the valid voxels. However, regional averaging can also combine pathological and healthy areas, as the borders of pathology may not coincide with tissue boundaries derived from a T1 parcellation which are used to define the regions. This means that pathology could be underestimated or missed completely.

In this chapter, existing image synthesis frameworks are applied to the generation of R_1 from ASL-CBF, to determine whether they are better suited to the task than the physiologically based linear regression model used previously. Multi-atlas propagation with image fusion (IF) and deep learning based regression (DL) using deep convolutional neural networks (CNNs) are data driven techniques which have been widely used for medical image synthesis and make less strict assumptions about the relationship between ASL-CBF and PET- R_1 than linear regression. Consequently they should be better able to generate reasonable estimates of R_1 , even for imperfect ASL-CBF maps. They are compared with the linear regression (LR) methods developed in chapters 5 and 7.

All three approaches (LR, image fusion (IF) and deep learning (DL)) are ‘supervised’ in that they use a database of ‘training’ subjects which have both the input ASL-CBF data and corresponding output PET- R_1 , however the assumptions made are different. For LR, a direct linear relationship is enforced between the input ASL-CBF maps and the output R_1 estimate. This is a strong assumption which means that errors in the input will be propagated

to the output. It also does not allow for inter-subject differences in the mapping as a single LR is applied to the whole dataset.

The IF method assumes that if the input ASL-CBF map for the test subject has a high local similarity to that of a subject within the training database, then their output R_1 map must also be similar in that region. This is a weaker assumption than that of the LR method, as it only relies on the relationship between input and output being consistent between subjects if their inputs are similar. It also means that the input data is not used directly so ASL-CBF artefacts do not propagate into the R_1 estimation. However, it does rely on the database containing subjects which are similar to the test subject, and that areas of ASL-CBF artefact which have a high local similarity also have similar R_1 values.

For DL based methods the relationship between the input and output is learnt from the data. Consequently, the only assumption is that there is some complex set of functions relating the input ASL-CBF to the output PET- R_1 which can be learnt from the training data.

This chapter compares the performance of several implementations of these three of image synthesis approaches for the generation of PET- R_1 maps, where multiple implementations are achieved by changing the number of modalities included in the input data. Both the accuracy in R_1 and BP_{ND} estimation are used to determine whether the evaluated methods are sufficiently accurate to allow the RT-SRTM method to be applied on a voxelwise basis.

8.2 R_1 estimation approaches

Three different types of image synthesis approach were applied to the generation of PET- R_1 maps; LR, IF and DL. From these three approaches a total of five methods were used to derive R_1 from ASL-CBF, where the extra methods are achieved by adding input modalities to the training set. For each method a database of subjects with 60 minutes of simultaneously acquired PET data,

ASL and (optionally) T1-w MRI and dynamic PET frames are used. The methods are implemented as described in the following sections.

8.2.1 Linear regression

Linear regression (LR) (Scott et al., 2016) As described in chapter 7, briefly, a single linear regression was performed on all voxels within the brain mask between PET- R_1 and ASL-CBF on the database of training subjects to derive the relationship between the two parameters. This relationship derived from all voxels combined was then applied to an unseen ASL-CBF map from the test set, see figure 8.1.

Regional linear regression (rLR) (Scott et al., 2018c) As described in chapter 7, briefly, regions were defined using T1 based tissue parcellations generated using GIF (Cardoso et al., 2015) which were combined into 16 regions (excluding the reference region cerebellar grey matter), see section 5.2.3. Linear regression was then performed between PET- R_1 and ASL-CBF for each voxel for each of the 16 regions on the database of training subjects to derive the relationship between the two parameters. This results in region specific linear regression coefficients, as apposed to LR which produces a single linear regression. The regional regression coefficients are then applied to the ASL-CBF map of the test subjects.

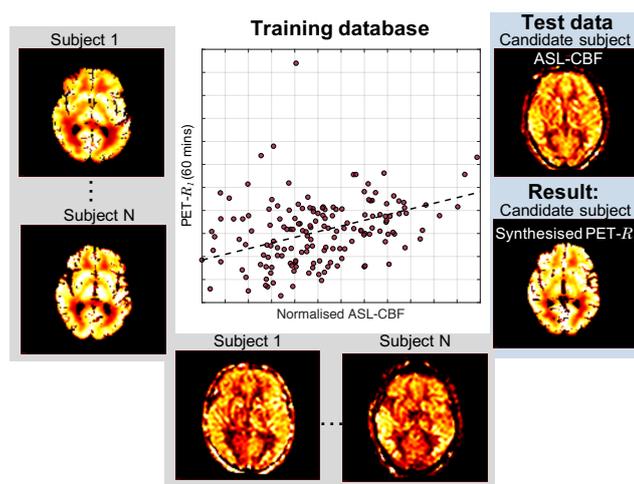


Figure 8.1: Generation of R_1 maps from ASL-CBF maps using linear regression (LR). Result image is for illustration only.

8.2.2 Multi-atlas propagation with image fusion (IF) (Scott et al., 2017)

Multi-atlas propagation with image fusion has been used for a range of image synthesis tasks, including the generation of pseudo CT data, which is required for attenuation correction in PET-MR imaging (Burgos et al., 2015). The framework developed by Burgos et al. (2015), which is available as part of NiftyPipe¹, was applied to this R_1 synthesis task by changing the training database. The adaptation of this framework to R_1 synthesis was initially described in Scott et al. (2017), and is explored in more detail here.

Let $A = \{A_1, \dots, A_N\}$ be the training database which comprises N subjects. Here A_i consists of M_i input image modalities and an output image O_i as $A_i = \{I_{1_i}, \dots, I_{M_i}; O_i\}$. Given a test subject $S_T = \{S_{T_1}, \dots, S_{T_M}\}$, the aim is to estimate O_T . Based on the terminology used in Burgos et al. (2015), if $N=1$ then this is a single atlas method, if $M=1$ then this is a single contrast method. In this case $M = 2$ where I_{1_i} is the structural T1-w image, I_{2_i} is the ASL-CBF map, and O_i is the corresponding R_1 map. The T1-w image data was included to improve image registration and to allow the image similarity to be calculated even in regions where the ASL-CBF map contains no valid data.

The basic steps involved in multi-atlas multi-contrast image synthesis are outlined in the following sections, and are summarised in figure 8.2.

8.2.2.1 Database construction

For each of the N subjects, all modalities within A_i are spatially aligned with one another and with O_i via rigid registration to ensure spatial correspondence, see data pre-processing section 8.3.4 for details. This data now constitutes the training database.

¹<http://cmictig.cs.ucl.ac.uk/wiki/index.php/NiftyPipe>

8.2.2.2 Registration to test subject

For $i = (1, \dots, N)$, each training T1-w image (I_{1_i}) is non-rigidly registered to the test T1-w image S_{T_1} to get the deformation D_i , where an affine registration is used to initialise the non-rigid registration (Modat et al., 2010). Only the T1-w images were used for registration, and not the ASL-CBF maps, as the ASL-CBF maps are difficult to register due to noise and artefacts, see section 8.3.4. D_i is then applied to the ASL-CBF map (I_{2_i}) and R_1 image (O_i).

8.2.2.3 Similarity calculation

The local similarity between each of the registered input training database images and the candidate subject is then calculated. For each training subject i and voxel v the convolutional ROI- locally normalised cross correlation (LNCC) between the training T1-w image (I_{1_i}) to the test T1-w image (S_{T_1}) is calculated as:

$$\text{ROI-LNCC}_{i,v}(S_{T_1}, I_{1_i}) = \frac{\langle S_{T_1}, D_i I_{1_i} \rangle_v}{\sigma(S_{T_1})_v \sigma(D_i I_{1_i})_v}$$

$$\text{where } \langle S, DI \rangle_v = \overline{S \cdot DI}_v - \overline{S}_v \cdot \overline{DI}_v, \quad \sigma(S_v) = \sqrt{\overline{S_v^2} - \overline{S}_v^2} \quad (8.1)$$

and the convolution is calculated using a Gaussian kernel G_{σ_G} with a standard deviation $\sigma_G = 5\text{mm}$ and a density normalisation term Ω to ensure that valid estimates are only produced where the image field of views (FOVs) overlap:

$$\overline{S}_v = \frac{[G_{\sigma_G} \otimes S]_v}{[G_{\sigma_G} \otimes \Omega]_v}, \quad \Omega = 1 \text{ where FOV's overlap, } = 0 \text{ otherwise} \quad (8.2)$$

The ROI-LNCC between the training ASL-CBF maps (I_{2_i}) and the test ASL-CBF map (S_{T_2}) are also calculated and the similarity metrics from the two modalities are combined into a single value for each voxel called the multivariate (MV) ROI-LNCC as:

$$\text{MV-ROI-LNCC}_{i,v} = \text{ROI-LNCC}_{i,v}(S_{T_1}, I_{1_i}) + \text{ROI-LNCC}_{i,v}(S_{T_2}, I_{2_i}) \quad (8.3)$$

Calculating the similarity between each of the database images and the candidate images using an ROI size which is smaller than the image volume increases the accuracy and flexibility of the technique, as two subjects may resemble one another in one area, but not in another. It also means that the technique is less reliant on the accuracy of the registration.

8.2.2.4 Image fusion

For each voxel v , the N training datasets are ranked based on the MV-ROI-LNCC similarity, and the rank (R) is converted to a weight using an exponential decay function as:

$$W_{i,v} = e^{-\beta R_{i,v}}, \quad \text{where } \beta=0.5 \text{ by default} \quad (8.4)$$

For each voxel, the synthesised output value (O_{T_v}) is generated through a weighted sum of the registered output database image intensities using:

$$O_{T_v} = \frac{\sum_{i=1}^N W_{i,v} D_i(O_i)}{\sum_{i=1}^N W_{i,v}} \quad (8.5)$$

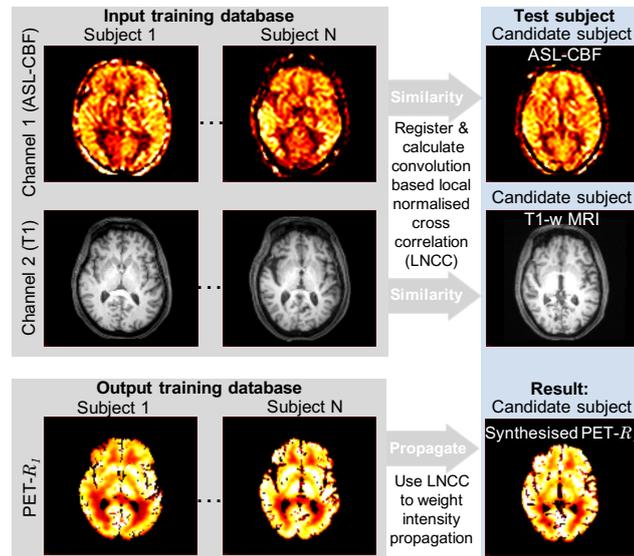


Figure 8.2: Registration transforms input database images into the candidate subject space. The local similarity between the images is calculated and is used to weight the propagation of R_1 information from the output database into candidate space. Images are for illustration only and don't represent actual results.

8.2.3 Deep learning based regression

Deep learning is a branch of machine learning methods which has exploded in popularity in medical imaging over recent years due to its successful application to a wide range of image processing tasks. According to the comprehensive review by Litjens et al. (2017) deep convolutional neural network (CNN)s trained end-to-end using supervised learning are the most popular choice in medical imaging. As a result, an open source implementation called ‘highres3Dnet’ (Li et al., 2017) available in NiftyNet (Gibson et al., 2018) was selected as a suitable deep learning based approach for R_1 synthesis. A brief overview of deep neural networks (largely adapted from Goodfellow et al. (2016)) and details of the implementation used are given below.

8.2.3.1 Neural network organisation

Network layers. Multi-layer neural networks consist of individual neurons, or units, which are organised into layers. Information is passed from one layer to the next starting from the input layer which reads in the data, through the hidden layers, to the output layer which produces the result, as shown in figure 8.3a. The units in the hidden layers can perform simple non-linear operations, and by combining them in parallel as a layer they are able to learn complex functions. By increasing the number of hidden layers, known as the depth, increasingly abstract features can be learnt from the training data. The width of a network is the number of units per layer, and the depth is the number of hidden layers.

Pre-activation. Each unit takes a set of d inputs $\vec{x} = [x_1, \dots, x_d]$ and produces an output $h(\vec{x})$ as shown in figure 8.3b. The units perform this calculation in two stages: firstly the pre-activation, a , is calculated based on a weighted sum of the inputs x and a bias b , as in equation (8.6). The weights, $\vec{w} = [w_1, \dots, w_d]$, determine the strength of the connection of the neuron to its inputs. Both the weights and the bias (a scalar value for a single neuron) are parameters of the model which are tuned during training.

$$a(\vec{x}) = b + \sum_{i=1}^d w_i x_i = b + \vec{w}^T \vec{x} \quad (8.6)$$

Alternatively, for a convolutional layer in a CNN, the pre-activation is calculated by convolving the input with a kernel to produce the output (feature map), as opposed to matrix multiplication. The kernel itself is a multi-dimensional array of parameters which are adapted by the learning algorithm, which are analogous to the weights and bias for conventional neural network layers.

Post-activation. The neuron output (or post-activation) $h(\vec{x})$ is equal to the activation function $g(a(\vec{x}))$ which operates on the pre-activation, $a(\vec{x})$. The activation function introduces non-linearity into the network. The default function used in NiftyNet is the rectified linear unit (ReLU) where $h(\vec{x}) = \max(0, a(\vec{x}))$.

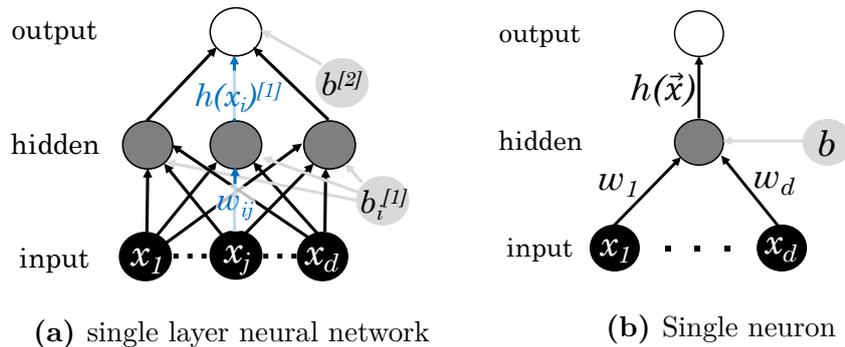


Figure 8.3: schematics to show the building blocks of feed-forward neural networks where each circle represents a unit or neuron

CNN properties. CNNs are specifically designed for inputs with a grid-like structure such as medical images and have several desirable properties. The use of single convolution kernel per layer, rather than a multiplication by a different weight for each layer input, means that each element of the kernel is used at every position across the inputs, and is therefore shared. This reduces the computational cost of the network and also makes it equivariant to translations such that if the input is translated, then the output will be translated by the same amount. This is useful for images in which there is a

feature which occurs in multiple locations across the image.

Translational equivariance can be extended to invariance using pooling. The pooling operation, which is applied after the post-activation, is a function which combines a group of outputs from a spatial neighbourhood into a single value. One example is max pooling, where the maximum output is used, but there are many different options. If pooling is applied across layers with different kernels, then the network can learn which transformations to be invariant to.

Sparsity can also be introduced by using a kernel which is smaller than the input to the layer. This reduces the number of connections compared to a fully connected network which increases training efficiency while reducing the memory requirements.

8.2.3.2 Training a Convolutional Neural Network

For end-to-end supervised learning, the parameters are learnt during iterative training as follows:

Initialisation. The parameters of the convolution kernels must be initialised to begin training. Random initialisation is commonly used to ensure that units learn as independently as possible.

Forward propagation. The input training data is fed through the network to generate an output for each instance of the input training data at the current iteration (i). To speed up the training process, instead of passing each training example through the network, a random subset, or mini-batch, of the data can be propagated through.

Loss calculation and backpropagation. The output from the network can then be compared to the true output by evaluating the cost function to give the loss at the current iteration. The direction in which the model parameters should change is found using backpropagation. In backpropagation the gradient of the cost function is calculated with respect to the model parameters using the chain rule. If the gradient is calculated over a mini-batch of the training data then this is called mini-batch gradient descent.

Parameter optimisation. The parameters (w_i) are updated by an amount proportional to the gradient (G) as in equation (8.7). The size of the step (η) taken when updating the parameters is determined by the learning rate and is a hyperparameter which can be set at the beginning by the user. Several methods, such as adaptive moment estimation (Adam), have been proposed where the learning rate is adapted for each weight based on how well it is learning (as determined by the gradient). Adam is the default option in NiftyNet and was used here to minimise the loss quickly while avoiding overshooting.

$$w_{i+1} = w_i + \eta G \quad (8.7)$$

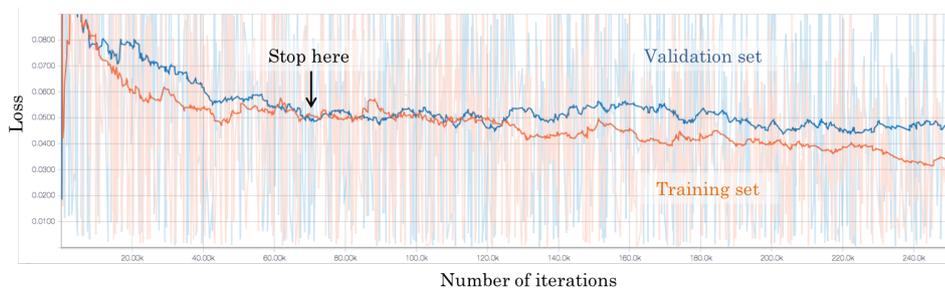


Figure 8.4: Plot generated by Tensorboard showing the training set (orange) and validation set (blue) loss as a function of training iterations. The arrow indicates the point at which the model starts to overfit to the training data and training should be stopped.

Stopping criteria. The network parameters can be updated at each iteration until a stopping criterion is satisfied. To avoid under or overfitting to the training data a technique called early stopping can be used. Here, the data is divided into validation, training and testing sets. The network is trained on the training data and after a fixed number of iterations, inference is performed on the validation set and the loss is calculated. By comparing the training loss and the validation loss as a function of the number of iterations, as shown in figure 8.4, the point at which the network begins to overfit to the training data can be determined. This point is where the validation loss (blue) flattens out or begins increasing as the model can no longer generalise to the unseen data, this occurs at around 70,000 iterations in figure 8.4. By contrast, the training

set loss (orange) will continue to decrease. This method provides a stopping criteria which aims to maximise the generalisability of the network, although it does reduce the number of available training examples due to the use of a validation set.

Inference. Once the model has been trained the model parameters are fixed and the test data can be passed through to generate an estimate of the output.

8.2.3.3 Regularisation.

Deep neural networks typically contain a very large number of parameters, and consequently can be prone to overfitting, as the network can simply learn the training data exactly. Ideally the training dataset would be sufficiently large to mitigate this problem, however this is not practical as datasets of that magnitude are not available, particularly in medical imaging. There are a variety of regularisation methods which can be used, which include:

Data augmentation. This can be used to artificially increase the size of an existing training dataset. New data is generated by applying a transformation to the existing training data; such as a small translation, or adding noise to the images. The augmented data is then included in the training dataset.

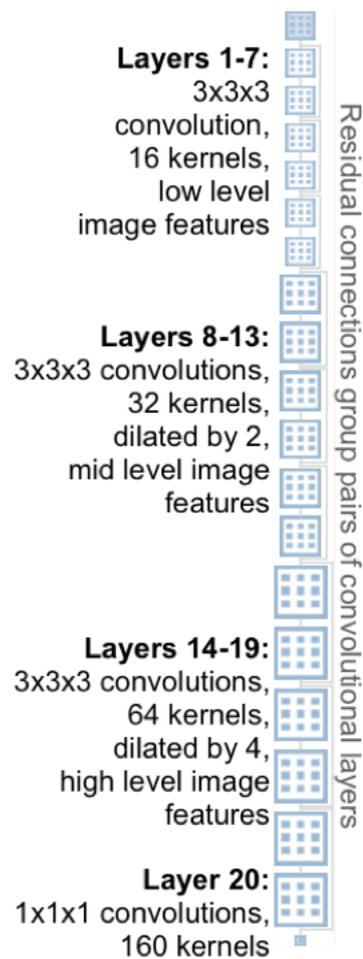


Figure 8.5: Overview of ‘Highres3Dnet’ architecture showing the 20 convolutional layers of the network and the paired residual connections, adapted from Li et al. (2017).

Dropout. Dropout works on the principle that the generalisation error can be reduced by training several different models separately then combining the results, since different models are unlikely to make the same mistakes. For dropout, this is implemented by randomly excluding a fraction of the non output layer units each time a batch is processed. This essentially means that a different model is trained each time, and is implemented by setting a probability of each unit being included. Dropout can also be performed at inference to give a probabilistic output (Nair et al., 2018).

8.2.3.4 ‘Highres3Dnet’ architecture and training

For both of the deep learning based methods in this chapter, ‘Highres3Dnet’ (Li et al., 2017) is used to perform image regression for R_1 synthesis. ‘Highres3Dnet’ is a convolutional neural network with 20 convolutional layers, which uses a stack of residual dilated convolutions with increasingly large dilation factors as illustrated in figure 8.5. Dilated convolutions are an alternative to pooling which allow the receptive field to be increased without downsampling the input, and without increasing the number of parameters. For dilated convolutions the kernel parameters are spaced out by a dilation factor with zeros in the gaps (Yu and Koltun, 2015). This leads to an exponential increase in the receptive field through the layers, while only linearly increasing the number of parameters. The residual connections shown in grey in figure 8.5 allow the deep network to be trained more efficiently by mimicking a series of shallow networks, as proposed by He et al. (2016).

To train the networks, adaptive moment estimation (Adam) with an initial learning rate of 10^{-3} was used with a root mean square error loss function and a mini-batch size of 2. The network parameters were initialised randomly (a different network was trained for each fold of the cross-validation, see section 8.3.5) and trained for a maximum of 250,000 iterations. The training patch size was $56 \times 56 \times 56$ voxels and a smoothed brain mask was used to mask the loss calculation. Random rotation and scaling transformations of $\pm 10\%$ were used for training data augmentation. All inputs were 3D image volumes.

Deep learning (DL) method (Scott et al., 2018b). For this method ‘High-res3Dnet’ was trained using the T1-w MRI data and the ASL-CBF maps as inputs, as shown in figure 8.6.

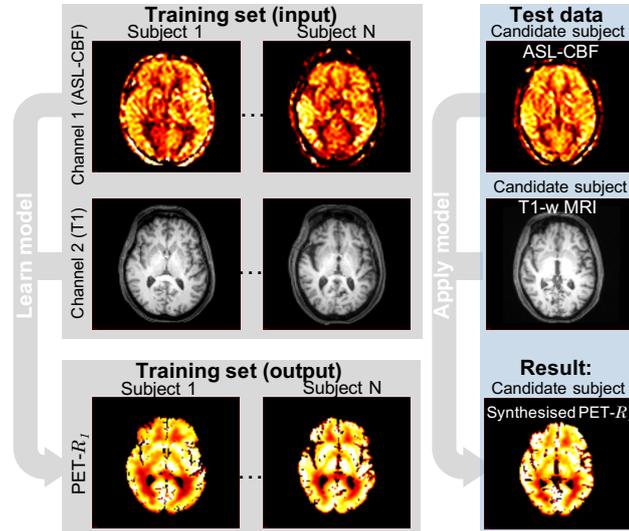


Figure 8.6: The workflow for the DL method, where the model is trained using ASL-CBF and T1-w MRI as the input with PET- R_1 as the output. Inference can then be applied to the test set using the trained model. Images are for illustration only and don’t represent actual results.

Deep learning plus (DL+) method. This network was trained in the same way as the DL method, however there was an additional input which was a static PET frame acquired at $t = 20, 25$ minutes. This was included to give some additional information about the tracer delivery in addition to the ASL-CBF data.

8.3 Experimental set-up

8.3.1 RT-SRTM implementation

The RT-SRTM framework, described in detail in section 5.2.2, requires three steps to estimate BP_{ND} which were implemented as follows:

Step 1) Estimation of R_1 from ASL-CBF. This is performed using one of the methods described in section 8.2.

Step 2) Extrapolation of PET reference region, $C_R(t)$. To compute the convolution term in equation 3.20, C_R must be known from injection, at $t = 0$,

so C_R must be extrapolated from the incomplete measured PET data. The *PCA* C_R method (Scott et al., 2018c) was used, where PCA is applied to the C_R of the training subjects and the model is then fitted to the available data for the test subject, see chapter 7.2.2 for more details.

Step 3) PK modelling on reduced PET data. As shown in section 5.2.2 equation (5.5), basis functions are generated using the extrapolated C_R from step 2), and C_T is fitted by fixing R_1 from step 1) to give BP_{ND} and K_2 . Dynamic data from $t=20,50$ mins was used as optimised in chapter 7.

8.3.2 Comparison methods

Two types of comparative analysis are performed in this chapter. Firstly, a comparison is performed between the 5 different R_1 estimation approaches, described in section 8.2, with respect to the gold standard, R_1^* . The second part of the comparative analysis is between the BP_{ND} estimated using the RT-SRTM, using the 5 different R_1 estimation methods, with the gold standard BP_{ND}^* and the clinical standard $SUVR-1$.

Gold standard R_1^* and BP_{ND}^* estimation. As described in section 5.2.1 the linearised simplified reference tissue model (SRTM) is used for gold standard PET quantification (Gunn et al., 1997), using dynamic PET data acquired from $t=0,60$ minutes.

$SUVR-1$ calculation. $SUVR$ was calculated by summing dynamic frames to create a 10 minute static image, then dividing through by the mean cerebellar grey matter value. $SUVR-1$ was used for comparison with BP_{ND} by subtracting 1 from the $SUVR$. As in chapter 7 the timing of the PET data used to calculate $SUVR$ was optimised by comparing estimates to the gold standard to ensure that it was a fair comparison. Here, this was performed on a voxel-wise basis across the whole brain, then grey matter and white matter voxels only. Regional analysis using the same 16 regions as in the previous chapters was also optimised in this chapter to allow for comparison.

8.3.3 Data

Data were acquired from 75 subjects participating in three separate studies: 57 were cognitively normal subjects participating in Insight 46, a neuroimaging sub-study of the MRC National Survey of Health and Development (Lane et al., 2017) (1946), 4 were clinically diagnosed with probable Young Onset Alzheimer’s disease (YOAD) and 14 were participating in the biomarkers and outcomes in cerebral amyloid angiopathy (BOCAA) study where 5 were controls. All subjects underwent simultaneous amyloid PET and multi-modal MRI on a Siemens Biograph mMR 3T PET-MRI scanner. List mode PET data were acquired for 60 minutes following intravenous injection of [^{18}F]-florbetapir. More details on each study can be found in appendix A.1.

8.3.4 Data pre-processing

PET and MRI data acquisition, reconstruction and processing was performed as described in 6.2, however some adjustments were made to facilitate voxelwise analysis.

Firstly, the PET frames were smoothed using a using a 3mm FWHM Gaussian filter to reduce noise in the time activity curves. The FWHM was determined empirically to give optimal results for kinetic modelling on the gold standard data through visual assessment of a subset of the parameter maps. The filter was then applied to all PET data used.

Voxelwise analysis requires the ASL-CBF maps to be spatially aligned with the PET- R_1 maps. This was performed by first affinely registering the T1-w MRI image to the non-attenuation corrected PET reconstruction ($t=0,60$ minutes). The registered T1-w image in PET space was then used as the target for the ASL registration. Since the ASL-CBF is noisy and contains artefacts, the saturation recovery image ($TR=1\text{s}$) was used as the floating image, in the place of the ASL-CBF map, as it contains anatomical information which is similar to the T1-w image and is in the same space as the CBF map and acquired at a similar time. The transformation found is then applied to the CBF map to propagate it into PET space. Rigid registration was

used throughout as all registrations are intra-subject and was performed using NiftyReg (Modat et al., 2014b).

Finally, the registered ASL-CBF maps were smoothed to match the resolution of the PET data. The smoothing kernel was determined by assessing the voxelwise correlation between the PET- R_1 and ASL-CBF using Gaussian filters with FWHM values ranging between 0 and 3mm. The kernel which resulted in the highest correlation was selected.

8.3.5 Validation and Analysis

Cross-validation. Each R_1 estimation method, described in section 8.2, was implemented using five-fold cross-validation. In each fold, 15 unseen subjects are tested (test set) with 60 subjects in the database (training set). For the DL based methods a validation set of 7 subjects, drawn from the training set, was also used to determine when to stop training the model. Consequently, for DL and DL+, each fold contained 53 training subjects, 7 validation subjects, and 15 subjects for testing.

The folds were determined using stratified sampling to ensure that the distribution of subject groups in each fold represented the overall subject population used. The data were nominally classed as negative-subject (1946 subjects and BOCAA controls who were $a\beta$ negative), positive-subject (1946 subjects who were $a\beta$ positive), positive-patient (YOAD subjects who were $a\beta$ positive) and BOCAA-patient. The BOCAA patients were stratified separately as it was not trivial to classify them as globally $a\beta$ -positive or negative due to other pathologies. See appendix A.1 for more details of the studies used.

The training set was used both for the generation of R_1 , and for the extrapolation of the reference region, C_R .

Statistical Analysis. Calculation of the MSE and ME, to quantify accuracy and bias respectively, was performed as described in section 5.2.5. Individual comparisons between methods were performed using the paired, 2-tailed Wilcoxon signed-rank test for MSE (as the data are not normally distributed), and paired 2-tailed t-tests for ME. To account for multiple comparisons Fried-

man’s test with Bonferroni post-hoc corrections was performed for MSE, and one way ANOVA with Tukey’s Honest Significant Difference was applied to ME. For each statistic quoted it is implicit that the accompanying p -value is derived by controlling for multiple comparisons, unless it is explicitly stated that ‘individual comparison’ was applied. In all cases the threshold for significance was set at $p \leq 0.05$. It should be noted that the Bonferroni correction results in a loss of power and is used as a strict measure of statistical significance to highlight methods which give strongly different results as apposed to incremental methodological improvements.

Voxelwise analysis. For voxelwise analysis the errors between estimated and the true parameter values have been calculated in two ways; for all voxels across the whole brain (WB), and for grey matter (GM) voxels only. This is because voxelwise analysis of the whole brain is dominated by white matter voxels which are not of principal importance for amyloid- β imaging. However, the WB errors are still useful for seeing how the methods perform across different tissues and are useful to assess the generalisability to other PET tracers. It is worth noting that in the deep learning based regression methods, equal weighting was given to all brain voxels during training, whereas the $SUVR-1$ method was optimised separately for the whole brain and for grey matter only. The long computation time required to train separate networks using grey matter voxels only was not performed, since reasonable results were gained using the whole brain.

8.4 Results and discussion

8.4.1 Pre-processing and SUVR optimisation

8.4.1.1 ASL-CBF map smoothing

Figure 8.7 shows both the Spearman (non-parametric) and Pearson (linear) correlations between the true PET- R_1 and the relative ASL-CBF maps using a range of Gaussian smoothing kernels between 0 and 3mm. The correlation was calculated on a voxelwise basis for all brain voxels and for grey matter

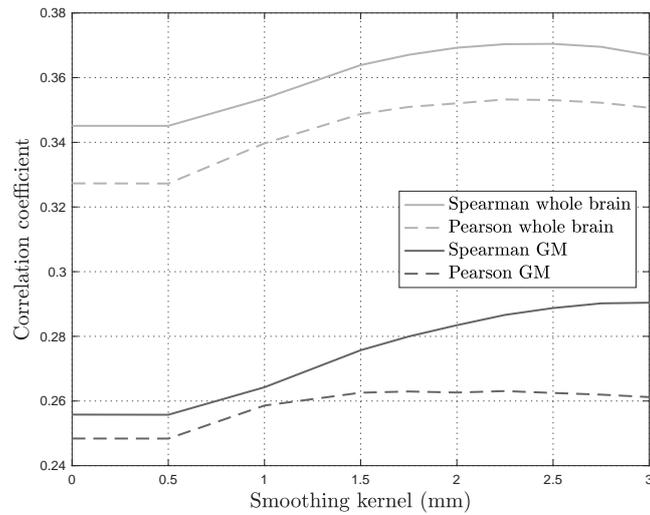


Figure 8.7: The voxelwise correlation coefficient between relative ASL-CBF and PET- R_1 as a function of the Gaussian smoothing kernel applied to the ASL-CBF map. This was used to determine the kernel used to smooth the ASL-CBF maps in the database.

voxels only. The plot in figure 8.7 shows that the optimal FWHM for the Gaussian smoothing kernel is 2.5mm in most cases, although the Spearman correlation in grey matter appears to continue to increase up to 3mm. This suggests that the underlying resolution of both the PET and the ASL data is similar, as the PET data was smoothed with a 3mm Gaussian kernel. It is important to note that the ASL acquisition was not isotropic and that the data was reconstructed into 2x2x4mm voxels. This suggests that an anisotropic smoothing kernel could be useful here, however this was considered to be too complex to optimise for a small gain in the correlation between the datasets. Consequently an isotropic 2.5mm Gaussian kernel was applied to all of the ASL data used in subsequent sections.

8.4.1.2 $SUVR-1$ timing optimisation

The optimal timing window for $SUVR-1$ calculation was estimated by calculating $SUVR-1$ at different timing windows and comparing them to the gold standard BP_{ND}^* . Table 8.1 shows the mean squared error (MSE) between $SUVR-1$ and BP_{ND} for each timing window, which demonstrates that

the optimal timing for $SUVR-1$ calculation depends on which brain tissues are of interest. For whole brain voxelwise analysis, where there are a large proportion of white matter voxels, the MSE is lowest for 25:35 minutes post injection. This is also true if only white matter voxels are included, however if only grey matter voxels are used then the optimal timing window is much later, at $t=50,60$ minutes post injection.

The $t=50,60$ minute timing window was also found to be optimal for regional analysis (denoted 16 regions in table 8.1). This is because only 2 out of 16 regions consist of white matter. This supports what was found in the previous chapter and the recommendations in the literature, as it is generally grey matter quantification which is important for amyloid PET imaging.

It should be noted that calculating $SUVR-1$ at timepoints earlier than 50 minutes post-injection means that the tracer has not reached a steady-state, see chapter 3.3. This is not only a violation of the model assumptions, but it also means that the value of $SUVR-1$ is changing rapidly, and is therefore strongly time dependent, see figure 7.7. This is not ideal in general practice, as it leads to a high likelihood of errors, but this was performed here to represent the best performance of $SUVR-1$ to give the fairest comparison between methods.

| Window | Whole brain | Cortical GM | White matter | 16 regions |
|------------|---------------|---------------|---------------|---------------|
| 20:30 mins | 0.0323 | 0.0303 | 0.0363 | 0.0121 |
| 25:35 mins | 0.0281 | 0.0251 | 0.0299 | 0.0129 |
| 30:40 mins | 0.0314 | 0.0234 | 0.0394 | 0.0140 |
| 35:45 mins | 0.0365 | 0.0234 | 0.0526 | 0.0143 |
| 40:50 mins | 0.0403 | 0.0236 | 0.0629 | 0.0137 |
| 45:55 mins | 0.0413 | 0.0233 | 0.0670 | 0.0123 |
| 50:60 mins | 0.0403 | 0.0228 | 0.0664 | 0.0110 |

Table 8.1: Mean squared error between $SUVR-1$ and BP_{ND} averaged over 75 subjects and estimated at different timing windows.

8.4.2 Regional R_1 estimation comparison

Table 8.2 shows the average regional errors in the estimation of R_1 for the five different methods tested: LR, rLR, IF, DL and DL+, compared to the

gold standard R_1^* . The LR method produced R_1 estimates with a statistically significantly higher MSE ($p < 0.001$) and ME ($p < 0.001$) than the other four methods when corrected for multiple comparisons, except for the ME in the DL method which was only significant for individual comparisons ($p = 0.151$ when corrected for multiple comparisons). The poor performance of the LR method is due to the strong assumption that there is a single linear relationship between ASL-CBF and PET- R_1 which has previously been shown to be an oversimplification (see chapter 6) as the relationship is regionally dependent. Furthermore, the relationship may not be linear, particularly in the presence of artefacts.

When comparing the four best performing methods, DL+ had the lowest MSE, however they were all found to be statistically equivalent for individual and multiple comparisons. The DL and DL+ methods are implemented in using the same approach except that the DL+ method includes an additional input, consisting of a single PET frame, both for training and inference. This does not result in a significant reduction in the MSE of the R_1 estimates from DL+ compared to DL, however it does greatly reduce the bias (ME). Table 8.2 shows that the rLR method produces R_1 estimates with the lowest ME, followed by IF. However, the ME of the rLR, IF and DL+ methods are statistically equivalent for both single comparisons ($p \geq 0.709$) and with corrections for multiple comparisons.

Overall, the DL+ method produced R_1 estimates with the lowest MSE and a low ME which suggests that the DL+ is the best performing method for regional R_1 estimation. This is likely to be due to the fact that the DL based methods impose the weakest assumption about the relationship between the input data with R_1 . Furthermore, the addition of the single PET frame gives the network information relating to the extraction of the tracer as well as the blood flow from ASL-CBF. This is explored in more detail in the voxelwise analysis in section 8.4.3.

The fact that the DL+ method does not produce an MSE that is sta-

| Method | Mean Square Error | Mean Error |
|--------|-------------------|----------------|
| LR | 0.0193 | -0.0484 |
| rLR | 0.0084 | -0.0017 |
| IF | 0.0080 | -0.0034 |
| DL | 0.0079 | -0.0217 |
| DL+ | 0.0059 | -0.0036 |

Table 8.2: Regional mean squared error (MSE) and mean error (ME) in R_1 estimates averaged over 75 subjects.

tistically significantly lower than the rLR, IF and DL methods demonstrates that all of these methods are sufficiently accurate for regional analysis. This suggests that there is limited benefit in increasing the complexity of the R_1 synthesis method for regional R_1 estimation.

Whilst table 8.2 shows that these R_1 estimation methods can be used on a regional basis, the R_1 parameter maps in figure 8.8 demonstrate why regional analysis may not be appropriate for all patients. Figure 8.8b shows local R_1 reductions in the anterior portion of the brain, thought to be due to previous stroke in this patient, who was participating in the BOCAA study. This reduction is not visualised in the regional R_1 map shown in figure 8.8a, as the regional boundaries defined by the T1 parcellation do not match with the diseased area. Consequently the drop in R_1 is averaged out over several regions. This effect is noticeable in many of the diseased subjects used in this study. This highlights the necessity of voxelwise kinetic modelling for such patients, which is assessed in the following sections.

Figure 8.8c also shows that this particular subject (BOCAA-17) may not be suitable for quantification using the gold standard SRTM, since the reference tissue, cerebellar grey matter, shows a reduction in R_1 in the inferior portion. This leads to R_1 values which are much higher than those seen in other subjects (see figure 8.14a) and was not seen on any of the other subjects. Despite the violations of the SRTM model assumptions this subject was not excluded from analysis as such atypical subjects are useful for showing the strengths and weaknesses of the different techniques. The invalidity of the model should, however, be considered when assessing the results for this subject.

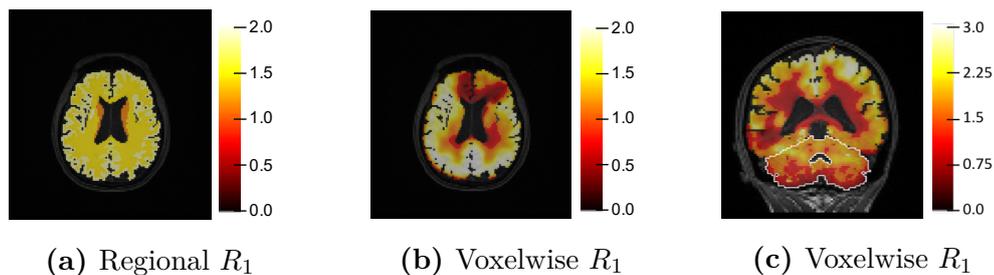


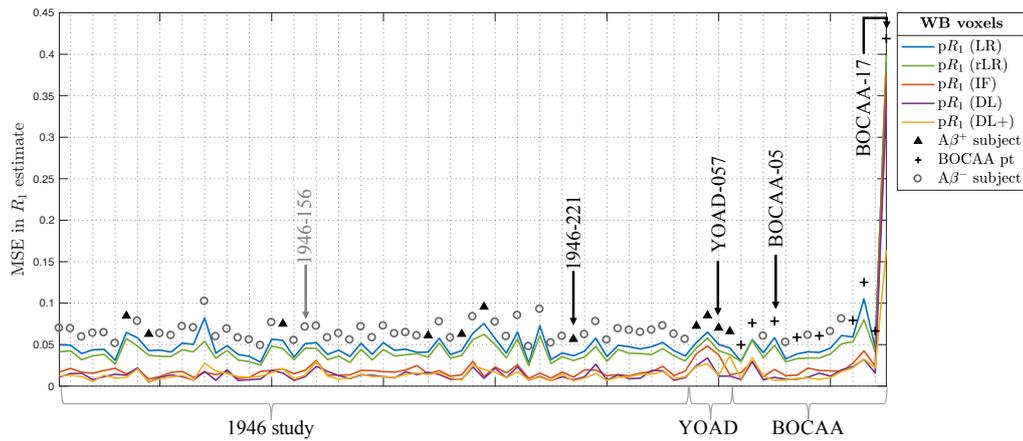
Figure 8.8: Parameter maps for subject BOCAA-017, fitted from $t=0:60$ mins of PET data showing that regional analysis (a) does not adequately describe the tracer delivery for this patient due to stroke induced R_1 reductions which is not confined to a single region, shown by the voxelwise R_1 map in b). c) shows reduced tracer delivery in the inferior portion of the cerebellum (outlined in white) affecting much of the grey matter. This demonstrates that grey matter cerebellum may not be a suitable reference region for this subject.

| Method | WB voxels | | GM voxels | |
|--------|---------------|----------------|---------------|----------------|
| | MSE | ME | MSE | ME |
| LR | 0.0524 | -0.0002 | 0.0519 | -0.0482 |
| rLR | 0.0455 | -0.0011 | 0.0463 | -0.0011 |
| IF | 0.0232 | -0.0005 | 0.0306 | -0.0112 |
| DL | 0.0181 | -0.0134 | 0.0230 | -0.0355 |
| DL+ | 0.0160 | -0.0063 | 0.0199 | -0.0028 |

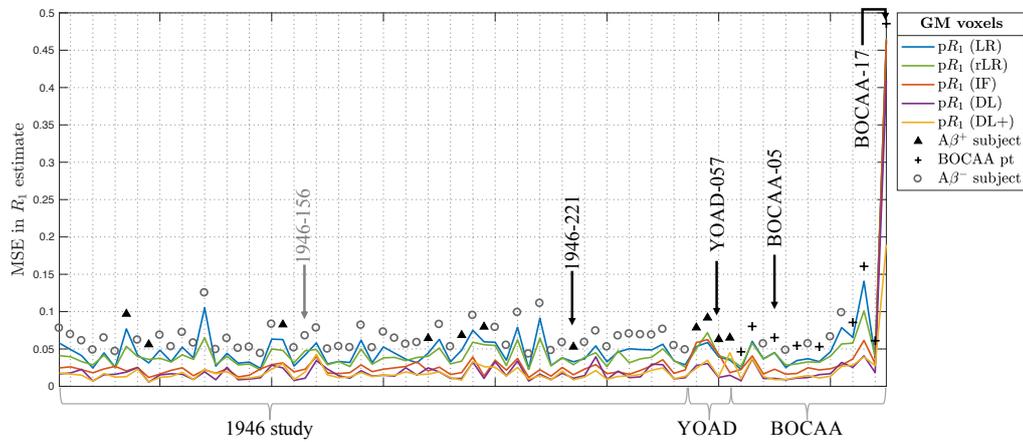
Table 8.3: Voxel mean squared error (MSE) in R_1 estimates averaged over 75 subjects using all brain voxels (WB) and grey matter (GM) voxels only.

8.4.3 Voxelwise R_1 estimation comparison

Table 8.3 shows the errors in voxelwise R_1 estimation for the five different methods tested compared to the gold standard R_1^* value. Both for the whole brain and for grey matter voxels only the DL and DL+ methods have a significantly lower MSE than the other methods, even when correcting for multiple comparisons ($p < 0.001$). When directly comparing the DL and DL+ methods they were found to have a statistically equivalent MSE (individual and multiple comparisons $p > 0.05$), however, as for the regional case, the inclusion of the PET data into the model leads to a reduction in the ME. For individual comparison between the methods, the reduction in ME reaches statistical significance for the grey matter voxels ($p = 0.006$) and approaches significance across the whole brain ($p = 0.065$), but the significance does not survive multiple comparison correction ($p \geq 0.0891$). Overall, when performing multiple



(a) All voxels across the whole brain (WB)



(b) Cortical grey matter (GM) voxels only

Figure 8.9: MSE in estimated voxelwise R_1 (lines) plotted by subject along the x-axis. Shapes above the lines indicate the subject classification and curved brackets show the study the subject came from.

comparisons all methods had a statistically equivalent ME for the whole brain ($p = 0.5442$), and in grey matter only LR had significantly higher ME than the other methods ($p \leq 0.037$). As a result, the DL+ appears to be the optimal method to derive R_1 when averaging across all subjects in the dataset used.

The plots in figure 8.9 show the error in R_1 estimation for each method across the 75 subjects included in this analysis. These plots demonstrate how the techniques perform across the different cohorts and the disease status of the subjects. Figure 8.9 clearly demonstrates that the novel three image synthesis approaches (IF, DL and DL+), which don't use the ASL-CBF maps directly,

perform much better than the LR and rLR methods, due to a combination of factors. Firstly, these image synthesis approaches benefit from the fact that they do not use the ASL-CBF values themselves, hence artefacts are not propagated directly into the estimated R_1 image. Moreover, these approaches also relax the assumptions made about the relationship between the ASL-CBF maps and the PET- R_1 ; instead of assuming a direct, linear relationship between the two parameters the IF method assumes that if a patch within the candidate ASL-CBF maps and T1-w images looks similar to a patch in the database, then the corresponding R_1 values for that patch must also be similar. Furthermore, the deep learning based methods make no assumptions about the data other than that the input images (ASL-CBF maps and T1-w images for DL, and the ASL-CBF maps, T1-w images and a single PET frame for DL+) contain the information required to estimate the output R_1 , and hence that the relationship between input and output can be learnt. This makes the deep learning methods particularly suitable for this task, as here the exact relationship between the ASL-CBF and the PET- R_1 is unknown and complex due to imperfect quantification in both modalities. Finally, the novel image synthesis methods (IF, DL and DL+) take the spatial context of a voxel into account, rather than considering the voxel in isolation as in the linear regression based techniques. Whilst the rLR method does use the brain region within the regression to improve the R_1 estimate, it doesn't consider where in the region the voxel is located, nor does it use any information from neighbouring voxels.

The superior performance of the image synthesis methods are evident in two example subjects from the 1946 study shown in figures 8.10a (amyloid- β negative) and 8.11a (amyloid- β positive) where the novel image synthesis methods (IF, DL, DL+) are much closer to the true R_1^* map, as they are better able to capture variation across the image. Introducing the regional information into the rLR method does improve the R_1 estimation and increase the contrast between regions, but without knowing the context of each voxel

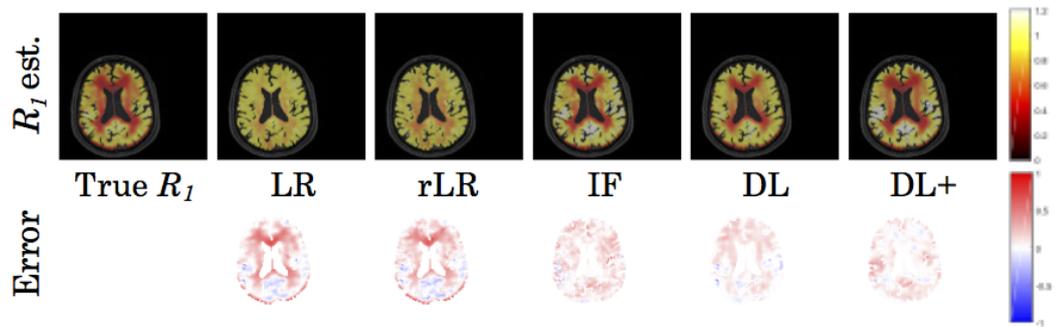
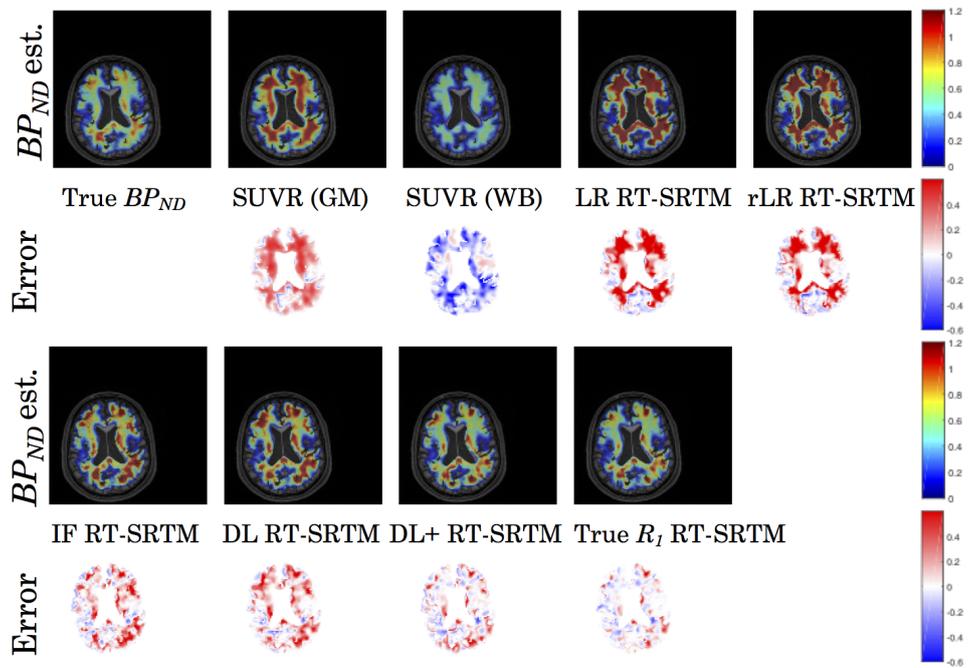
(a) R_1 parameter maps(b) BP_{ND} parameter maps

Figure 8.10: Parameter maps for subject 1946-156 who is a healthy amyloid- β negative ($A\beta^-$) subject.

and simply applying the regional linear regression to a single value, the local variation can not be adequately described and the resulting image closely resembles a regional average.

The IF, DL and DL+ methods are also able to reproduce the PVEs which occur in the PET data largely as a result of the low spatial resolution of the PET acquisition, see section 2.2.2. Ideally the PET images would contain minimal PVEs following post-reconstruction or within-reconstruction partial volume correction (PVC), however the optimal implementation of PVC techniques is still a matter of debate within the literature and as such was considered beyond the scope of this work (Thomas et al., 2011; Schwarz et al., 2017). The PVEs here manifest as a reduction in R_1 in the posterior edges of the brain due to spill out effects, as demonstrated in figure 8.10a. The ASL-CBF maps were smoothed during the pre-processing stage in an attempt to match the spatial resolution of the PET data, however the Gaussian smoothing kernel was selected based on the value which gave the best correlation between ASL-CBF and PET- R_1 , rather than through visual inspection. Consequently the LR and rLR methods do not reproduce the partial volume effects seen in the PET data.

For the subjects from the 1946 study, who are age matched, healthy older controls, figure 8.9 shows that the three novel image fusion methods (IF, DL and DL+) perform similarly well for both whole brain and grey matter voxels. However, it also shows that when applied to more heterogeneous subjects, as found in the YOAD and BOCAA cohorts, the IF method does not perform as well as the two deep learning based methods. This is particularly evident in figure 8.9b which shows the MSE by subject within cortical grey matter voxels. Figure 8.9b shows a marked increase in the MSE for the YOAD subjects using the IF method, as demonstrated by the parameter maps for an example subject in figure 8.12a. Here, the reduced R_1 in the posterior of the brain is not reproduced in the R_1 map generated by the IF method. It is likely that this reduction in R_1 is not present in the synthesised image as there were

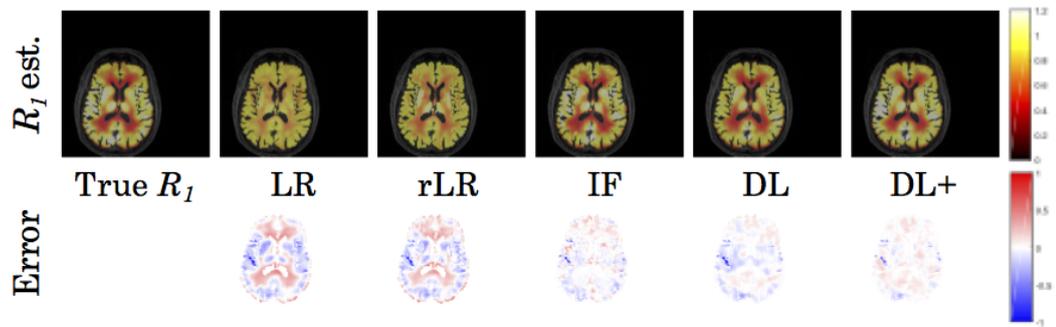
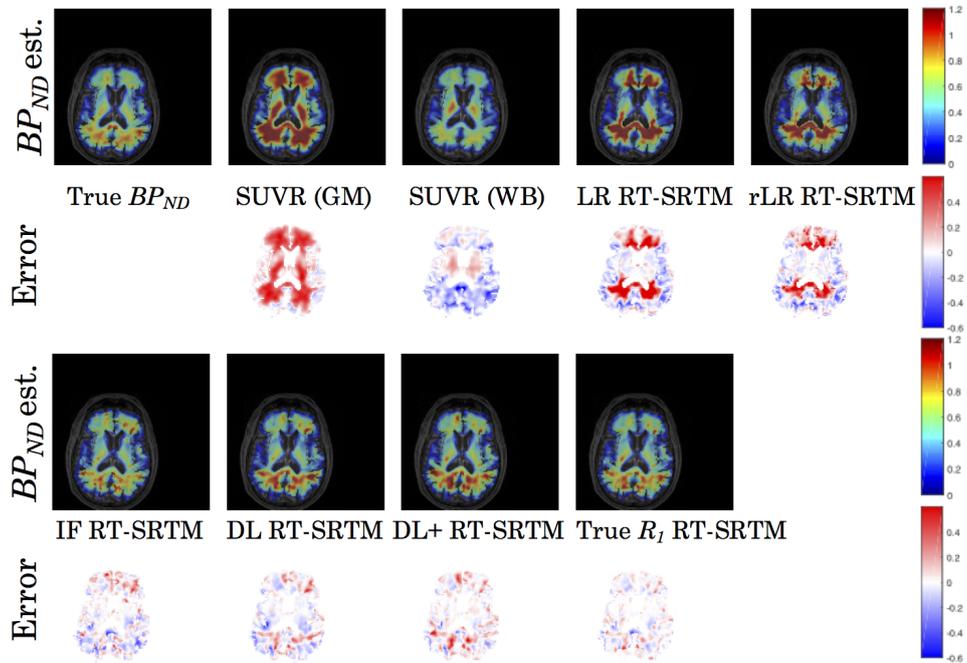
(a) R_1 parameter maps(b) BP_{ND} parameter maps

Figure 8.11: Parameter maps for subject 1946-221 who is a healthy amyloid- β positive ($A\beta^+$) subject.

no images in the training database which showed the same effect, so the R_1 information which was propagated through from subjects which did not match the candidate image. This could be improved by using a database with more YOAD subjects, however collating databases of a sufficient size and range of diseases is always challenging due to limited data availability. Conversely, since the deep learning based methods do not assume that the images look similar, but instead learn the relationship between the input and output data they are better able to cope with test images which differ from the training data, although they require large training sets to avoid overfitting.

The deep learning methods, however, do rely on the fact that the input data contains all the information required to estimate the output. Figure 8.13a shows that the ASL-CBF maps and the structural T1-w images alone are not enough to describe the reduced tracer delivery, highlighted by the dotted box, as the DL method overestimates R_1 in this region. However, when some PET data is included in the model, as in the DL+ method, then the extra information improves the model and so it produces a more accurate estimate of the R_1 map.

The example subjects shown in figures 8.12a and 8.13a demonstrate that deep learning based methods are better able to produce accurate R_1 estimates for test subjects that are not similar to the subjects in the training database compared to the IF method. However, there is a limit to how different the test image can be to the training database, beyond which the R_1 estimation becomes unreliable. This is demonstrated for the BOCAA subject shown in figure 8.14a. Here the true R_1^* image shows that the R_1 values across much of the brain are significantly higher than any of the other subjects due to reduced tracer delivery in the inferior portion of the cerebellum as discussed in section 8.4.2. Furthermore, this subject shows two areas of greatly reduced R_1 in the frontal regions of the brain which are likely to have been induced by a previous stroke. The ASL-CBF map did not show a noticeable difference in blood flow between these very high and low R_1 areas, therefore the linear

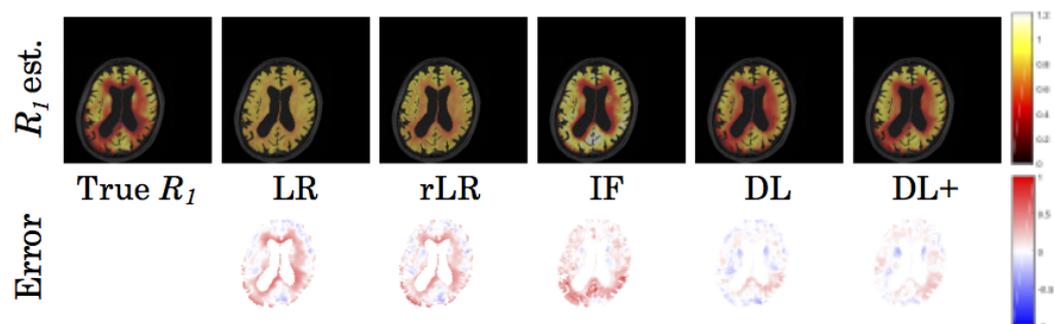
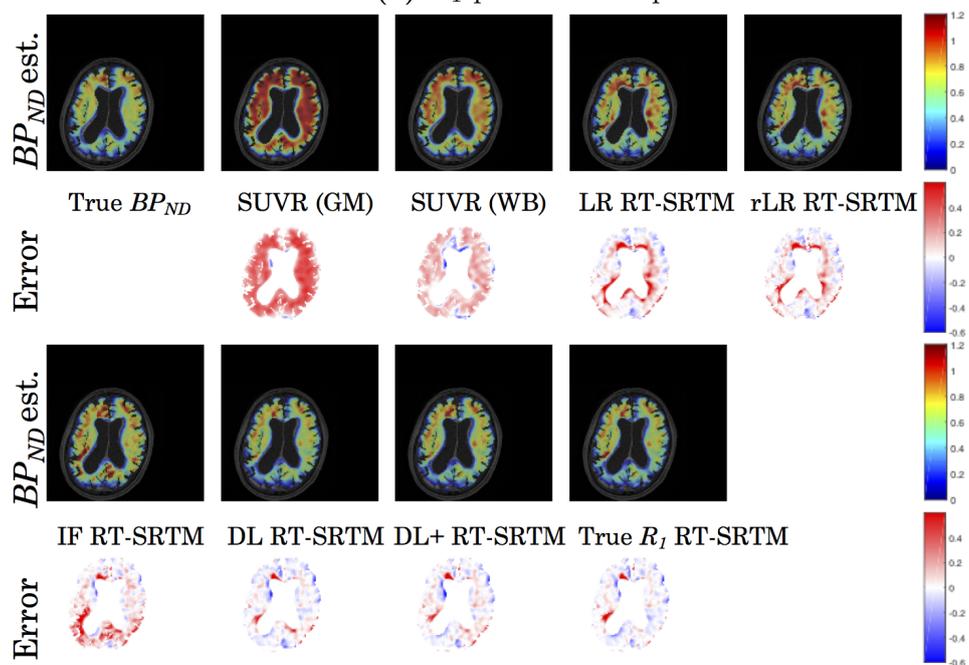
(a) R_1 parameter maps(b) BP_{ND} parameter maps

Figure 8.12: Parameter maps for subject YOAD-057 who is a amyloid- β positive ($A\beta^+$) patient with clinically diagnosed AD.

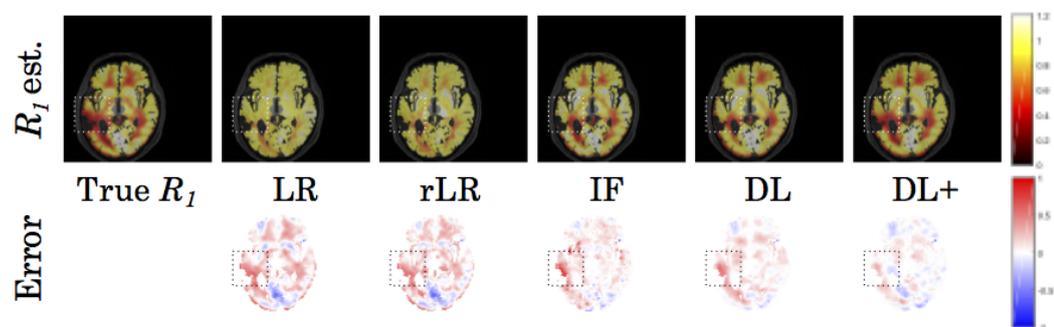
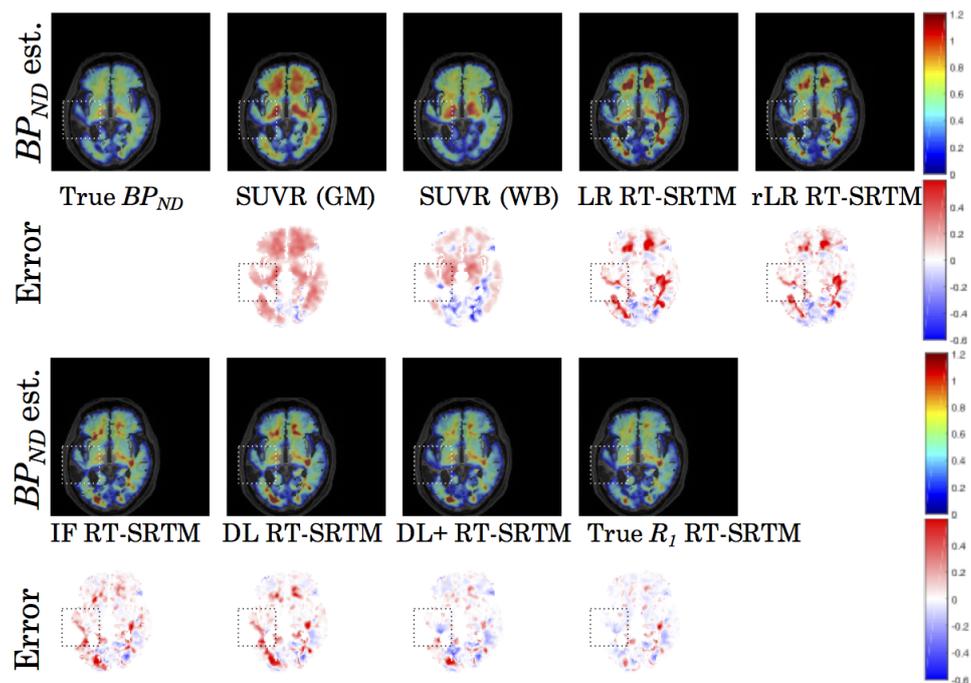
(a) R_1 parameter maps(b) BP_{ND} parameter maps

Figure 8.13: Parameter maps for subject BOCAA-05 who has cerebral amyloid angiopathy.

regression methods were not able to accurately estimate R_1 in these areas. The T1-w image also didn't show any structural changes in these regions so neither the IF nor the DL method were able to capture this effect. Only the introduction of the PET data in the DL+ method creates R_1 maps which appear to reproduce the pattern seen. However this did still result in a general overestimation of the R_1 map which may be related to the violation of the SRTM for this subject, as previously discussed. This subject represents the limit of the techniques tested as it contains patterns which don't exist in the rest of the dataset.

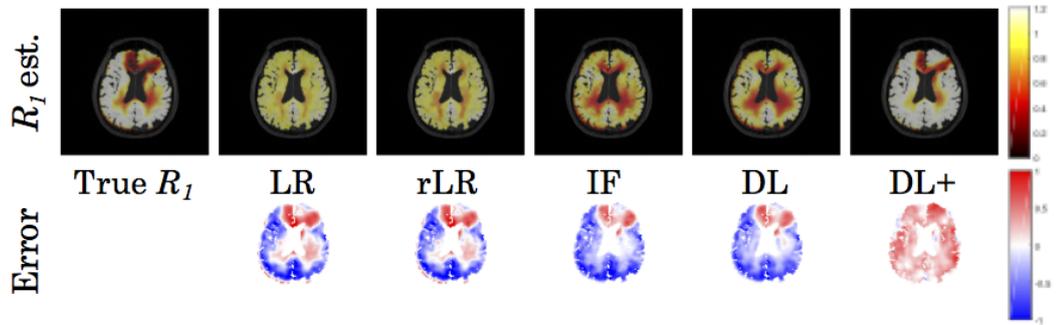
8.4.4 Voxelwise BP_{ND} estimation comparison

Once the estimated R_1 maps have been generated they can be used within the RT-SRTM to produce BP_{ND} maps. In addition to using the synthesised R_1 maps with the RT-SRTM to estimate BP_{ND} , the BP_{ND} is also calculated using the RT-SRTM with the gold standard R_1^* (denoted 'true' in some figures for clarity) to demonstrate the upper limit of BP_{ND} estimation performance using this model. The outputs from the RT-SRTM methods are compared to the optimised $SUVR-1$ as well as the gold standard BP_{ND}^* .

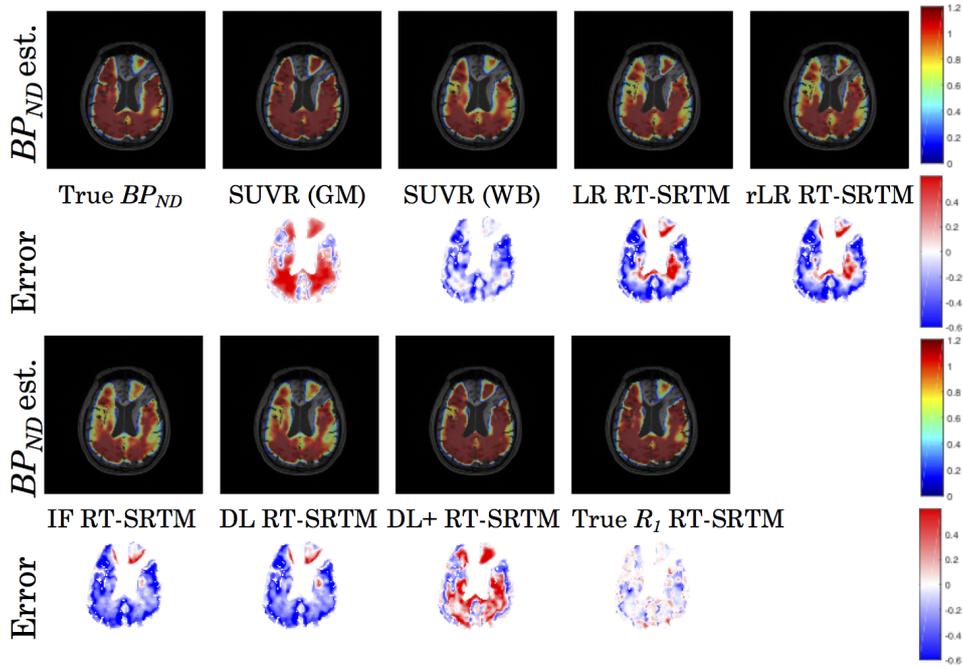
| Method | All brain voxels | | GM voxels | |
|-----------------|------------------|---------------|---------------|---------------|
| | MSE | ME | MSE | ME |
| $SUVR-1$ (opt) | 0.0217 | 0.0250 | 0.0228 | 0.0129 |
| LR RT-SRTM | 0.0334 | 0.0370 | 0.0227 | -0.0519 |
| rLR RT-SRTM | 0.0262 | 0.0244 | 0.0229 | -0.0358 |
| IF RT-SRTM | 0.0129 | 0.0106 | 0.0197 | -0.0387 |
| DL RT-SRTM | 0.0120 | 0.0090 | 0.0183 | -0.0446 |
| DL+ RT-SRTM | 0.0109 | 0.0154 | 0.0176 | -0.0312 |
| R_1^* RT-SRTM | 0.0054 | 0.0084 | 0.0133 | -0.0312 |

Table 8.4: Voxel mean squared error (MSE) in BP_{ND} estimates averaged over 75 subjects.

Table 8.4 shows the error in BP_{ND} against the gold standard BP_{ND}^* for all seven methods implemented. This shows that, for voxelwise analysis, the LR and rLR methods no longer out-perform $SUVR-1$. Here, the LR method performs significantly worse than $SUVR-1$ ($p = 0.002$), and the rLR method



(a) R_1 parameter maps



(b) BP_{ND} parameter maps

Figure 8.14: Parameter maps for subject BOCAA-17 who has cerebral amyloid angiopathy.

is statistically equivalent. The simulations performed in section 7.3.7 showed that voxelwise analysis requires higher accuracy in R_1 estimation compared to regional analysis to achieve the same BP_{ND} accuracy. This is due to the higher noise in the voxelwise dynamic PET data compared to the regional averages, which makes it more difficult to fit the model. The results in table 8.4 suggest that the R_1 estimates using LR and rLR are not sufficiently accurate for the voxelwise RT-SRTM.

The IF, DL and DL+ methods all produce a significantly lower mean MSE than $SUVR-1$ ($p \leq 0.001$) across the whole brain, although for grey matter only the IF method has a statistically equivalent MSE to $SUVR-1$. The lower MSE for the IF, DL and DL+ demonstrates that the improved R_1 estimation achieved with these methods translates into better BP_{ND} estimates. As for the errors in R_1 estimation, see section 8.4.3, the DL and DL+ methods both have a lower MSE than the IF approach when performing individual comparison ($p \leq 0.001$), but the significance did not survive correction for multiple comparisons. Again, the DL and DL+ methods were found to be statistically equivalent to one another ($p \geq 0.086$), despite the DL+ method showing a lower mean MSE. Similarly to the bias observed in R_1 estimation, individual comparisons of the two deep learning methods showed that they have statistically equivalent bias' across the whole brain ($p = 0.469$) and the DL+ method had a significantly lower bias when only GM voxels were considered ($p = 0.003$), although this did not survive correction for multiple comparisons.

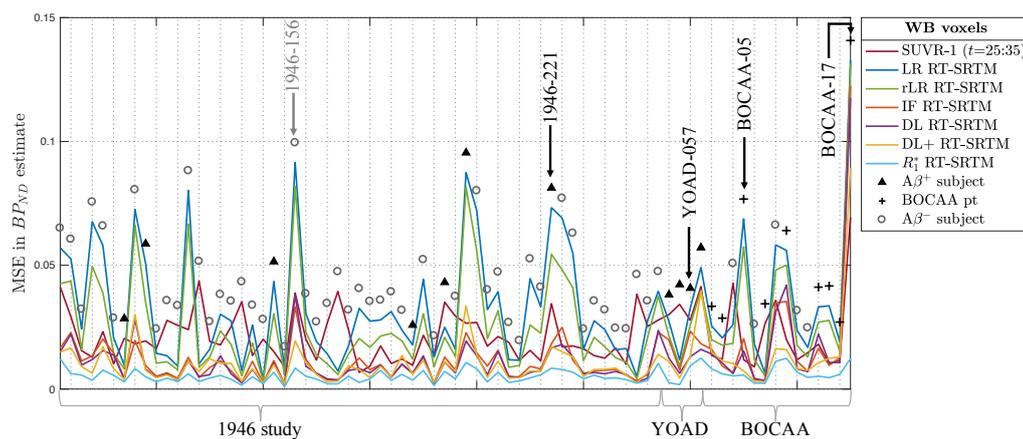
The plot of the MSE in the BP_{ND} estimate by subject, shown in figure 8.15, shows how the methods perform across the different patient groups. When all the brain voxels are included in the analysis, figure 8.15a shows that the LR and rLR methods introduce large errors in the majority of the subjects. This again demonstrates that these methods can not be used on a voxelwise basis. Figure 8.15b, which includes cortical GM voxels only, shows that $SUVR-1$ yields a particularly poor estimate of BP_{ND} in the amyloid- β positive subjects, most notably in the YOAD cohort. This is illustrated by

the BP_{ND} maps in figure 8.12b, where the $SUVR-1$ method overestimated the level of tracer binding. This confirms what was seen on a regional basis in section 7.3.5, which showed that the bias in $SUVR-1$ increases with BP_{ND} ; since the YOAD subjects show the highest BP_{ND} , they also show the highest $SUVR-1$ bias.

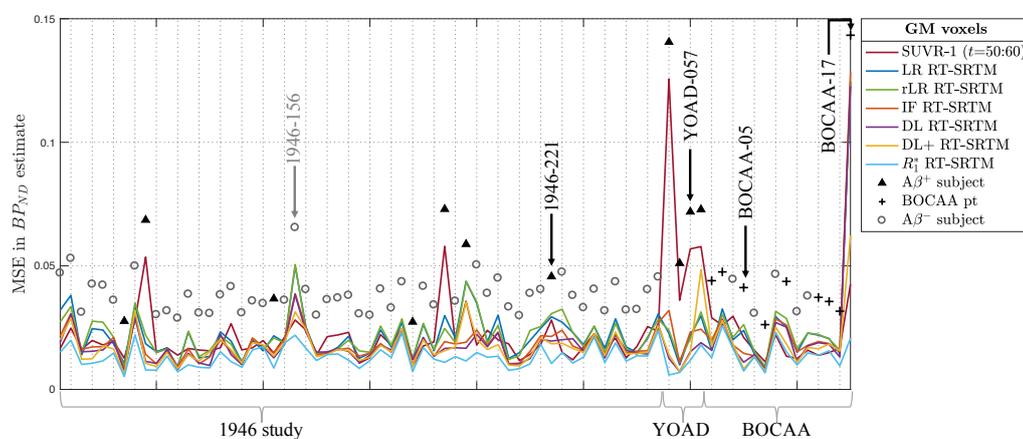
Whilst the difference in MSE between the DL and DL+ methods did not reach statistical significance, figure 8.13b shows that the addition of the PET data is useful in regions of low tracer extraction. The R_1 maps in figure 8.13a showed that only the DL+ method was able to reproduce the reduction in R_1 highlighted within the dotted box. This improved R_1 estimation accuracy in turn improves the BP_{ND} map in figure 8.13b where the same area is highlighted by the dotted box. This shows that the DL+ method performs very similarly to the RT-SRTM using the gold standard ‘true’ R_1^* maps, which represents the upper limit of the performance of this technique.

However, as was the case when analysing the R_1 maps, the BP_{ND} estimates for the BOCAA patient shown in figure 8.14b are poor for all methods. Whilst the DL+ method appears to give a reasonable estimate of the BP_{ND} distribution, the error map shows that the overestimation of the R_1 leads to large errors in the BP_{ND} map. For this subject (BOCAA-17), figure 8.15 shows that $SUVR-1$ generally produced BP_{ND} estimates with a lower MSE than the RT-SRTM methods. Only by improving the R_1 estimate for this subject, and getting a result close to the gold standard R_1^* can the RT-SRTM achieve a low MSE across all subjects used, as shown in figure 8.15. However, it is important to note that the SRTM is not an appropriate model for this patient, as the images show abnormal tracer uptake in the cerebellar grey matter, which means that this is not a valid reference region. If this subject were excluded then figure 8.15 shows that the DL and DL+ methods provide robust estimates of BP_{ND} across all the subject groups.

Table 8.5 shows the error in BP_{ND} estimation when the data from subject BOCAA-17 are excluded. Compared to table 8.4 it is clear that for all methods



(a) All brain (WB) voxels



(b) Grey matter (GM) voxels

Figure 8.15: MSE in estimated voxelwise BP_{ND} (lines) plotted by subject along the x-axis. Shapes above the lines indicate the subject classification and curved brackets show the study the subject came from.

the errors have reduced, although the change for the R_1^* method is minimal, but the overall trends remain the same. The IF, DL and DL+ methods all perform significantly better than $SUVR-1$ ($p \leq 0.001$) across the whole brain, and the DL and DL+ methods have a significantly lower average MSE in grey matter ($p \leq 0.002$). Individual comparison suggests that the deep learning methods both perform significantly better than IF ($p < 0.001$), but the significance is lost when accounting to multiple comparisons. Despite this, the deep learning methods are not able to approach the BP_{ND}^* accuracy achieved when using R_1^* .

| Method | All brain voxels | | GM voxels | |
|---------------------|------------------|---------------|---------------|----------------|
| | MSE | ME | MSE | ME |
| <i>SUVR-1</i> (opt) | 0.0211 | 0.0273 | 0.0225 | 0.0125 |
| LR RT-SRTM | 0.0320 | 0.0400 | 0.0213 | -0.0493 |
| rLR RT-SRTM | 0.0248 | 0.0274 | 0.0216 | -0.0329 |
| IF RT-SRTM | 0.0114 | 0.0141 | 0.0182 | -0.0356 |
| DL RT-SRTM | 0.0105 | 0.0123 | 0.0169 | -0.0417 |
| DL+ RT-SRTM | 0.0098 | 0.0145 | 0.0170 | -0.0315 |
| R_1^* RT-SRTM | <i>0.0054</i> | <i>0.0087</i> | <i>0.0132</i> | <i>-0.0312</i> |

Table 8.5: Voxel mean squared error (MSE) in BP_{ND} estimates averaged over 74 subjects (excluding BOCAA17).

8.5 Conclusions

In this chapter five different methods (LR, rLR, IF, DL and DL+) for generating R_1 maps from ASL-CBF maps are compared on a varied cohort of subjects. Both the quality of the derived R_1 maps and the corresponding BP_{ND} resulting from the RT-SRTM method were used to evaluate the methods. Regional analysis of the derived R_1 maps showed that the simplest method, LR, had a significantly higher MSE than the other methods as the rigid assumption that there is a single linear relationship between ASL-CBF and PET- R_1 is not valid for this data. The most complex DL+ method was found to have the lowest MSE, however the MSE was not significantly lower than the other novel image synthesis methods (IF and DL). This is because regional averaging allows local variation and high error voxels to be averaged out, and all methods can perform reasonably well. Regional averaging was also shown to mask pathological areas which do not necessarily follow the boundaries of a structural brain segmentation used to define regions. Consequently further analysis was performed on a voxelwise basis.

Comparison of the voxelwise R_1 maps from the different methods shows that, without regional averaging, the deep learning methods (DL and DL+) perform significantly better than the other methods. This is because these methods fully exploit all of the information in the data given by learning the relationship between the input and the output images with no assumptions

imposed on the data. This means that the deep learning based methods generalise better than IF across subjects which appear different from the majority of the training set, as represented by the YOAD and BOCAA cohorts. Furthermore, the addition of a single PET frame into the DL+ method reduces the bias in this method for the GM voxels, although the MSE is not significantly lower than the DL method.

It is important to note that the areas of reduced tracer delivery seen in the BOCAA-17 subject was not visible on the ASL-CBF maps. From the available data it is not clear whether this is because these areas have normal blood flow but reduced tracer extraction, or whether the ASL data is not of sufficient quality, either due to the acquisition itself or subject motion, to capture the reduced blood flow. This should be explored in the future.

When the R_1 estimates were used with the RT-SRTM to generate BP_{ND} and compared to $SUVR-1$ it became clear that the LR and rLR methods perform poorly for voxelwise analysis. These methods result in higher errors compared to BP_{ND}^* than a 10 minute $SUVR-1$ estimation. This may be due to the increased difficulty in fitting noisy voxelwise data which requires a more accurate R_1 estimation to constrain the fit and give a reasonable BP_{ND} estimate. However it is important to note that a fixed 20:50 minute timing window was used based on optimisation of the R_1^* RT-SRTM, see chapter 7. The optimisation was performed in this way to make it generalisable, rather than specific to the R_1 estimation method used, however it means that it may not be directly applicable to the other RT-SRTM methods, particularly as table 8.3 shows a negative bias in R_1 estimation for all methods used.

The IF, DL and DL+ methods performed significantly better than both LR, rLR, and $SUVR-1$, with the deep learning based methods performing best. However, whilst the deep learning methods perform well, and are a definite improvement over the currently used $SUVR-1$, using the true R_1^* still produces a better estimate of BP_{ND} , indicating that there is room for improvement.

The addition of the PET frame in the DL+ method did improve the BP_{ND} estimation visually, however the result did not reach statistical significance. One of the drawbacks of the DL and DL+ methods is that they use the ‘high-res3Dnet’ architecture which combines the input volumes in the first layer, so the low level features are combined. Architectures such as ‘scalenet’ (Fidon et al., 2017) have been proposed where the volumes are processed separately for a set number of layers, before the feature maps are combined. This allows the model to learn more from the individual images before combining them, and could potentially lead to an improved R_1 estimation in this context.

All of the proposed methods struggled with the subjects which were not represented by the training data. From a machine learning perspective the dataset used was challenging as it was relatively small, and the groups were unbalanced. Ideally more data from the YOAD and BOCAA studies would be included, however as this data is currently unavailable data augmentation techniques could potentially be employed to re-balance the training dataset by preferentially augmenting the under-represented groups.

Despite the limitations of the techniques discussed here, this work shows that advanced image synthesis techniques such as multi-atlas propagation with image fusion and deep learning based regression can be used out of the box to extend the application of the RT-SRTM to voxelwise analysis. This is essential for accurately assessing the disease status of a patient and means that the method could directly replace $SUVR-1$ estimation. However, it is important to note that while $SUVR-1$ was optimised for timing, a 10 minute acquisition window was used. This is reasonable for the GM or 16 region approach, where the later the acquisition the lower the error, but for whole brain analysis a longer acquisition window could potentially improve performance.

By further optimising and tailoring the deep learning approach to this specific problem there is the potential for the resulting BP_{ND} estimates to approach those achieved using R_1^* in the RT-SRTM, leading to BP_{ND} estimates which are almost indistinguishable from BP_{ND}^* , but with half the acquisition

time.

Chapter 9

Direct PET-MR binding potential map synthesis using CNNs

In this chapter a novel method to generate BP_{ND} from the dynamic PET frames and ASL data in a single step using deep learning (SS-DL) is tested. Since this approach does not require explicit R_1 estimation, nor does it need curve fitting or C_R extrapolation which means that fewer PET frames are required. For this reason the SS-DL method is evaluated on acquisitions as short as 10 minutes. The work in this chapter is an extension of the preliminary study published in Scott et al. (2018a).

9.1 Introduction

In the previous chapters the RT-SRTM has been shown to be an effective method of reducing the acquisition time required for PET quantification. However, the minimum length of the acquisition is determined by the number of datapoints required to accurately extrapolate the reference region (C_R) and to fit the PK model to the data. Theoretically, the number of datapoints could be increased for the same acquisition time by reducing the frame length, but this would result in an increase in noise. As shown in the simulations in chapter 7, increased noise makes the model more difficult to fit so there is no

advantage to increasing the number of datapoints. Direct parametric reconstruction could be used to reduce noise, though the effect is limited and comes at a high computational cost (Jiao et al., 2017). Consequently, the RT-SRTM has been constrained to a 30 minute minimum acquisition time, which is still intolerable for many patients.

A further limitation of the RT-SRTM is that it estimates R_1 from ASL-CBF independently from dynamic PET fitting for target density (BP_{ND}) and washout rate (k_2). This implementation cannot explicitly model the known influence of CBF on washout, due to high uncertainty in washout estimation, and the complex relationship which is dependent on the local tissue tracer kinetics (Wu and Carson, 2002). Thus the information from the ASL-CBF maps is being underutilised.

In this chapter a deep learning approach which achieves PET quantification for a short acquisition time in a single step (SS-DL) is evaluated. The noise sensitive voxelwise PK curve fitting step of the RT-SRTM is avoided through the use of deep convolutional neural networks which enforce spatial regularisation across the receptive field. The SS-DL approach also negates the need for explicit modelling between CBF, tracer delivery and tracer washout, as these relationships are learnt from the data and modelled in conjunction with the dynamic PET data. Furthermore it avoids C_R extrapolation, further facilitating acquisition time reduction.

The SS-DL framework for direct BP_{ND} synthesis, initially presented in Scott et al. (2018a), was the first time in which deep learning has been successfully applied to replace PET pharmacokinetic modelling. This is largely due to the availability of robust models, such as the SRTM, which can adequately describe standard dynamic PET data such that deep learning offers no advantage over the well understood models. Furthermore, deep learning is not typically an attractive approach to this type of analysis as there is a lack of one-to-one mapping between model parameters and dynamic PET data, so a single time-activity curve could be described by various combinations of

parameter estimates.

However, standard PK models are designed to describe PET data acquired continuously from injection to washout and are thus not ideal for adaptation to PET data acquired after injection with ASL-CBF maps to compensate. Furthermore, the incorporation of ASL-CBF constrains the parameter estimation, simplifying the mapping between the time activity curve and the parameter estimates.

In this work deep learning was chosen for its versatility; it is able to model the underlying relationship between ASL-CBF and the delivery, binding and washout of the PET tracer without explicit feature extraction. By exploiting all of the PET and MRI information, and avoiding voxelwise fitting, this framework could potentially provide more robust estimates of target density with a shorter acquisition time.

9.2 Methods

9.2.1 Deep Learning Framework for BP_{ND} Estimation (SS-DL)

The SS-DL performs regression of BP_{ND} from dynamic PET and MRI data directly. The network was implemented in NiftyNet (Gibson et al., 2018) using the ‘highresnet’ convolutional neural network described in 8.2.3.4. All inputs were 3D image volumes: the ASL-CBF maps, the structural T1 weighted MRI, and the each frame of the dynamic PET data, as shown in figure 9.1. The SS-DL was tested on various different acquisition window lengths and timings which determine the number of PET frames used, such that a different model was trained for each acquisition window. The model hyperparameters were set as in section 8.2.3.4 and all the models were trained for 250,000 iterations using early stopping to select the optimal model for testing.

9.2.2 Comparative BP_{ND} Estimation Methods

Short Acquisition Methods. The SS-DL method was compared to the RT-SRTM using the R_1 synthesis methods evaluated in chapter 8, namely linear regression (LR), regional linear regression (rLR), image fusion (IF), deep learning (DL) and true R_1 (R_1^*). The DL+ method used in the previous chapter, which included a PET frame as an input, was excluded here as this would require a new R_1 estimation to be generated for each timing window evaluated. Since the upper limit of RT-SRTM accuracy is demonstrated using the true R_1 (T), inclusion of the DL+ method was deemed unnecessary. The clinical standard, $SUVR-1$, was also included for comparison and was calculated as described in section 8.3.2.

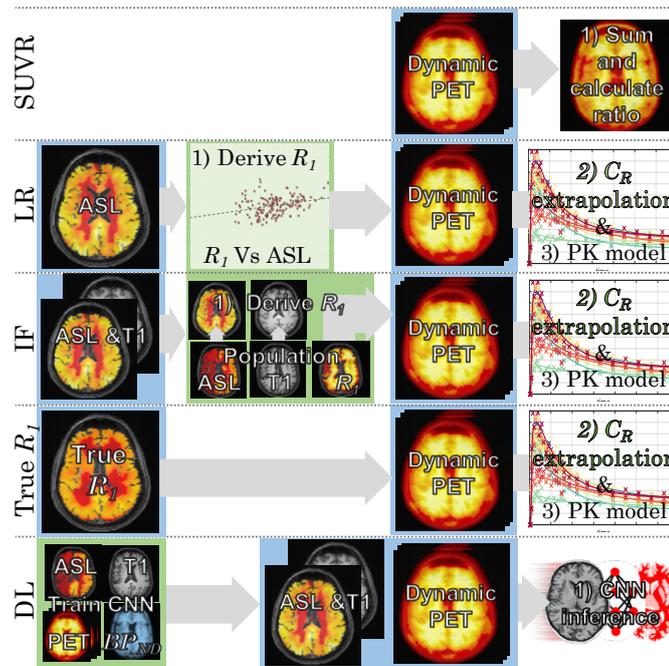


Figure 9.1: Overview of methods tested where true R_1 is R_1^* . Blue boxes indicate input data from the test subject and green boxes show population data from the training database.

Gold Standard BP_{ND} estimation. As described in section 5.2.1 the linearised simplified reference tissue model (SRTM) is used for gold standard R_1 (R_1^*) and BP_{ND} (BP_{ND}^*) quantification Gunn et al. (1997), using dynamic PET data acquired from $t=0,60$ minutes. Statistical analysis was performed

between SS-DL, RT-SRTM and $SUVR-1$ and the gold standard BP_{ND}^* was performed as described in section 8.3.5.

9.2.3 Acquisition window definition

For gold standard PK modelling the scan starts at tracer injection, $t_s=0$, with a duration of $t_d=60$ minutes. However, for the short acquisition methods $t_s>0$, and t_d is chosen to fit clinical requirements. The timing window, $t = t_s:t_s+t_d$, was optimised separately for each RT-SRTM method, SS-DL and $SUVR-1$ at different t_d 's. This was performed from $t_s > 20$ minutes, as this period is recommended for routine clinical scans using $[^{18}\text{F}]$ -florbetapir.

9.2.4 Data

The same datasets analysed in the previous chapter (chapter 8) was also used here, where a full description of the data used can be found in 8.3.3. In this chapter, one BOCAA subject was excluded from testing as it was suspected that their pathology was affecting the reference region, see section 8.4.2, and the impact of this on the quantification has already been assessed. Briefly, data from 74 subjects participating in three separate studies were included: 57 were cognitively normal subjects participating in Insight 46 (Lane et al., 2017) (1946), 4 were clinically diagnosed with probable Young Onset Alzheimer's disease (YOAD) and 13 were participating in the biomarkers and outcomes in cerebral amyloid angiopathy study (BOCAA) where 5 were controls and 8 were patients. Details of database construction and data pre-processing can also be found in chapter 8 along with the division of subjects into 5 folds for cross-validation.

9.3 Results and discussion

9.3.1 Method Comparison Over Different Timing Windows

To assess the performance of the different methods (SS-DL, RT-SRTM and $SUVR-1$) as a function of scan start time (t_s) and duration (t_d), the average

MSE in BP_{ND} estimation was calculated using one of the 5 folds ($n = 15$). The errors for each combination of t_s and t_d were calculated across all brain voxels, as shown in figure 9.2, and for grey matter voxels only in figure 9.3.

For both the whole brain and grey matter only analysis, the RT-SRTM methods (LR, rLR, IF, DL, R_1^*) show a strong dependence on t_d as the length of the scan determines the number of data points available for both reference region extrapolation and for fitting the pharmacokinetic model. Even in the best case, where R_1 has been estimated perfectly (R_1^* RT-SRTM, see figures 9.2e and 9.3e) the method requires 25-30 minutes of PET data to estimate BP_{ND} with reasonable accuracy. This demonstrates that the RT-SRTM is only suitable for patients who can tolerate a scan of approximately half an hour. The results also show that shifting t_s to a later time improves the BP_{ND} estimate as the later data contains more information about the binding and washout of the tracer, and as R_1 is set prior to model fitting the early PET data does not provide new information.

For the SS-DL method, the results in figures 9.2f and 9.3f show that the method produces BP_{ND} estimates with a low MSE across all permutations of t_s and t_d . The method demonstrates a minimal dependence on the length and timing of the scan which allows flexibility in the scheduling of the scan, and could mean that the patient can be scanned and discharged more quickly. For instance, if a 10 minute scan were performed 20 minutes after injection then the patient could potentially leave 30 minutes after the injection. This would be advantageous both for the patient and for the department where the patient is scanned.

The fact that increasing the number of PET frames available does not greatly reduce the MSE in the BP_{ND} estimate makes the SS-DL more versatile, but it also indicates that the model is not taking full advantage of the available data. This is because the PET frames are input as independent images and the model has no information about their ordering and the fact that they are a consecutive temporal series. Consequently the PET frames appear similar to

one another and the model does not use extra frames to improve the estimate of BP_{ND} . The frame independent behaviour shown here would be desirable for correcting a static PET image (or $SUVR-1$) for blood flow induced bias, however it does not exploit all of the information available in this dynamic dataset.

For the $SUVR-1$ method, the results for all brain voxels in figure 9.2g shows the trade off between the white matter voxels, in which an earlier acquisition is preferable as the binding is overestimated when the tracer is in a steady-state, and grey matter voxels, where a later acquisition when the tracer is in a steady-state is preferable (see figure 9.3g). This highlights that whole brain quantification using $SUVR-1$ requires either a compromise on acquisition timing between the tissue types, or separate acquisitions for grey or white matter. Whilst in general it is the grey matter that is of interest for this tracer, this may not be the case for other tracers with similar properties, and also have implications for regions at the boundary between grey and white matter, and regions in which the tissues are mixed due to PVE.

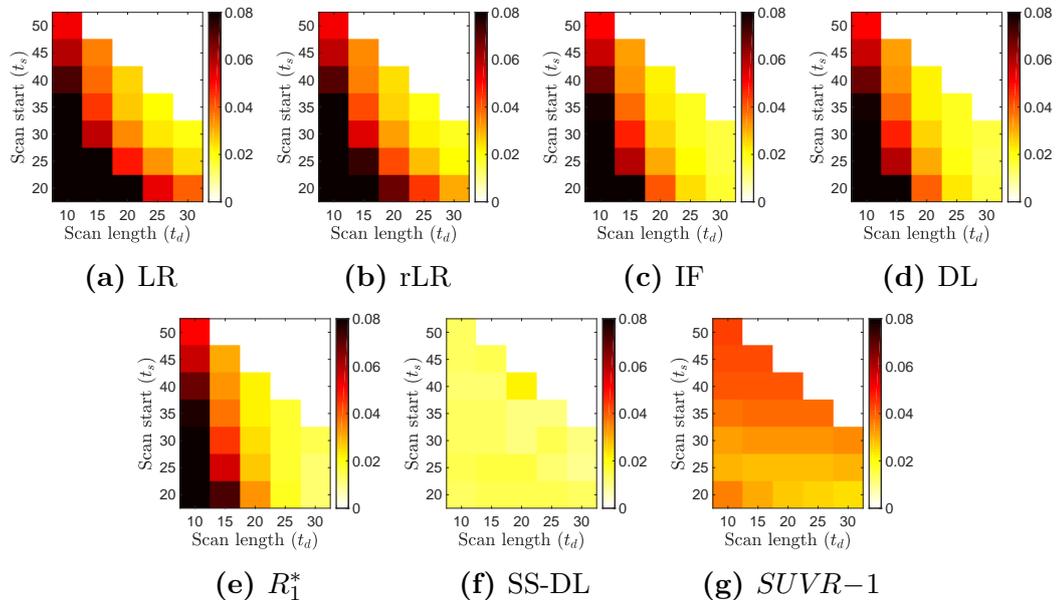


Figure 9.2: MSE for whole brain at different timing windows.

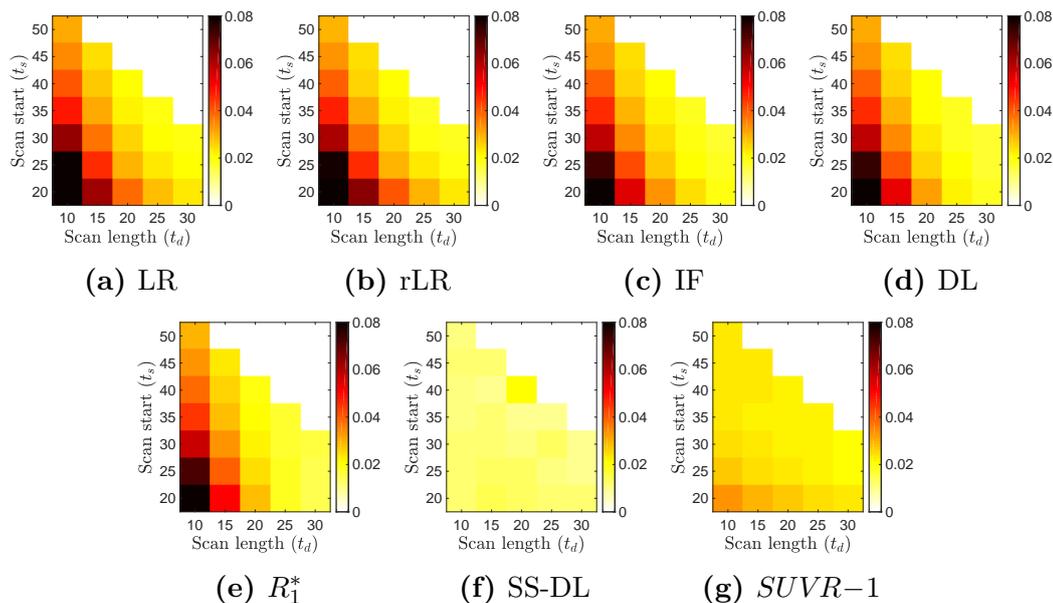


Figure 9.3: MSE for GM voxels at different timing windows.

9.3.2 Optimised Timing Window Method Comparison

To quantitatively compare the different methods, two different scan durations were selected, prior to data processing, for in-depth analysis. A t_d of 30 minutes was used for comparison with previous chapters and to represent a typical MRI scan protocol length for dementia. A shorter t_d of 15 minutes was also included to represent a scan length which would be tolerable for most patients whilst allowing enough time to acquire essential MRI data. Essential MRI data includes a high resolution 3D T1-w image (5 minutes), data for attenuation correction (either 3D T2-w for pCT (Burgos et al., 2013) 5 minutes, or UTE 2 minutes), and ASL (5 minutes), see appendix A.2.

Table 9.1 shows the optimal start times (t_s) for each method for a scan duration t_d of either 15 or 30 minutes. The t_s values were selected as the times which gave the lowest average MSE across the single fold of data used to generate figures 9.2 and 9.3 and was defined both for all brain voxels and for GM voxels alone. Table 9.1 shows the importance of optimising the timing window by method, as the 20:50 minute window selected in chapter 7 only applies to the R_1^* RT-SRTM.

For most methods the start times for all brain voxels and for GM alone

| Method | $t_d=30$ mins | | $t_d=15$ mins | |
|-----------------|---------------|----|---------------|----|
| | WB | GM | WB | GM |
| LR RT-SRTM | 30 | 30 | 45 | 45 |
| rLR RT-SRTM | 30 | 30 | 45 | 45 |
| IF RT-SRTM | 25 | 30 | 45 | 45 |
| DL RT-SRTM | 25 | 25 | 45 | 45 |
| R_1^* RT-SRTM | 20 | 20 | 45 | 45 |
| SS-DL | 25 | 25 | 40 | 40 |
| $SUVR-1$ | 20 | 30 | 25 | 35 |

Table 9.1: Start time (t_s) in minutes since injection which gives the lowest MSE for a given scan duration (t_d) for each method.

are the same or similar. However, as noted in the previous section, the optimal timing for $SUVR-1$ is strongly tissue dependent such that the optimal t_s for grey matter is different from that for the whole brain.

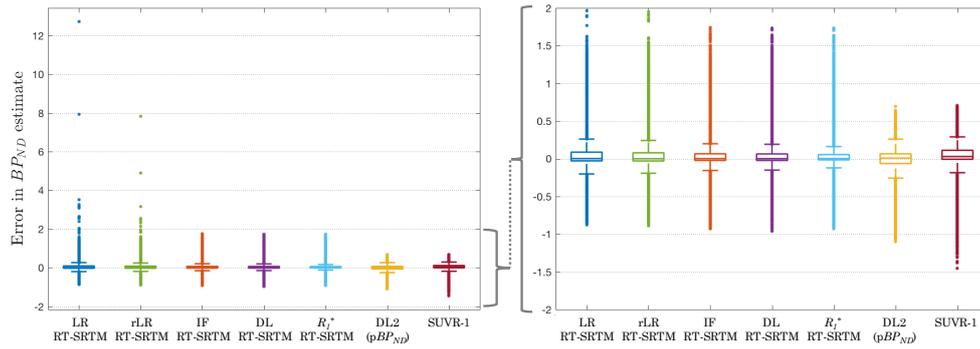
Conversely, when comparing the optimal t_s between the acquisition lengths, the start time does not vary greatly for $SUVR-1$, whereas the other methods (SS-DL and RT-SRTM) all shift much later. This is likely to be due to the fact that these methods all incorporate the blood flow information from the ASL data and as such are minimally reliant on the early part of the data, and the washout information later on is more important.

The optimised timings for a 30 minute and 15 minute scan shown in table 9.1 were used to generate the results for the following sections.

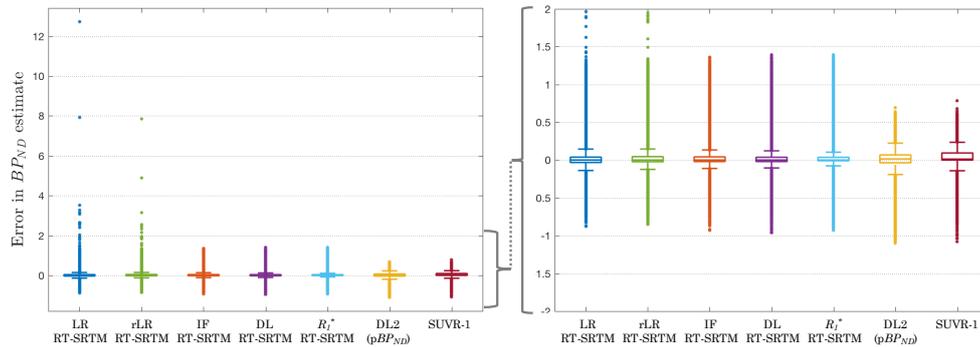
9.3.2.1 30 minute optimised acquisition.

Table 9.2 shows the MSE across subjects for the 30 minute acquisition window. Comparison with table 8.5 shows that extending the acquisition time for $SUVR-1$, and optimising the timing window for each RT-SRTM technique reduces the MSE of these methods, although the bias in $SUVR-1$ is increased. This was considered to be reasonable since the primary metric used to compare performance against other methods here is MSE, and the optimal timing of all methods was selected using MSE.

Table 9.2 shows that the $SUVR-1$ method has a higher mean MSE and bias compared to the other techniques, both across the whole brain and grey



(a) All brain voxels (DL2=SS-DL)



(b) GM voxels only (DL2=SS-DL)

Figure 9.4: Boxplots of voxelwise error using 30 minutes PET data.

matter only. When performing individual comparison between the techniques then this difference is statistically significant in all cases ($p < 0.001$), however for the whole brain the difference between LR and $SUVR-1$ does not persist when corrected for multiple comparisons and the methods are statistically equivalent. The rLR method is no longer out performed by $SUVR-1$, due to the timing window optimisation which was not performed in chapter 8, suggesting that voxelwise analysis using this approach is feasible. The SS-DL method performs equivalently to the rLR RT-SRTM method in terms of MSE, but statistically worse than the IF and DL approaches ($p < 0.017$).

When the SS-DL was initially implemented using only ‘healthy’ subjects in Scott et al. (2018c), it performed equivalently to the best RT-SRTM method. The reduction in performance seen in this chapter is likely to be due to the increased variability in the dataset and the fact that the training set is now unbalanced, with far more healthy subjects than diseased. Furthermore, the

| Method | WB voxels | | GM voxels | |
|-----------------|---------------|---------------|---------------|---------------|
| | MSE | ME | MSE | ME |
| LR RT-SRTM | 0.0131 | 0.0130 | 0.0087 | -0.0083 |
| rLR RT-SRTM | 0.0118 | 0.0079 | 0.0080 | 0.0007 |
| IF RT-SRTM | 0.0087 | 0.0061 | 0.0072 | 0.0000 |
| DL RT-SRTM | 0.0082 | 0.0045 | 0.0069 | -0.0038 |
| R_1^* RT-SRTM | 0.0054 | 0.0087 | 0.0046 | 0.0072 |
| SS-DL | 0.0129 | 0.0007 | 0.0116 | 0.0112 |
| $SUVR-1$ | 0.0182 | 0.0372 | 0.0123 | 0.0389 |

Table 9.2: Voxel mean squared error (MSE) in BP_{ND} estimates averaged over 74 subjects using all brain voxels (WB) and grey matter (GM) voxels only using 30 minutes of PET data.

BOCAA subjects contribute less than 15% of the subjects and are very different from the other datasets due to the additional pathologies, such as strokes, which affect this cohort. This makes it difficult for the network to learn an accurate model which can be applied to all the data included as the loss is calculated over all the training data which is dominated by healthy subjects. Training set data augmentation could be used to rebalance the dataset prior training to address this issue. The only data augmentation currently available in NiftyNet (the platform on which the deep learning architectures are implemented) is performed ‘on-the-fly’ and has no capabilities to rebalance the dataset as all images have an equal probability of being augmented. The database imbalance has less of an influence on the RT-SRTM methods as they utilise a physiological model which can accurately be applied to all subjects.

The bias for the SS-DL method appears to be lower than the RT-SRTM methods, but this does not reach statistical significance ($p \geq 0.905$). Figure 9.4a shows that the SS-DL method does, however, produce a lower variance in the error. It should be noted that figure 9.4a shows the error for each voxel used in the analysis (approximately 17 million points for whole brain analysis), therefore the variation in the RT-SRTM methods appears much worse, due to a small proportion of outliers.

9.3.2.2 15 minute optimised acquisition.

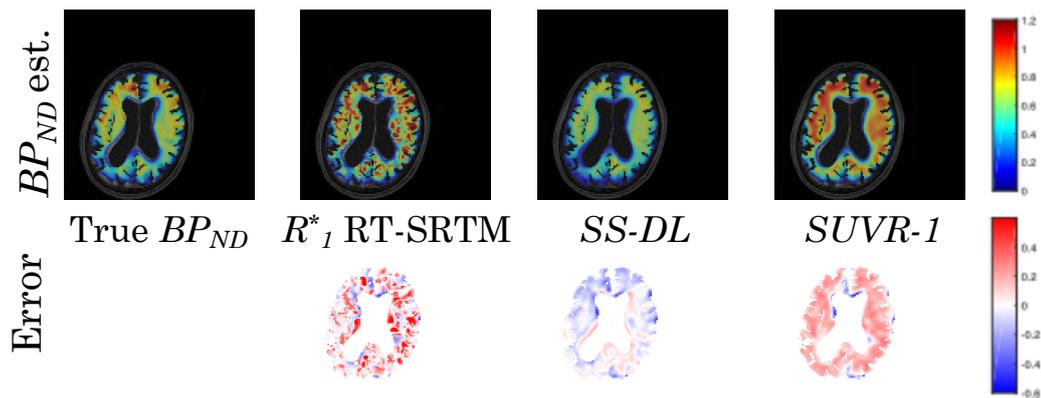
When the scan time is reduced to 15 minutes, where just 3 PET frames are available, the MSE for the RT-SRTM methods increases, as shown by comparison of table 9.2 with table 9.3. The increase in BP_{ND} estimation error, which occurs even when using the true R_1 parameter (R_1^*), is due to the reduction in the number of frames available for the extrapolation of the reference region (C_R) and for kinetic modelling. By contrast, the SS-DL and $SUVR-1$ methods maintain their performance levels as these methods are less dependent on the number of PET frames available, as demonstrated in section 9.3.2.

For a 15 minute acquisition the SS-DL method has a significantly lower MSE and ME than all other methods ($p < 0.005$) across all brain voxels. This shows that this methodology is effective at combining the PET and MRI data to improve the estimate of BP_{ND} . This is demonstrated in figure 9.6a where the voxelwise error in BP_{ND} is shown to be reduced for the SS-DL method (yellow) with minimal bias.

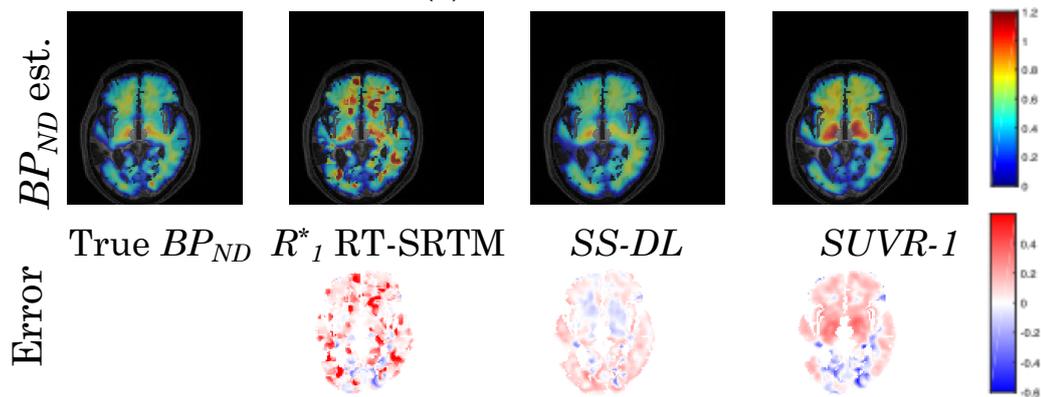
When including just grey matter voxels within the analysis, the SS-DL method has a significantly lower bias than all other techniques ($p \leq 0.020$) except LR ($p = 0.053$), with the most marked difference compared to $SUVR-1$ ($p < 0.001$). Whilst the SS-DL method has a lower mean MSE than all other methods, this difference is only significant when comparing individually to the RT-SRTM methods ($p \leq 0.015$), and not for $SUVR-1$ ($p = 0.292$) or when accounting for multiple comparisons.

The results in table 9.3 demonstrate that the SS-DL method is an effective method for estimating BP_{ND} and at short acquisition times can be considered to be a way of correcting $SUVR-1$ data. Whilst it still reduces bias, the SS-DL method does not bring such large improvements in accuracy when analysed for GM voxels alone compared to the whole brain. This is due to the fact that when the model is trained, an equal weight is given to all voxels within the brain, and table 9.1 shows that the same model is used for WB and GM. Since the brain contains predominantly white matter voxels (approximately twice as

many WM voxels as GM voxels in the data used here), the errors in this region contributes the most to the cost function. Consequently, if only grey matter voxels are of interest, the BP_{ND} estimation can be improved by adjusting the voxel weighting, such that the model training is mostly influenced by errors in grey matter regions.



(a) YOAD-057



(b) BOCAA-5

Figure 9.5: Voxelwise estimates of BP_{ND} for a 15 minute acquisition.

Figure 9.5 shows the estimated BP_{ND} for two example subjects, who were also analysed in chapter 8. Figure 9.5a shows a patient from the YOAD study and demonstrates that the RT-SRTM method, even when using the true R^*_1 , is corrupted by noise, due to the limited data available. By contrast, the SS-DL method is of a similar noise level to the true BP_{ND} image and is able to reproduce the increased binding in the anterior portion of the brain. As seen

| Method | WB voxels | | GM voxels | |
|-----------------|---------------|----------------|---------------|---------------|
| | MSE | ME | MSE | ME |
| LR RT-SRTM | 0.0309 | 0.0652 | 0.0161 | 0.0285 |
| rLR RT-SRTM | 0.0297 | 0.0621 | 0.0155 | 0.0325 |
| IF RT-SRTM | 0.0274 | 0.0603 | 0.0152 | 0.0322 |
| DL RT-SRTM | 0.0274 | 0.0602 | 0.0150 | 0.0308 |
| R_1^* RT-SRTM | 0.0266 | 0.0589 | 0.0145 | 0.0323 |
| SS-DL | 0.0137 | -0.0027 | 0.0121 | 0.0075 |
| $SUVR-1$ | 0.0217 | 0.0455 | 0.0139 | 0.0442 |

Table 9.3: Voxel mean squared error (MSE) in BP_{ND} estimates averaged over 74 subjects using all brain voxels (WB) and grey matter (GM) voxels only, for 15 minutes of PET data

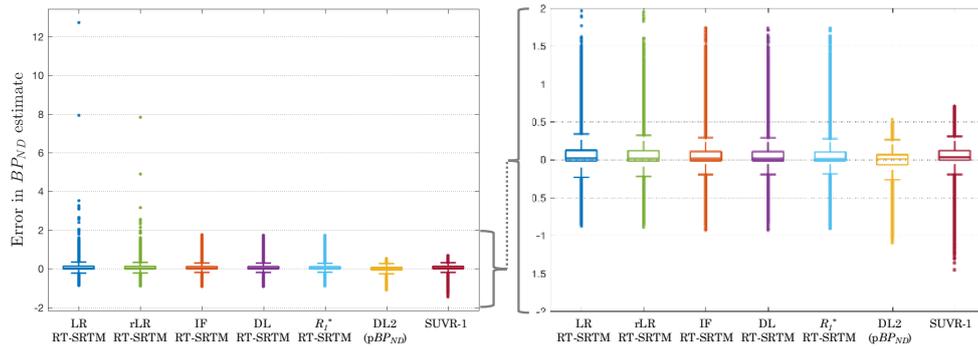
previously, the $SUVR-1$ method leads to an overestimation of the tracer binding across the brain. Similarly, figure 9.5b shows a patient from the BOCAA study, which again demonstrates that the RT-SRTM method is dominated by noise, $SUVR-1$ leaves to an overestimation, and SS-DL produces a reasonable estimate of the BP_{ND} .

9.4 Conclusions

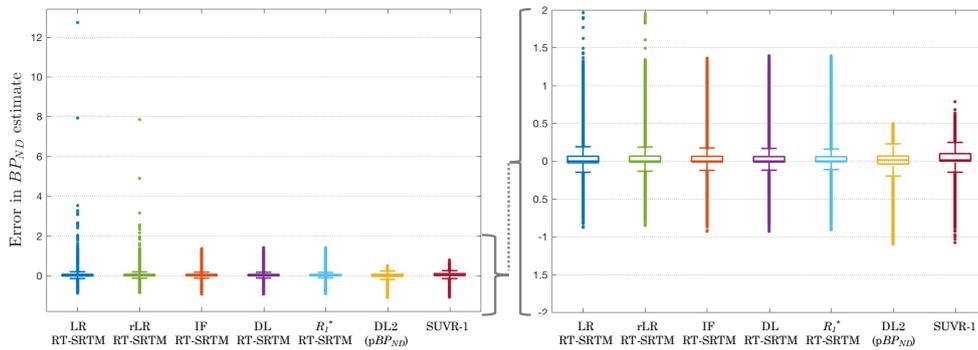
This chapter evaluates the SS-DL; a deep learning approach to PET BP_{ND} estimation, which combines dynamic PET data with MRI blood flow and structural images in a single step. The SS-DL method performed consistently better than the clinical standard $SUVR-1$, showing that deep convolutional neural networks have the potential to be useful for PK modelling for the first time.

For a 30 minute PET-MR acquisition there is minimal benefit in using the SS-DL method over the RT-SRTM as the RT-SRTM generally performs better, although the RT-SRTM has less flexibility in the acquisition start time which could affect clinical implementation.

However, when the acquisition time is reduced to 15 minutes, there is not sufficient data to fit the RT-SRTM and the performance is worse than that of the $SUVR-1$ method. By contrast the SS-DL method maintains similar performance to a 30 minute acquisition, producing significantly more accurate results than $SUVR-1$ across the whole brain, and has a significantly



(a) Including all brain voxels (approx. 17 million)



(b) Including grey matter voxels only (approx. 8 million voxels)

Figure 9.6: Voxelwise error using 15 minutes PET data.

lower bias in grey matter voxels. These results show that the SS-DL approach has the potential to be an improved alternative to $SUVR-1$ which can be acquired within a comparable time frame. However, the work performed here was preliminary proof-of-concept work and there are several changes which could be made to significantly improve performance.

The simplest way to improve SS-DL performance could be to optimise grey matter voxels separately when applying grey matter only analysis, such that parameter optimisation during training is not dominated by white matter voxels. Furthermore, rebalancing of the training data, either through data augmentation or through patient selection, could greatly improve the generalisation of the model.

Finally, the network architecture of the SS-DL was not adapted for input PET frames as a time series, and the fact that the PET frames are sequential is ignored. recurrent neural networks (RNN)s are designed for sequential data,

and deep RNNs have been shown to be useful in speech recognition (Graves et al., 2013) with promising results in medical imaging (Litjens et al., 2017). As such deep RNNs could potentially be adapted and applied to this task to account for the temporal ordering of the PET data and consequently improve the BP_{ND} estimation.

Chapter 10

General Conclusions

10.1 Summary

PET can provide quantitative measurements of amyloid- β *in vivo*, however this requires dynamic data to be acquired from tracer injection for 60 minutes or more to allow accurate fitting of standard PK models. The simplified estimate of amyloid- β estimation, SUVR, only requires a 10 minute static scan, but it is biased by changes in blood flow and the violation of the tracer equilibrium assumption. Consequently, there is no methodology which can produce sufficiently robust and accurate estimates of amyloid- β burden within a clinically feasible time frame for longitudinal studies.

The aim of this work was to exploit the availability of simultaneously acquired MRI data from a PET-MR scanner to reduce the acquisition time required for accurate quantification. To do this new models were required to incorporate blood flow information measured using ASL-MRI with the dynamic PET data. The main approach developed in this thesis adapted the established SRTM such that the early phase of the dynamic PET data was no longer needed, resulting in the reduced acquisition time SRTM (RT-SRTM). The RT-SRTM is applied in three stages 1) PET- R_1 is estimated from ASL-CBF, 2) the reference region TAC (C_R) is extrapolated back to $t = 0$ at tracer injection to estimate the missing early phase of the data, then 3) the RT-SRTM is fitted to the measured dynamic PET data using fixed R_1 and extrapolated

C_R .

The RT-SRTM was initially implemented on a regional basis where average parameter values are estimated across a region defined by the tissue parcellation. Here, linear regression on a population of PET- R_1 ASL-CBF pairs was used to estimate the subject R_1 , and the mean population C_R was scaled to the patient's C_R to extrapolate the reference region before fitting the RT-SRTM (see chapter 5 for details). This demonstrated that reasonable BP_{ND} estimates could be derived using a 30 minute PET-MR acquisition with the RT-SRTM with a significantly lower bias and error than using SUVR. Each step of the RT-SRTM was then optimised in chapters 6 and 7. For step 1) linear regression using region as a co-variate was found to be more suitable for R_1 extrapolation as the relationship between PET- R_1 and ASL-CBF is region dependant. For step 2) Principal component analysis (PCA) was found to be better for extrapolating C_R than scaling the population mean value. Finally, for step 3) it was found that fitting data acquired at $t = 20, 50$ minutes post injection gave the lowest MSE, assuming a perfect estimation of R_1 . The $t = 20, 50$ minute time frame provided enough early data to get a good estimate of C_R , whilst leaving enough late data to estimate BP_{ND} . The optimised RT-SRTM using a 30 minute PET-MR acquisition was able to produce BP_{ND} estimates with excellent correlation to the gold standard ($r = 0.896$), with significantly lower MSE and mean error (ME) ($p < 0.001$) which were not correlated with blood flow (R_1^*).

The RT-SRTM was then extended to voxelwise analysis in chapter 8. Here two other methods were introduced and evaluated for the estimation of R_1 , which were selected to avoid artefacts from the ASL-CBF estimates propagating into the R_1 estimate. Here it was demonstrated that multi-atlas propagation with image fusion and deep learning based regression can be used to generate more accurate voxelwise estimates of R_1 than linear regression, which led to significantly improved estimates of BP_{ND} .

Whilst chapter 8 demonstrated that the RT-SRTM can provide excellent

BP_{ND} estimates for a 30 minute acquisition, the need for C_R extrapolation and voxelwise curve fitting makes it difficult to further reduce the acquisition time without increasing the error in BP_{ND} . In chapter 9 the SS-DL method was proposed, where a deep CNN was used to learn the relationship between the input PET-MR data and the output BP_{ND} parameter map. For a 30 minute acquisition the SS-DL method was statistically equivalent to the RT-SRTM, as there was adequate data for C_R extrapolation and curve fitting. However, when the acquisition time was reduced to 15 minutes, the RT-SRTM method broke down, as there was insufficient data to fit the model. By contrast, the SS-DL method maintained its performance, generating significantly better estimates of BP_{ND} than SUVR. The SS-DL method has the potential to provide a direct alternative approach to current clinical practice with improved accuracy and reduced bias.

10.2 Future Research Directions

The work contained in this thesis produced two potential approaches to exploiting simultaneous PET-MR acquisition to reduce the acquisition time required for accurate quantification. It represents a proof of concept which demonstrates that both novel methods developed, the RT-SRTM for a 30 minute acquisition and the SS-DL for a 15 minute acquisition, could be used for longitudinal studies. However, longitudinal data wasn't available at the time to validate this claim and consequently cross-sectional validation was performed instead. The 1946 study, used extensively in this work, is currently acquiring a second time-point approximately 2 years after the first which provides an excellent resource for longitudinal validation in the future.

Furthermore, there were several factors which limited the assessment of the generalisability of the techniques. Firstly, there was a paucity of high quality data from a diseased population and the data from only 4 clinically diagnosed subjects participating in the YOAD study were included. The YOAD study has now concluded and many more datasets will be made available,

which will better balance the databases used in both the RT-SRTM and SS-DL methods and allow a more detailed assessment of the performance on the techniques on these subjects.

Finally, an assessment of generalisability is limited by the fact that the work was all carried out on a single PET-MR scanner (Siemens mMR at UCLH) with one tracer ($[^{18}\text{F}]$ -florbetapir) and one imaging protocol (60 minute dynamic PET with pCASL ASL 3D GRASE readout). Thanks to the commissioning of various new PET-MR systems during the course of this work, there is the potential for multi-centre studies such as the dementias platform UK for validation, and harmonisation between systems is already underway¹.

In theory both the RT-SRTM and SS-DL method could be applied to other tracers, such as those that target tau or other amyloid tracers. For such tracers the acquisition time required for quantification is often much longer than that for $[^{18}\text{F}]$ -florbetapir, which provides an extra incentive to applying reduced acquisition techniques. However this does also make it more difficult to find full dynamic data for gold standard comparisons and for generating a database, so this is suggested as future work.

For both the RT-SRTM and SS-DL methods, using deep CNNs to learn the model instead of imposing a model based on flawed assumptions was found to be advantageous. However, the CNN architecture used for these tasks was not tailored towards the problem being tackled. The use of an existing architecture was ideal for quickly demonstrating that the approach had potential, however better results could be achieved using architectures better designed for the task. A possible improvement for the SS-DL methods could be to take into account the temporal ordering of the PET frames, as is done explicitly in standard PK modelling. Further work developing a recurrent neural network suitable for the task could lead to a robust and generalisable method which can directly substitute for SUVR whilst maintaining quantitative accuracy.

¹<https://www.dementiasplatform.uk/news-and-blogs/news/dpuk-brain-imaging-harmonisation-study-launch>

Appendix A

Additional data information

A.1 Data sets used

The candidate took no part in recruiting or scanning the subjects enrolled in the studies used, however performed extensive data processing, see section B.

A.1.1 Insight46 (1946)

Imaging data were collected from 462 healthy volunteers participating in Insight 46, a neuroimaging sub-study of the Medical Research Council National Survey of Health and Development (Lane et al., 2017). This study was conducted in line with the principles of the Declaration of Helsinki and ethical approval was obtained from the National Research Ethics Service (ref 14/LO/1173). Written informed consent was obtained from all participants. Across the whole cohort the mean age at scan was 70.7 years (range 69.3-71.9).

All subjects underwent 60 minutes of simultaneous amyloid PET and multi-modal MR imaging on a Siemens Biograph mMR PET/MR scanner. Of the 60 subjects used in this work, mean age 69.8 years (range 69.3-70.4 years), 56 had both PET and ASL data, and for 4 the ASL data were missing due to repetition of other scans. All subjects had an MMSE \geq 27 (mean = 29.3).

A.1.2 Young Onset Alzheimer's Disease (YOAD)

10 subjects from a study of young onset Alzheimer's Disease (YOAD) with an intermediate or high certainty diagnosis (McKhann et al., 2011) were scanned, 4

of which had both list mode PET and ASL data available and were used in this work. For these subjects the mean age at scan was 67.4 years (range 61.7-70.5), one of which was diagnosed with posterior cortical atrophy. This study was conducted in line with the principles of the Declaration of Helsinki and ethical approval was obtained from the London Queen Square Ethics Committee (ref 15/LO/1412). Written informed consent was obtained from all participants.

A.1.3 Biomarkers and Outcomes in Cerebral Amyloid Angiopathy (BOCAA)

15 subjects were scanned for the biomarkers and outcomes in cerebral amyloid angiopathy (BOCAA) pilot study, where cerebral amyloid angiopathy (CAA) is a small vessel disease associated spontaneous intracerebral haemorrhage. This study included 10 CAA patients (mean age 68.6 years, SD 3.0 years), who all had lobar microbleeds, and 5 age matched healthy controls (mean age 64.4 years, SD 6.5 years). However, list mode PET data were incomplete for one CAA patient consequently 14 subjects were available for use. Ethical approval for this study was granted by the NHS Health Research Authority London- Dulwich Research Ethics Committee (REC reference 15/LO/1443). Written informed consent was obtained from all participants.

A.2 Data protocols

For all three data-sets described in section A.1, all subjects underwent 60 minutes of simultaneous amyloid PET and multi-modal MR imaging on a Siemens Biograph mMR PET-MR scanner. List mode PET data were acquired for 60 minutes following intravenous injection of [^{18}F]-florbetapir, which targets amyloid- β . Figure A.1 shows the order in which the simultaneous PET/MR data were acquired and the duration of each acquisition. The ASL acquisition is performed last, approximately 55 minutes into the scan, and as such if any previous scans need to be repeated then the ASL is not acquired. This ensures that the subject does not exceed 60 minutes on the scanner. More details can be found in Lane et al. (2017).

The only dataset which differs from this protocol is the BOCAA study, which replaces the T2-w MRI with task based fMRI. This means that the pCT is derived using the T1-w MRI only, and it is assumed that the task itself does not alter the blood flow in the brain significantly enough to cause a measurable difference to the PET and ASL data.

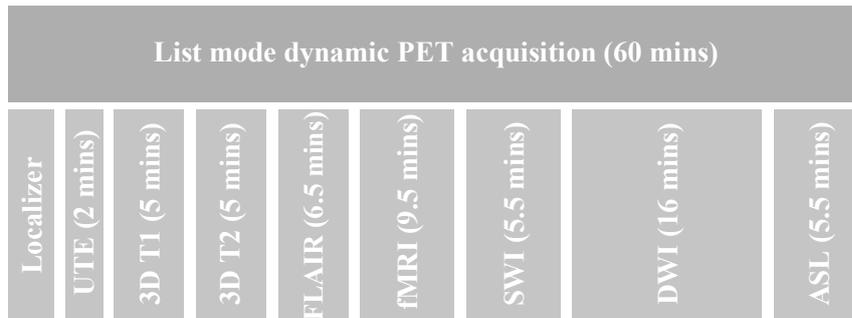


Figure A.1: Imaging protocol showing simultaneous PET and multi-parametric imaging.

A.3 Subject amyloid status

The analysis in this section was performed by Cash et al. (2017). To determine whether subjects were globally amyloid positive, mean cortical grey matter SUVR values were calculated with both whole cerebellum and eroded white matter reference regions. To determine the cut point, which defines whether a subject is positive or negative, a Gaussian mixture model was fitted to the SUVR of 462 subjects from the Insight 46 study for each reference region. The cut-points along with the SUVR information for the subjects used in the study can be found in table A.1. In this work, SUVR using the whole cerebellum as the reference region was used to define whether a subject was amyloid positive or negative. The cut-points defined using the Insight46 data were also applied to the YOAD and BOCAA data.

| | Whole cerebellum | Eroded white matter |
|------------|------------------|---------------------|
| Cut-point | 1.077 | 0.610 |
| Mean | 1.016 | 0.576 |
| Min | 0.791 | 0.469 |
| Max | 1.740 | 0.892 |
| $a\beta^+$ | 11 (17%) | 12 (19%) |

Table A.1: Summary of mean cortical grey matter SUVR values in the dataset used for two different reference regions with the number of subjects defined as amyloid positive ($a\beta^+$).

Appendix B

Data processing

B.1 PET reconstruction

Excluding the data used in chapter 5 (where vendor software was used on some datasets were reconstructed by Jieqing Jiao) all PET reconstruction was carried out by the candidate as follows; the data were reconstructed using NiftyPET (Markiewicz et al., 2017) which is an open source reconstruction platform. The software was developed by Pawel Markiewicz over the time course of this PhD to support dynamic data reconstruction, where the candidate was a test user. Whilst many software versions were developed during this time, the parameters used were kept the same to maintain consistency between reconstructed datasets.

The PET data were reconstructed as follows: the pseudo CT (pCT) μ -map, generated as in section B.3, was propagated into PET space by rigidly registering the T1-weighted images to a full 60-minute non-attenuation corrected reconstructed PET image and applying the transformation to the pCT. Dynamic PET data were binned into 31 time frames ($15\text{s} \times 4$, $30\text{s} \times 8$, $60\text{s} \times 9$, $180\text{s} \times 2$, $300\text{s} \times 8$), and reconstructed into $2 \times 2 \times 2\text{mm}$ voxels. An ordered subset expectation maximisation (OSEM) algorithm was used with 4 iterations, 14 subsets, with corrections for dead-time, attenuation, scatter, randoms and normalisation. These reconstruction parameters were selected for quantitative accuracy as validated in Markiewicz et al. (2017).

B.2 ASL quantification

All ASL quantification was carried out by the candidate using the open source software using NiftyFit (Melbourne et al., 2016) as follows; S_0 maps were estimated by fitting saturation recovery images acquired with the same sequence at three different saturation times (1,2,4s). CBF maps were computed with equation (B.1) (Buxton et al., 1998)

$$\text{CBF} = \frac{6000\lambda \Delta S}{2\alpha S_0} \frac{e^{\text{PLD}/T1_{\text{blood}}}}{T1_{\text{blood}}(1 - e^{-\tau/T1_{\text{blood}}})} [\text{ml}/100\text{g}/\text{min}] \quad (\text{B.1})$$

with 0.9 ml/g for the plasma/tissue partition coefficient (λ), a blood T1 of 1650 ms ($T1_{\text{blood}}$), and a labelling efficiency of 0.85 (α) as recommended in the ASL consensus paper (Alsop et al., 2015).

Quality control on each subject was performed by the candidate, where motion and artefacts were identified. In this work no motion correction or registration between images was performed.

B.3 pCT generation

In this thesis most pCT based μ -maps were generated using an automated pipeline set up by Ninon Burgos, however the pCTs for all BOCAA and YOAD subjects were generated by the candidate. pCT based μ -maps were generated as described in Burgos et al. (2015) using multi-atlas propagation with image fusion followed by bilinear mapping to convert CT images from Hounsfield units to attenuation coefficients.

B.4 Pharmacokinetic modelling

All kinetic modelling software used in this thesis was developed by the candidate based on an in-house MATLAB implementation of the SRTM with basis functions following the description in Gunn et al. (1997). For all implementations of SRTM used in this work, 100 basis functions were generated with $15e^{-5} \leq \theta \leq 0.1s^{-1}$.

Bibliography

David C. Alsop, John A. Detre, Xavier Golay, Matthias Günther, Jeroen Hendrikse, Luis Hernandez-Garcia, Hanzhang Lu, Bradley J. Macintosh, Laura M. Parkes, Marion Smits, Matthias J P Van Osch, Danny J J Wang, Eric C. Wong, and Greg Zaharchuk. Recommended implementation of arterial spin-labeled Perfusion mri for clinical applications: A consensus of the ISMRM Perfusion Study group and the European consortium for ASL in dementia. *Magnetic Resonance in Medicine*, 73(1):102–116, 2015. ISSN 15222594. doi: 10.1002/mrm.25197.

Isadora L Alves, Antoon TM Willemsen, Rudi A Dierckx, Ana Maria M da Silva, and Michel Koole. Dual time-point imaging for post-dose binding potential estimation applied to a [^{11}C]raclopride PET dose occupancy study. *Journal of Cerebral Blood Flow & Metabolism*, 37(3):866–876, 2017. ISSN 0271-678X. doi: 10.1177/0271678X16644463. URL <http://journals.sagepub.com/doi/10.1177/0271678X16644463>.

D. L. Bailey, B. J. Pichler, B. Gückel, G. Antoch, H. Barthel, Z. M. Bhujwala, S. Biskup, S. Biswal, M. Bitzer, R. Boellaard, R. F. Braren, C. Brendle, K. Brindle, A. Chiti, C. la Fougère, R. Gillies, V. Goh, M. Goyen, M. Hacker, L. Heukamp, G. M. Knudsen, A. M. Krackhardt, I. Law, J. C. Morris, K. Nikolaou, J. Nuyts, A. A. Ordonez, K. Pantel, H. H. Quick, K. Riklund, O. Sabri, B. Sattler, E. G.C. Troost, M. Zaiss, L. Zender, and Thomas Beyer. Combined PET/MRI: Global Warming—Summary Report of the 6th International Workshop on PET/MRI, March 27–29, 2017,

- Tübingen, Germany. In *Molecular Imaging and Biology*, volume 20, pages 4–20. Springer New York LLC, feb 2018. doi: 10.1007/s11307-017-1123-5.
- Dale L Bailey, Michael N Maisey, David W Townsend, and Peter E Valk. *Positron emission tomography*. Springer, 2005.
- S. L. Baker, S. N. Lockhart, J. C. Price, M. He, R. H. Huesman, D. Schonhaut, J. Faria, G. Rabinovici, and W. J. Jagust. Reference tissue-based kinetic evaluation of 18F-AV-1451 in aging and dementia. *Journal of Nuclear Medicine*, 5279:1–30, 2016. ISSN 0161-5505. doi: 10.2967/jnumed.116.175273. URL <http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.116.175273>.
- Randall J. Bateman, Chengjie Xiong, Tammie L.S. Benzinger, Anne M. Fagan, Alison Goate, Nick C. Fox, Daniel S. Marcus, Nigel J. Cairns, Xianyun Xie, Tyler M. Blazey, David M. Holtzman, Anna Santacruz, Virginia Buckles, Angela Oliver, Krista Moulder, Paul S. Aisen, Bernardino Ghetti, William E. Klunk, Eric McDade, Ralph N. Martins, Colin L. Masters, Richard Mayeux, John M. Ringman, Martin N. Rossor, Peter R. Schofield, Reisa A. Sperling, Stephen Salloway, and John C. Morris. Clinical and biomarker changes in dominantly inherited Alzheimer’s disease. *New England Journal of Medicine*, 367(9):795–804, aug 2012. ISSN 15334406. doi: 10.1056/NEJMoa1202753.
- Marije R. Benedictus, Annebet E. Leeuwis, Maja A A Binnewijzend, Joost P A Kuijjer, Philip Scheltens, Frederik Barkhof, Wiesje M. van der Flier, and Niels D. Prins. Lower cerebral blood flow is associated with faster cognitive decline in Alzheimer’s disease. *European Radiology*, pages 1169–1175, 2016. ISSN 14321084. doi: 10.1007/s00330-016-4450-z.
- Murat Bilgel, Lori Beason-Held, Yang An, Yun Zhou, Dean F Wong, and Susan M Resnick. Longitudinal evaluation of surrogates of regional cerebral blood flow computed from dynamic amyloid PET imaging. *Jour-*

- nal of Cerebral Blood Flow and Metabolism*, 2019. ISSN 15597016. doi: 10.1177/0271678X19830537.
- Gunnar Blomqvist, S Pauli, S Farde, L Eriksson, A Person, and C Halidin. *Dynamic models for reversible ligand binding*, volume 21. 03 edition, 1989. ISBN 0792302540.
- Ronald Boellaard, Paul Knaapen, Abraham Rijbroek, Gert J.J. Luurtsema, and Adriaan A. Lammertsma. Evaluation of basis function and linear least squares methods for generating parametric blood flow images using 15O-water and positron emission tomography. *Molecular Imaging and Biology*, 7(4):273–285, jul 2006. ISSN 18602002. doi: 10.1007/s11307-005-0007-2.
- Ajna Borogovac and Iris Asllani. Arterial spin labeling (ASL) fMRI: Advantages, theoretical constrains and experimental challenges in neurosciences. *International Journal of Biomedical Imaging*, 2012, 2012. ISSN 16874188. doi: 10.1155/2012/818456.
- Santiago Bullich, Henryk Barthel, Norman Koglin, Georg A Becker, Susan De Santi, Aleksandar Jovalekic, Andrew W Stephens, and Osama Sabri. Validation of Noninvasive Tracer Kinetic Analysis of 18 F-Florbetaben PET Using a Dual-ÅTime-Window Acquisition Protocol. *Journal of Nuclear Medicine*, 59(7):1104–1110, jul 2018. ISSN 0161-5505. doi: 10.2967/jnumed.117.200964. URL <http://jnm.snmjournals.org/lookup/doi/10.2967/jnumed.117.200964>.
- C Burger, G Goerres, S Schoenes, A Buck, A H R Lonn, and G K Von Schulthess. PET attenuation coefficients from CT images: experimental evaluation of the transformation of CT into PET 511-keV attenuation coefficients. *European journal of nuclear medicine and molecular imaging*, 29(7):922–7, jul 2002. ISSN 1619-7070. doi: 10.1007/s00259-002-0796-3. URL <http://www.ncbi.nlm.nih.gov/pubmed/12111133>.
- Ninon Burgos, Manuel Jorge Cardoso, Marc Modat, Stefano Pedemonte, John

- Dickson, Anna Barnes, John S. Duncan, David Atkinson, Simon R. Arridge, Brian F. Hutton, and Sebastien Ourselin. Attenuation correction synthesis for hybrid PET-MR scanners. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8149 LNCS(PART 1):147–154, 2013. ISSN 03029743. doi: 10.1007/978-3-642-40811-3_19.
- Ninon Burgos, M Jorge Cardoso, Kris Thielemans, Marc Modat, John Dickson, Jonathan M Schott, David Atkinson, Simon R Arridge, Brian F Hutton, and Sébastien Ourselin. Multi-contrast attenuation map synthesis for PET/MR scanners: assessment on FDG and Florbetapir PET tracers. *European journal of nuclear medicine and molecular imaging*, pages 1447–1458, 2015. ISSN 1619-7089. doi: 10.1007/s00259-015-3082-x. URL <http://www.ncbi.nlm.nih.gov/pubmed/26105119>.
- R B Buxton, L R Frank, E C Wong, B Siewert, S Warach, and R R Edelman. A general kinetic model for quantitative perfusion imaging with arterial spin labeling. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 40(3):383–396, 1998. ISSN 0740-3194. doi: 10.1002/mrm.1910400308.
- M. Jorge Cardoso, Marc Modat, Robin Wolz, Andrew Melbourne, David Cash, Daniel Rueckert, and Sebastien Ourselin. Geodesic Information Flows: Spatially-Variant Graphs and Their Application to Segmentation and Fusion. *IEEE Transactions on Medical Imaging*, pages 1976–1988, 2015. ISSN 0278-0062. doi: 10.1109/TMI.2015.2418298. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7086081>.
- R E Carson, M a Channing, R G Blasberg, B B Dunn, R M Cohen, K C Rice, and P Herscovitch. Comparison of bolus and infusion methods for receptor quantitation: application to [18F]cyclofoxy and positron emission tomography. *Journal of cerebral blood flow and metabolism : official journal*

- of the International Society of Cerebral Blood Flow and Metabolism*, 13(1): 24–42, 1993. ISSN 0271-678X. doi: 10.1038/jcbfm.1993.6.
- David M Cash, Ninon Burgos, John Dickson, Daniel Beasley, Pawel Markiewicz, Christopher A Lane, Thomas Parker, Anna Barnes, David L Thomas, M Jorge Cardoso, Ian B Malone, Thomas Veale, David Wallon, Jana Klimova, Kjell Erlandsson, Andrew Wong, Marcus Richards, Nick C Fox, Sébastien Ourselin, and Jonathan M. Schott. A comparison of techniques for quantifying amyloid burden on a combined PET/MR scanner. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 13: 1100–1101, 2017.
- Ciprian Catana. Motion correction options in PET/MRI. *Seminars in Nuclear Medicine*, 45(3):212–223, may 2015. ISSN 15584623. doi: 10.1053/j.semnuclmed.2015.01.001.
- Michael Chappell, Bradley MacIntosh, and Thomas Okell. *Introduction to Perfusion Quantification using Arterial Spin Labeling*. Oxford Oxford University Press 2018, 2018.
- Y. J. Chen, B. L. Rosario, W. Mowrey, C. M. Laymon, X. Lu, O. L. Lopez, W. E. Klunk, B. J. Lopresti, C. a. Mathis, and J. C. Price. PiB Relative Delivery (R1) as a Proxy of Relative Cerebral Blood Flow: Quantitative Evaluation Using Single Session 15O-Water and 11C-PiB PET. *Journal of Nuclear Medicine*, pages 1199–1205, 2015. ISSN 0161-5505. doi: 10.2967/jnumed.114.152405. URL <http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.114.152405>.
- Patricia Clement, Henk Jan Mutsaerts, Lena VáclavÁř, Eidrees Ghariq, Francesca B. Pizzini, Marion Smits, Marjan Acou, Jorge Jovicich, Ritva Vanninen, Mervi Kononen, Roland Wiest, Egill Rostrup, António J. Bastos-Leite, Elna Marie Larsson, and Eric Achten. Variability of physiological brain perfusion in healthy subjects – A systematic review of modifiers.

- Considerations for multi-center ASL studies. *Journal of Cerebral Blood Flow and Metabolism*, 2017. ISSN 15597016. doi: 10.1177/0271678X17702156.
- T F Cootes, C J Taylor, D H Cooper, and J Graham. Active shape models: their training and application. *Comp. Vis. and Image Understanding*, 61(1): 38–59, 1995.
- Christian Crone. The Permeability of Capillaries in Various Organs as Determined by Use of the Indicator Diffusion Method. *Acta Physiologica Scandinavica*, 58(4):292–305, 1963. ISSN 1365201X. doi: 10.1111/j.1748-1716.1963.tb02652.x.
- Zsolt Cselényi and Lars Farde. Quantification of blood flow-dependent component in estimates of beta-amyloid load obtained using quasi-steady-state standardized uptake value ratio. *Journal of Cerebral Blood Flow & Metabolism*, (March):1–9, 2015. ISSN 0271-678X. doi: 10.1038/jcbfm.2015.66. URL <http://www.nature.com/doifinder/10.1038/jcbfm.2015.66>.
- Gaspar Delso, Sebastian Fürst, Björn Jakoby, Ralf Ladebeck, Carl Ganter, Stephan G. Nekolla, Markus Schwaiger, and Sibylle I. Ziegler. Performance measurements of the siemens mMR integrated whole-body PET/MR scanner. *Journal of Nuclear Medicine*, 52(12):1914–1922, dec 2011. ISSN 01615505. doi: 10.2967/jnumed.111.092726.
- Kjell Erlandsson, Irène Buvat, P Hendrik Pretorius, Benjamin a Thomas, and Brian F Hutton. A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Physics in medicine and biology*, 57(21):R119–59, nov 2012. ISSN 1361-6560. doi: 10.1088/0031-9155/57/21/R119. URL <http://www.ncbi.nlm.nih.gov/pubmed/23073343>.
- A. P. Fan, H. Jahanian, S. J. Holdsworth, and G. Zaharchuk. Comparison of cerebral blood flow measurement with [15O]-water positron emission tomography and arterial spin labeling magnetic resonance imaging: A systematic

- review. *Journal of Cerebral Blood Flow & Metabolism*, 2016. ISSN 0271-678X. doi: 10.1177/0271678X16636393. URL <http://jcb.sagepub.com/lookup/doi/10.1177/0271678X16636393>.
- Audrey P. Fan, Jia Guo, Mohammad M. Khalighi, Praveen K. Gulaka, Bin Shen, Jun Hyung Park, Harsh Gandhi, Dawn Holley, Omar Rutledge, Prachi Singh, Tom Haywood, Gary K. Steinberg, Frederick T. Chin, and Greg Zaharchuk. Long-Delay Arterial Spin Labeling Provides More Accurate Cerebral Blood Flow Measurements in Moyamoya Patients: A Simultaneous Positron Emission Tomography/MRI Study. *Stroke*, 48(9):2441–2449, 2017. ISSN 15244628. doi: 10.1161/STROKEAHA.117.017773.
- Lucas Fidon, Wenqi Li, Luis C Garcia-Peraza-Herrera, Jinendra Ekanayake, Neil Kitchen, Sebastien Ourselin, and Tom Vercauteren. Scalable multi-modal convolutional networks for brain tumour segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 285–293. Springer, 2017.
- Eli Gibson, Wenqi Li, Carole Sudre, Lucas Fidon, Dzhoshkun I. Shakir, Guotai Wang, Zach Eaton-Rosen, Robert Gray, Tom Doel, Yipeng Hu, Tom Whyn-tie, Parashkev Nachev, Marc Modat, Dean C. Barratt, Sébastien Ourselin, M. Jorge Cardoso, and Tom Vercauteren. NiftyNet: a deep-learning platform for medical imaging. *Computer Methods and Programs in Biomedicine*, 158:113–122, 2018. ISSN 18727565. doi: 10.1016/j.cmpb.2018.01.025. URL <https://doi.org/10.1016/j.cmpb.2018.01.025>.
- R Goetti, G Warnock, F P Kuhn, R Guggenberger, R O Gorman, A Buck, N Khan, and I Scheer. Quantitative Cerebral Perfusion Imaging in Children and Young Adults with Moyamoya Disease : Comparison of Arterial Spin-Labeling MRI and H₂[15O]-PET. *Ajnr*, 35(May):1022–1028, 2014. ISSN 1936959X. doi: 10.3174/ajnr.A3799.
- Xavier Golay. Arterial spin labeling-MRI: acquisition and analysis tech-

- niques. In Peter B Barker, Xavier Golay, and Gregory Zaharchuk, editors, *Clinical Perfusion MRI Techniques and Applications*, chapter 3, pages 38–57. Cambridge University Press, 2013. ISBN 0849300649. doi: <http://dx.doi.org/10.1017/CBO9781139004053.005>.
- Sandeep SV Golla, Sander CJ Verfaillie, Ronald Boellaard, Sofie M Adriaanse, Marissa D Zwan, Robert C Schuit, Tessa Timmers, Colin Groot, Patrick Schober, Philip Scheltens, Wiesje M van der Flier, Albert D Windhorst, Bart NM van Berckel, and Adriaan A Lammertsma. Quantification of [18 F]florbetapir: A test–retest tracer kinetic modelling study. *Journal of Cerebral Blood Flow & Metabolism*, page 0271678X1878362, 2018. ISSN 0271-678X. doi: 10.1177/0271678X18783628. URL <http://journals.sagepub.com/doi/10.1177/0271678X18783628>.
- Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- Alex Graves, Abdel Rahman Mohamed, and Geoffrey Hinton. Speech recognition with deep recurrent neural networks. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 6645–6649, oct 2013. ISBN 9781479903566. doi: 10.1109/ICASSP.2013.6638947.
- R N Gunn, a a Lammertsma, S P Hume, and V J Cunningham. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *NeuroImage*, 6(4):279–287, 1997. ISSN 1053-8119. doi: 10.1006/nimg.1997.0303.
- Matthias Günther, Koichi Oshio, and David A. Feinberg. Single-shot 3D imaging techniques improve arterial spin labeling perfusion measurements. *Magnetic Resonance in Medicine*, 54(2):491–498, 2005. ISSN 07403194. doi: 10.1002/mrm.20580.
- Andreas Hahn, Martin Schain, Maria Erlandsson, Petter Sjölin, Gregory M.

- James, Olof T. Strandberg, Douglas Hägerström, Rupert Lanzenberger, Jonas Jögi, Tomas G. Olsson, Ruben Smith, and Oskar Hansson. Modeling strategies for quantification of in vivo 18F-AV-1451 binding in patients with tau pathology. *Journal of Nuclear Medicine*, 58(4):623–631, 2017. ISSN 2159662X. doi: 10.2967/jnumed.116.174508. URL <http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.116.174508>.
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- Fiona Heeman, Maqsood Yaquub, Isadora Lopes Alves, Kerstin Heurling, Johannes Berkhof, Juan Domingo Gispert, Santiago Bullich, Christopher Foley, and Adriaan A. Lammertsma. Optimized dual-time-window protocols for quantitative [18F]flutemetamol and [18F]florbetaben PET studies. *EJNMMI Research*, 9(1):1–14, 2019. ISSN 2191219X. doi: 10.1186/s13550-019-0499-4.
- D. F R Heijtel, H. J M M Mutsaerts, E. Bakker, P. Schober, M. F. Stevens, E. T. Petersen, B. N M van Berckel, C. B L M Majoie, J. Booij, M. J P van Osch, E. VanBavel, R. Boellaard, A. A. Lammertsma, and A. J. Nederveen. Accuracy and precision of pseudo-continuous arterial spin labeling perfusion during baseline and hypercapnia: A head-to-head comparison with 15O H2O positron emission tomography. *NeuroImage*, 92:182–192, 2014. ISSN 10959572. doi: 10.1016/j.neuroimage.2014.02.011. URL <http://dx.doi.org/10.1016/j.neuroimage.2014.02.011>.
- Karl Herrup. The case for rejecting the amyloid cascade hypothesis. *Nature Neuroscience*, 18(6):794–799, 2015. ISSN 1097-6256. doi: 10.1038/nn.4017. URL <http://www.nature.com/doi/10.1038/nn.4017>.
- Ing Tsung Hsiao, Chin Chang Huang, Chia Ju Hsieh, Wen Chun Hsu, Shiaw Pyng Wey, Tzu Chen Yen, Mei Ping Kung, and Kun Ju Lin. Corre-

- lation of early-phase 18F-florbetapir (AV-45/Amyvid) PET images to FDG images: Preliminary studies. *European Journal of Nuclear Medicine and Molecular Imaging*, 39(4):613–620, apr 2012. ISSN 1619-7089. doi: 10.1007/s00259-011-2051-2. URL <http://www.ncbi.nlm.nih.gov/pubmed/22270508>.
- Masanori Ichise, Hiroshi Toyama, Rb Innis, and Re Carson. Strategies to improve neuroreceptor parameter estimation by linear regression analysis. *Journal of Cerebral Blood . . .*, 22:1271–1281, 2002. doi: 10.1097/01.WCB.0000038000.34930.4E. URL <http://www.nature.com/jcbfm/journal/v22/n10/abs/9591323a.html>.
- Masanori Ichise, Jeih-San Liow, Jian-Qiang Lu, Akihiro Takano, Kendra Model, Hiroshi Toyama, Tetsuya Suhara, Kazutoshi Suzuki, Robert B Innis, and Richard E Carson. Linearized reference tissue parametric imaging methods: application to [11C]DASB positron emission tomography studies of the serotonin transporter in human brain. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, 23(9):1096–1112, 2003. ISSN 0271-678X. doi: 10.1097/01.WCB.0000085441.37552.CA.
- Robert B Innis, Vincent J Cunningham, Jacques Delforge, Masahiro Fujita, Albert Gjedde, Roger N Gunn, James Holden, Sylvain Houle, Sung-Cheng Huang, Masanori Ichise, Hidehiro Iida, Hiroshi Ito, Yuichi Kimura, Robert A Koeppe, Gitte M Knudsen, Juhani Knuuti, Adriaan A Lammertsma, Marc Laruelle, Jean Logan, Ralph Paul Maguire, Mark A Mintun, Evan D Morris, Ramin Parsey, Julie C Price, Mark Slifstein, Vesna Sossi, Tetsuya Suhara, John R Votaw, Dean F Wong, and Richard E Carson. Consensus Nomenclature for in vivo Imaging of Reversibly Binding Radioligands. *Journal of Cerebral Blood Flow & Metabolism*, 27(9):1533–1539, 2007. doi: 10.1038/sj.jcbfm.9600493. URL <http://journals.sagepub.com/doi/10.1038/sj.jcbfm.9600493>.

Clifford R. Jack, David S. Knopman, William J. Jagust, Leslie M. Shaw, Paul S. Aisen, Michael W. Weiner, Ronald C. Petersen, and John Q. Trojanowski. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology*, 9(1):119–128, 2010. ISSN 14744422. doi: 10.1016/S1474-4422(09)70299-6.

Clifford R. Jack, David S. Knopman, William J. Jagust, Ronald C. Petersen, Michael W. Weiner, Paul S. Aisen, Leslie M. Shaw, Prashanthi Vemuri, Heather J. Wiste, Stephen D. Weigand, Timothy G. Lesnick, Vernon S. Pankratz, Michael C. Donohue, and John Q. Trojanowski. Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *The Lancet Neurology*, 12(2):207–216, 2013. ISSN 14744422. doi: 10.1016/S1474-4422(12)70291-0.

Jieqing Jiao, Alexandre Bousse, Kris Thielemans, Ninon Burgos, Philip S.J. Weston, Jonathan M. Schott, David Atkinson, Simon R. Arridge, Brian F. Hutton, Pawel Markiewicz, and Sebastien Ourselin. Direct Parametric Reconstruction with Joint Motion Estimation/Correction for Dynamic Brain PET Data. *IEEE Transactions on Medical Imaging*, 36(1):203–213, jan 2017. ISSN 1558254X. doi: 10.1109/TMI.2016.2594150.

C L Joachim, J H Morris, and D J Selkoe. Diffuse senile plaques occur commonly in the cerebellum in Alzheimer's disease. *The American journal of pathology*, 135(2):309–319, 1989. ISSN 0002-9440.

Eric Karran, Marc Mercken, and Bart De Strooper. The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nature reviews. Drug discovery*, 10(9):698–712, 2011. ISSN 1474-1776. doi: 10.1038/nrd3505. URL <http://dx.doi.org/10.1038/nrd3505>.

Vincent Keereman, Yves Fierens, Tom Broux, Yves De Deene, Max Lonneux, and Stefaan Vandenberghe. MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences. *Journal of nuclear medicine : official*

- publication, *Society of Nuclear Medicine*, 51(5):812–8, may 2010. ISSN 1535-5667. doi: 10.2967/jnumed.109.065425. URL <http://www.ncbi.nlm.nih.gov/pubmed/20439508>.
- Emily Kilroy, Liana Apostolova, Collin Liu, Lirong Yan, John Ringman, and Danny J.J. Wang. Reliability of two-dimensional and three-dimensional pseudo-continuous arterial spin labeling perfusion MRI in elderly populations: Comparison with 15o-water positron emission tomography. *Journal of Magnetic Resonance Imaging*, 39(4):931–939, 2014. ISSN 15222586. doi: 10.1002/jmri.24246.
- William E. Klunk, Henry Engler, Agneta Nordberg, Yanming Wang, Gunnar Blomqvist, Daniel P. Holt, Mats Bergström, Irina Savitcheva, Guo Feng Huang, Sergio Estrada, Birgitta Ausén, Manik L. Debnath, Julien Barletta, Julie C. Price, Johan Sandell, Brian J. Lopresti, Anders Wall, Pernilla Koivisto, Gunnar Antoni, Chester a. Mathis, and Bengt Långström. Imaging Brain Amyloid in Alzheimer’s Disease with Pittsburgh Compound-B. *Annals of Neurology*, 55(3):306–319, 2004. ISSN 03645134. doi: 10.1002/ana.20009.
- R a Koeppe, K a Frey, G K Mulholland, M R Kilbourn, a Buck, K S Lee, and D E Kuhl. [11C]tropanyl benzilate-binding to muscarinic cholinergic receptors: methodology and kinetic modeling alternatives. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, 14(1):85–99, 1994. ISSN 0271-678X. doi: 10.1038/jcbfm.1994.13.
- Claes N. Ladefoged, Ian Law, Udunna Anazodo, Keith St. Lawrence, David Izquierdo-Garcia, Ciprian Catana, Ninon Burgos, M. Jorge Cardoso, Sebastien Ourselin, Brian Hutton, Inés Mérida, Nicolas Costes, Alexander Hammers, Didier Benoit, Søren Holm, Meher Juttukonda, Hongyu An, Jorge Cabello, Mathias Lukas, Stephan Nekolla, Sibylle Ziegler, Matthias Fenchel, Bjoern Jakoby, Michael E. Casey, Tammie Benzinger, Liselotte Højgaard, Adam E. Hansen, and Flemming L. Andersen. A multi-centre evaluation

- of eleven clinically feasible brain PET/MRI attenuation correction techniques using a large cohort of patients. *NeuroImage*, 147(June 2016):346–359, 2017. ISSN 10959572. doi: 10.1016/j.neuroimage.2016.12.010. URL <http://dx.doi.org/10.1016/j.neuroimage.2016.12.010>.
- Claes Nøhr Ladefoged, Lisbeth Marnier, Amalie Hindsholm, Ian Law, Liselotte Højgaard, and Flemming Littrup Andersen. Deep learning based attenuation correction of PET/MRI in pediatric brain tumor patients: Evaluation in a clinical setting. *Frontiers in Neuroscience*, 13(JAN), 2019. ISSN 1662453X. doi: 10.3389/fnins.2018.01005. URL [/pmc/articles/PMC6330282//pmc/articles/PMC6330282/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6330282/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6330282/).
- A. A. Lammertsma, C. J. Bench, S. P. Hume, S. Osman, K. Gunn, D. J. Brooks, and R. S.J. Frackowiak. Comparison of Methods for Analysis of Clinical [1C]Raclopride Studies. *Journal of Cerebral Blood Flow and Metabolism*, 16(1):42–52, 1996. ISSN 0271678X. doi: 10.1097/00004647-199601000-00005.
- AA Lammertsma and SP Hume. Simplified reference tissue model for PET receptor studies. *Neuroimage*, 158(4):153–158, 1996. URL <http://www.sciencedirect.com/science/article/pii/S105381199690066X>.
- Christopher A. Lane, Thomas D. Parker, Dave M. Cash, Kirsty Macpherson, Elizabeth Donnachie, Heidi Murray-Smith, Anna Barnes, Suzie Barker, Daniel G. Beasley, Jose Bras, David Brown, Ninon Burgos, Michelle Byford, M. Jorge Cardoso, Ana Carvalho, Jessica Collins, Enrico De Vita, John C. Dickson, Norah Epie, Miklos Espak, Susie M. D. Henley, Chandrashekar Hoskote, Michael Hutel, Jana Klimova, Ian B. Malone, Pawel Markiewicz, Andrew Melbourne, Marc Modat, Anette Schrag, Sachit Shah, Nikhil Sharma, Carole H. Sudre, David L. Thomas, Andrew Wong, Hui Zhang, John Hardy, Henrik Zetterberg, Sebastien Ourselin, Sebastian J. Crutch, Diana Kuh, Marcus Richards, Nick C. Fox, and Jonathan M. Schott. Study

- protocol: Insight 46 – a neuroscience sub-study of the MRC National Survey of Health and Development. *BMC Neurology*, 17(1):75, 2017. ISSN 1471-2377. doi: 10.1186/s12883-017-0846-x. URL <http://www.ncbi.nlm.nih.gov/pubmed/28420323><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5395844><http://bmcneuro.biomedcentral.com/articles/10.1186/s12883-017-0846-x>.
- Craig S. Levin, Sri Harsha Maramraju, Mohammad Mehdi Khalighi, Timothy W. Deller, Gaspar Delso, and Floris Jansen. Design Features and Mutual Compatibility Studies of the Time-of-Flight PET Capable GE SIGNA PET/MR System. *IEEE Transactions on Medical Imaging*, 35(8):1907–1914, aug 2016. ISSN 1558254X. doi: 10.1109/TMI.2016.2537811.
- Wenqi Li, Guotai Wang, Lucas Fidon, Sébastien Ourselin, M. Jorge Cardoso, and Tom Vercauteren. On the compactness, efficiency, and representation of 3d convolutional networks: Brain parcellation as a pretext task. In *International Conference on Information Processing in Medical Imaging (IPMI)*, 2017.
- Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra, Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A W M Van Der Laak, Bram Van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, 2017. doi: 10.1016/j.media.2017.07.005.
- Hong Liu-Seifert, Eric Siemers, Karen C. Holdridge, Scott W. Andersen, Ilya Lipkovich, Christopher Carlson, Gopalan Sethuraman, Sharon Hoog, Roza Hayduk, Rachelle Doody, and Paul Aisen. Delayed-start analysis: Mild Alzheimer’s disease patients in solanezumab trials, 3.5 years. *Alzheimer’s & Dementia: Translational Research & Clinical Interventions*, (July):1–11, 2015. ISSN 23528737. doi: 10.1016/j.trci.2015.06.006. URL <http://linkinghub.elsevier.com/retrieve/pii/S2352873715000177>.

- J Logan, J S Fowler, N D Volkow, a P Wolf, S L Dewey, D J Schlyer, R R MacGregor, R Hitzemann, B Bendriem, and S J Gatley. Graphical analysis of reversible radioligand binding from time-activity measurements applied to [N-11C-methyl]-(-)-cocaine PET studies in human subjects. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, 10(5):740–747, 1990. ISSN 0271-678X. doi: 10.1038/jcbfm.1990.127.
- J Logan, J S Fowler, N D Volkow, G J Wang, Y S Ding, and D L Alexoff. Distribution volume ratios without blood sampling from graphical analysis of PET data. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, 16(5): 834–840, 1996. ISSN 0271-678X. doi: 10.1097/00004647-199609000-00008.
- Brian J Lopresti, William E Klunk, Chester a Mathis, Jessica a Hoge, Scott K Ziolk, Xueling Lu, Carolyn C Meltzer, Kurt Schimmel, Nicholas D Tsopelas, Steven T DeKosky, and Julie C Price. Simplified quantification of Pittsburgh Compound B amyloid imaging PET studies: a comparative analysis. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, 46 (12):1959–1972, 2005.
- Pawel J. Markiewicz, Matthias J. Ehrhardt, Kjell Erlandsson, Philip J. Noonan, Anna Barnes, Jonathan M. Schott, David Atkinson, Simon R. Arridge, Brian F. Hutton, and Sebastien Ourselin. NiftyPET: a High-throughput Software Platform for High Quantitative Accuracy and Precision PET Imaging and Analysis. *Neuroinformatics*, 16:95–115, 2017. ISSN 15392791. doi: 10.1007/s12021-017-9352-y.
- Guy M McKhann, David S Knopman, Howard Chertkow, Bradley T Hyman, Clifford R Jack, Claudia H Kawas, William E Klunk, Walter J Koroshetz, Jennifer J Manly, Richard Mayeux, et al. The diagnosis of dementia due to alzheimerâ€™s disease: Recommendations from the national institute on aging-alzheimerâ€™s association workgroups on diagnostic guidelines for

- alzheimer's disease. *Alzheimer's & dementia: the journal of the Alzheimer's Association*, 7(3):263–269, 2011.
- Paul Meier and Kenneth L Zierler. On the Theory of the Indicator-Dilution Method for Measurement of Blood Flow and Volume. *Journal of Applied Physiology*, 6(12):731–744, 1954. ISSN 8750-7587. doi: 10.1152/jappl.1954.6.12.731. URL www.physiology.org/journal/jappl.
- Andrew Melbourne, Nicolas Toussaint, David Owen, Ivor Simpson, Thanasis Anthopoulos, Enrico De Vita, David Atkinson, and Sebastien Ourselin. NiftyFit: a Software Package for Multi-parametric Model-Fitting of 4D Magnetic Resonance Imaging Data. *Neuroinformatics*, 14(3):319–337, 2016. ISSN 15392791. doi: 10.1007/s12021-016-9297-6. URL <http://dx.doi.org/10.1007/s12021-016-9297-6>.
- Marc Modat, Gerard R. Ridgway, Zeike A. Taylor, Manja Lehmann, Josephine Barnes, David J. Hawkes, Nick C. Fox, and Sébastien Ourselin. Fast free-form deformation using graphics processing units. *Computer Methods and Programs in Biomedicine*, 98(3):278–284, 2010. ISSN 01692607. doi: 10.1016/j.cmpb.2009.09.002.
- Marc Modat, David M Cash, Pankaj Daga, Gavin P Winston, John S Duncan, and Sébastien Ourselin. Global image registration using a symmetric block-matching approach. *Journal of Medical Imaging*, 1(2):024003, 2014a. ISSN 2329-4302. doi: 10.1117/1.JMI.1.2.024003. URL <http://medicalimaging.spiedigitallibrary.org/article.aspx?doi=10.1117/1.JMI.1.2.024003>.
- Marc Modat, David M. Cash, Pankaj Daga, Gawin P. Winston, John S. Duncan, and Sébastien Ourselin. A symmetric block-matching framework for global registration. *SPIE Medical Imaging*, 9034:90341D, 2014b. ISSN 16057422. doi: 10.1117/12.2043652. URL <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2043652>.

- Elizabeth Morris, Anastasia Chalkidou, Alexander Hammers, Janet Peacock, Jennifer Summers, and Stephen Keevil. Diagnostic accuracy of ^{18}F amyloid PET tracers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *European Journal of Nuclear Medicine and Molecular Imaging*, pages 374–385, 2016. ISSN 16197089. doi: 10.1007/s00259-015-3228-x.
- William W Moses. Fundamental Limits of Spatial Resolution in PET. *Nuclear instruments & methods in physics research. Section A, Accelerators, spectrometers, detectors and associated equipment*, 648 Supple: S236–S240, 2011. ISSN 0168-9002. doi: 10.1016/j.nima.2010.11.092. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3144741&tool=pmcentrez&rendertype=abstract>.
- Tanya Nair, Doina Precup, Douglas L. Arnold, and Tal Arbel. Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11070 LNCS(DI):655–663, 2018. ISSN 16113349. doi: 10.1007/978-3-030-00928-1_74.
- J Nuyts, P Dupont, S Stroobants, R Benninck, L Mortelmans, and P Suetens. Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms. *IEEE transactions on medical imaging*, 18(5):393–403, may 1999. ISSN 0278-0062. doi: 10.1109/42.774167. URL <http://www.ncbi.nlm.nih.gov/pubmed/10416801>.
- John M Ollinger. *Physics in Medicine & Biology Model-based scatter correction for fully 3D PET*. 1996.
- Julie Ottoy, Jeroen Verhaeghe, Ellis Niemantsverdriet, Sebastiaan Engelborghs, Sigrid Stroobants, and Steven Staelens. A simulation study on the impact of the blood flow-dependent component in ^{18}F AV45 SUVR

- in Alzheimer's disease. *PLoS ONE*, 12(12):1–14, 2017a. ISSN 19326203. doi: 10.1371/journal.pone.0189155.
- Julie Ottoy, Jeroen Verhaeghe, Ellis Niemantsverdriet, Leonie Wyffels, Charisse Somers, Ellen De Roeck, Hanne Struyfs, Femke Soetewey, Steven Deleye, Tobi Van den Bossche, Sara Van Mossevelde, Sarah Ceyskens, Jan Versijpt, Sigrid Stroobants, Sebastiaan Engelborghs, and Steven Staelens. Validation of the Semiquantitative Static SUVR Method for ¹⁸F-AV45 PET by Pharmacokinetic Modeling with an Arterial Input Function. *Journal of Nuclear Medicine*, 58(9):1483–1489, 2017b. ISSN 0161-5505. doi: 10.2967/jnumed.116.184481. URL <http://jnm.snmjournals.org/lookup/doi/10.2967/jnumed.116.184481>.
- Nainesh Parikh, Kent P. Friedman, Shetal N. Shah, and Hersh Chandarana. Practical guide for implementing hybrid PET/MR clinical service: lessons learned from our experience. *Abdominal Imaging*, 40(6):1366–1373, aug 2015. ISSN 14320509. doi: 10.1007/s00261-015-0444-6.
- Laura M. Parkes and Paul S. Tofts. Improved accuracy of human cerebral blood perfusion measurements using arterial spin labeling: Accounting for capillary water permeability. *Magnetic Resonance in Medicine*, 48(1):27–41, 2002. ISSN 07403194. doi: 10.1002/mrm.10180.
- Laura M. Parkes, Waqar Rashid, Declan T. Chard, and Paul S. Tofts. Normal Cerebral Perfusion Measurements Using Arterial Spin Labeling: Reproducibility, Stability, and Age and Gender Effects. *Magnetic Resonance in Medicine*, 51(4):736–743, 2004. ISSN 07403194. doi: 10.1002/mrm.20023.
- Laura L.Boles Ponto, David J. Moser, Yusuf Menda, Emily L. Harlynn, Sean D. DeVries, Jacob J. Oleson, Vincent A. Magnotta, and Susan K. Schultz. Early Phase PIB-PET as a Surrogate for Global and Regional Cerebral Blood Flow Measures. *Journal of Neuroimaging*, 29(1):85–96, jan 2019. ISSN 15526569. doi: 10.1111/jon.12582.

- Julie C Price, William E Klunk, Brian J Lopresti, Xueling Lu, Jessica a Hoge, Scott K Ziolko, Daniel P Holt, Carolyn C Meltzer, Steven T DeKosky, and Chester a Mathis. Kinetic modeling of amyloid binding in humans using PET imaging and Pittsburgh Compound-B. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, 25(11):1528–1547, nov 2005. ISSN 0271-678X. doi: 10.1038/sj.jcbfm.9600146. URL <http://www.ncbi.nlm.nih.gov/pubmed/15944649>.
- Martin Prince, Anders Wimo, Maëlen Guerchet, Ali Gemma-Claire, Yu-Tzu Wu, and Matthew Prina. World Alzheimer Report 2015: The Global Impact of Dementia - An analysis of prevalence, incidence, cost and trends. *Alzheimer's Disease International*, page 84, 2015. ISSN 0963-7214. doi: 10.1111/j.0963-7214.2004.00293.x.
- EM Renkin. Transport of potassium-42 from blood to tissue in isolated mammalian skeletal muscles. *Am J Physiol*, 197(6):1205–1210, 1959. ISSN 0002-9513. URL <http://ajplegacy.physiology.org/content/197/6/1205.abstract>
<http://ajplegacy.physiology.org/content/197/6/1205.abstract?sid=586b9277-9912-41a8-93a4-2e126ea90447>.
- Anders Rodell, Joel Aanerud, Hans Braendgaard, and Albert Gjedde. Washout allometric reference method (WARM) for parametric analysis of [11C]PIB in human brains. *Frontiers in Aging Neuroscience*, 5(NOV), 2013. ISSN 16634365. doi: 10.3389/fnagi.2013.00045.
- Emilie Roncali and Simon R. Cherry. Application of silicon photomultipliers to positron emission tomography. *Annals of Biomedical Engineering*, 39(4): 1358–1377, apr 2011. ISSN 00906964. doi: 10.1007/s10439-011-0266-9.
- Cristian a Salinas, Graham E Searle, and Roger N Gunn. The simplified reference tissue model: model assumption violations and their impact on binding potential. *Journal of cerebral blood flow and metabolism : official*

- journal of the International Society of Cerebral Blood Flow and Metabolism*, 35(2):1–8, 2014. ISSN 1559-7016. doi: 10.1038/jcbfm.2014.202. URL <http://www.ncbi.nlm.nih.gov/pubmed/25425078>.
- Hasan Sari, Kjell Erlandsson, Lisbeth Marner, Ian Law, Henrik B.W. Larsson, Kris Thielemans, Sébastien Ourselin, Simon Arridge, David Atkinson, and Brian F. Hutton. Non-invasive kinetic modelling of PET tracers with radiometabolites using a constrained simultaneous estimation method: evaluation with 11 C-SB201745. *EJNMMI Research*, 8, 2018. ISSN 2191219X. doi: 10.1186/s13550-018-0412-6.
- Martin Schain, Francesca Zanderigo, J John Mann, and R Todd Ogden. Estimation of the binding potential BPND without a reference region or blood samples for brain PET studies. *NeuroImage*, 2016. ISSN 10538119. doi: 10.1016/j.neuroimage.2016.11.035. URL <http://linkinghub.elsevier.com/retrieve/pii/S1053811916306541>.
- Christopher G Schwarz, Matthew L Senjem, Jeffrey L Gunter, Nirubol Tosakulwong, Stephen D Weigand, Bradley J Kemp, Anthony J Spychalla, Prashanthi Vemuri, Ronald C Petersen, Val J Lowe, et al. Optimizing pib-pet suvr change-over-time measurement by a large-scale analysis of longitudinal reliability, plausibility, separability, and correlation with mmse. *Neuroimage*, 144:113–127, 2017.
- Catherine J Scott, Jieqing Jiao, Andrew Melbourne, Jonathan M Schott, Brian F Hutton, and Sébastien Ourselin. ASL-incorporated Pharmacokinetic Modelling of PET Data With Reduced Acquisition Time: Application to Amyloid Imaging. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, LNCS 9902:406–413, 2016. doi: 10.1007/978-3-319-46720-7. URL <http://link.springer.com/10.1007/978-3-319-46720-7>.
- Catherine J Scott, Jieqing Jiao, M Jorge Cardoso, Andrew Melbourne,

- Enrico De Vita, David L Thomas, Ninon Burgos, Pawel Markiewicz, Jonathan M. Schott, Brian F. Hutton, and Sébastien Ourselin. Short Acquisition Time PET Quantification Using MRI-Based Pharmacokinetic Parameter Synthesis. *In: Descoteaux M., Maier-Hein L., Franz A., Jannin P., Collins D., Duchesne S. (eds) MICCAI 2017. LNCS*, 10434:737–744, 2017. doi: 10.1007/978-3-319-66185-8. URL <http://link.springer.com/10.1007/978-3-319-66185-8>.
- Catherine J Scott, Jieqing Jiao, M Jorge Cardoso, Kerstin Kläser, Andrew Melbourne, Pawel Markiewicz, Jonathan M. Schott, Brian F. Hutton, and Sébastien Ourselin. Short Acquisition Time PET/MR Pharmacokinetic Modelling using CNNs. *In: Frangi, A., Schnabel, J., Davatzikos, Alberola-López, C., Fichtinger, G. (eds) MICCAI 2018. LNCS*, 11070 LNCS:48–56, sep 2018a. ISSN 16113349. doi: 10.1007/978-3-030-00928-1_6. URL https://link.springer.com/chapter/10.1007/978-3-030-00928-1_6.
- Catherine J Scott, Jieqing Jiao, Kerstin Kläser, M. Jorge Cardoso, Andrew Melbourne, Pawel J Markiewicz, Jonathan M Schott, Brian F Hutton, and Sébastien Ourselin. Reduced acquisition time PET/MR quantification: a comparison of tracer delivery estimation methods. *In Conference on PET and SPECT MR (PSMR) 2018*, 2018b.
- Catherine J Scott, Jieqing Jiao, Andrew Melbourne, Ninon Burgos, David M Cash, Enrico De Vita, Pawel J Markiewicz, Antoinette O'Connor, David L Thomas, Philip Sj Weston, Jonathan M Schott, Brian F Hutton, and Sébastien Ourselin. Reduced acquisition time PET pharmacokinetic modelling using simultaneous ASL&MRS: proof of concept. *Journal of Cerebral Blood Flow and Metabolism*, page 271678X18797343, sep 2018c. ISSN 15597016. doi: 10.1177/0271678X18797343. URL <http://www.ncbi.nlm.nih.gov/pubmed/30182792>.
- Catherine J Scott, Jieqing Jiao, Andrew Melbourne, Pawel J Markiewicz, Jonathan M Schott, Brian F Hutton, and Sébastien Ourselin. Reduced

- acquisition time amyloid PET quantification: sensitivity analysis. In *The XII International Symposium of Functional Neuroreceptor Mapping of the Living Brain (NRM) 2018*, 2018d.
- Catherine J Scott, Jieqing Jiao, Andrew Melbourne, Pawel J Markiewicz, Jonathan M Schott, Brian F Hutton, and Ourselin Sébastien. Reduced acquisition time PET quantification applied to [18f]-florbetapir. In *Human Amyloid Imaging Conference (HAI) 2018*, 2018e.
- Catherine J Scott, Jieqing Jiao, Andrew Melbourne, David L Thomas, Enrico De Vita, Pawel J Markiewicz, Jonathan M Schott, Brian F Hutton, and Sébastien Ourselin. Reduced acquisition time PET/MR quantification: optimisation for amyloid-pet data. In *Conference on PET and SPECT MR (PSMR) 2017*, 2018f.
- Mark Slifstein. Revisiting an old issue: the discrepancy between tissue ratio-derived binding parameters and kinetic modeling-derived parameters after a bolus of the serotonin transporter radioligand 123I-ADAM. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, 49(2): 176–8, 2008. ISSN 0161-5505. doi: 10.2967/jnumed.107.046631. URL <http://www.ncbi.nlm.nih.gov/pubmed/18245741>.
- J. Sojkova, J. Goh, M. Bilgel, B. Landman, X. Yang, Y. Zhou, Y. An, L. Beason-Held, M. Kraut, D. Wong, and S. Resnick. Voxel-wise relationships between distribution volume ratio and cerebral blood flow: implications for analysis of β -amyloid images. *Journal of Nuclear Medicine*, pages 1–26, 2015a. ISSN 0161-5505. doi: 10.2967/jnumed.114.151480. URL <http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.114.151480>.
- J Sojkova, J Goh, M Bilgel, B Landman, X Yang, Y Zhou, Y An, L L Beason-Held, M A Kraut, D F Wong, and S M Resnick. Voxelwise Relationships Between Distribution Volume Ratio and Cerebral Blood Flow: Implications

- for Analysis of beta-Amyloid Images. *J Nucl Med*, 56(7):1042–1047, 2015b. ISSN 0161-5505. doi: 10.2967/jnumed.114.151480.
- Tracy Ssali, Udunna C Anazodo, Jonathan D Thiessen, Frank S. Prato, and Keith St. Lawrence. A Non-invasive Method for Quantifying Cerebral Blood Flow by Hybrid PET/MR. *Journal of Nuclear Medicine*, page jnumed.117.203414, 2018. ISSN 0161-5505. doi: 10.2967/jnumed.117.203414. URL <http://jnm.snmjournals.org/lookup/doi/10.2967/jnumed.117.203414>.
- Rebecca M E Steketee, Esther E. Bron, Rozanna Meijboom, Gavin C. Houston, Stefan Klein, Henri J M M Mutsaerts, Carolina P. Mendez Orellana, Frank Jan de Jong, John C. van Swieten, Aad van der Lugt, and Marion Smits. Early-stage differentiation between presenile Alzheimer’s disease and frontotemporal dementia using arterial spin labeling MRI. *European Radiology*, 26(1):244–253, 2016. ISSN 14321084. doi: 10.1007/s00330-015-3789-x.
- Mohammed N. Tantawy, Carrie K. Jones, Ronald M. Baldwin, M. Sib Ansari, P. Jeffrey Conn, Robert M. Kessler, and Todd E. Peterson. [18F]Fallypride dopamine D2 receptor studies using delayed microPET scans and a modified Logan plot. *Nuclear Medicine and Biology*, 36(8):931–940, 2009. ISSN 09698051. doi: 10.1016/j.nucmedbio.2009.06.007. URL <http://dx.doi.org/10.1016/j.nucmedbio.2009.06.007>.
- Benjamin a Thomas, Kjell Erlandsson, Marc Modat, Lennart Thurfjell, Rik Vandenberghe, Sebastien Ourselin, and Brian F Hutton. The importance of appropriate partial volume correction for PET quantification in Alzheimer’s disease. *European journal of nuclear medicine and molecular imaging*, 38(6): 1104–19, jun 2011. ISSN 1619-7089. doi: 10.1007/s00259-011-1745-9. URL <http://www.ncbi.nlm.nih.gov/pubmed/21336694>.
- Bart N M van Berckel, Rik Ossenkuppele, Nelleke Tolboom, Maqsood Yaqub, Jessica C Foster-Dingley, Albert D Windhorst, Philip Scheltens, Adriaan a

- Lammertsma, and Ronald Boellaard. Longitudinal amyloid imaging using 11C-PiB: methodologic considerations. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, 54(9):1570–6, sep 2013. ISSN 1535-5667. doi: 10.2967/jnumed.112.113654. URL <http://www.ncbi.nlm.nih.gov/pubmed/23940304>.
- Larissa W van Golen, Joost P A Kuijer, Marc C Huisman, Richard G IJzerman, Frederik Barkhof, Michaela Diamant, and Adriaan a Lammertsma. Quantification of cerebral blood flow in healthy volunteers and type 1 diabetic patients: Comparison of MRI arterial spin labeling and [15 O]H 2 O positron emission tomography (PET). *Journal of magnetic resonance imaging : JMRI*, 40(6):1300–9, 2014. ISSN 10531807. doi: 10.1002/jmri.24484. URL <http://www.ncbi.nlm.nih.gov/pubmed/24214919>.
- M. J P Van Osch, Wouter M. Teeuwisse, M. A A Van Walderveen, Jeroen Hendrikse, Dennis A. Kies, and Mark A. Van Buchem. Can arterial spin labeling detect white matter perfusion signal? *Magnetic Resonance in Medicine*, 62(1):165–173, 2009. ISSN 07403194. doi: 10.1002/mrm.22002.
- Matthias J.P. van Osch, Wouter M. Teeuwisse, Zhensen Chen, Yuriko Suzuki, Michael Helle, and Sophie Schmid. Advances in arterial spin labelling MRI methods for measuring perfusion and collateral flow. *Journal of Cerebral Blood Flow and Metabolism*, 2017. ISSN 15597016. doi: 10.1177/0271678X17713434.
- Stefaan Vandenberghe and Paul K. Marsden. PET-MRI: A review of challenges and solutions in the development of integrated multimodality imaging, feb 2015. ISSN 13616560.
- Kathleen Vunckx, Ameya Atre, Kristof Baete, Anthonin Reilhac, Christophe M. Deroose, Koen Van Laere, and Johan Nuyts. Evaluation of three MRI-based anatomical priors for quantitative PET brain imaging.

- IEEE Transactions on Medical Imaging*, 31(3):599–612, mar 2012. ISSN 02780062. doi: 10.1109/TMI.2011.2173766.
- Esther Ah Warnert, Ashley D Harris, Kevin Murphy, Neeraj Saxena, Neeta Tailor, Nigel S Jenkins, Judith E Hall, and Richard G Wise. In vivo assessment of human brainstem cerebrovascular function: A multi-inversion time pulsed arterial spin labelling study. *Journal of Cerebral Blood Flow and Metabolism*, 34(6):956–963, 2014. ISSN 15597016. doi: 10.1038/jcbfm.2014.39. URL www.jcbfm.com.
- D. F. Wong, P. B. Rosenberg, Y. Zhou, A. Kumar, V. Raymont, H. T. Ravert, R. F. Dannals, A. Nandi, J. R. Brasic, W. Ye, J. Hilton, C. Lyketsos, H. F. Kung, a. D. Joshi, D. M. Skovronsky, and M. J. Pontecorvo. In Vivo Imaging of Amyloid Deposition in Alzheimer Disease Using the Radioligand 18F-AV-45 (Flobetapir F 18). *Journal of Nuclear Medicine*, 51(6):913–920, jun 2010. ISSN 1535-5667. doi: 10.2967/jnumed.109.069088. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3101877&tool=pmcentrez&rendertype=abstract>.
- K P Wong, D Feng, S R Meikle, and M J Fulham. Simultaneous estimation of physiological parameters and the input function—in vivo PET data. *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, 5(1):67–76, 2001. ISSN 1089-7771.
- Yanjun Wu and Richard E Carson. Noise reduction in the simplified reference tissue model for neuroreceptor functional imaging. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, 22(12):1440–1452, 2002. doi: 10.1097/01.WCB.0000033967.83623.34.
- Maqsood Yaqub, Nelleke Tolboom, Ronald Boellaard, Bart N M van Berckel, Erica W van Tilburg, Gert Luurtsema, Philip Scheltens, and Adriaan a Lam-

- mertsma. Simplified parametric methods for [11C]PIB studies. *NeuroImage*, 42(1):76–86, aug 2008. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2008.04.251. URL <http://www.ncbi.nlm.nih.gov/pubmed/18541442>.
- Frank Q. Ye, Karen Faith Berman, Timothy Ellmore, Giuseppe Esposito, John D. Van Horn, Yihong Yang, Jeff Duyn, Anne M. Smith, Joseph A. Frank, Daniel R. Weinberger, and Alan C. McLaughlin. H215O PET validation of steady-state arterial spin tagging cerebral blood flow measurements in humans. *Magnetic Resonance in Medicine*, 44(3):450–456, 2000. ISSN 07403194. doi: 10.1002/1522-2594(200009)44:3<450::AID-MRM16>3.0.CO;2-0.
- Fisher Yu and Vladlen Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*, 2015.
- H. Zaidi, N. Ojha, M. Morich, J. Griesmer, Z. Hu, P. Maniawski, O. Ratib, D. Izquierdo-Garcia, Z. A. Fayad, and L. Shao. Design and performance evaluation of a whole-body Ingenuity TF PET-MRI system. *Physics in Medicine and Biology*, 56(10):3091–3106, may 2011. ISSN 00319155. doi: 10.1088/0031-9155/56/10/013.
- Habib Zaidi and Kenneth F. Koral. Scatter modelling and compensation in emission tomography. *European Journal of Nuclear Medicine and Molecular Imaging*, 31(5):761–782, 2004. ISSN 16197070. doi: 10.1007/s00259-004-1495-z.
- Ke Zhang, Hans Herzog, Jörg Mauler, Christian Filss, Thomas W. Okell, Elena Rota Kops, Lutz Tellmann, Thomas Fischer, Burkhard Brocke, Walter Sturm, Heinz H. Coenen, and N. Jon Shah. Comparison of cerebral blood flow acquired by simultaneous [15O]water positron emission tomography and arterial spin labeling magnetic resonance imaging. *Journal of Cerebral Blood Flow and Metabolism*, 34(8):1373–1380, 2014. ISSN 15597016. doi: 10.1038/jcbfm.2014.92.

Yun Zhou, Susan M Resnick, Weiguo Ye, Hong Fan, Daniel P Holt, William E Klunk, Chester a Mathis, Robert Dannals, and Dean F Wong. Using a reference tissue model with spatial constraint to quantify [11C]Pittsburgh compound B PET for early diagnosis of Alzheimer's disease. *NeuroImage*, 36(2): 298-312, jun 2007. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2007.03.004. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2001263&tool=pmcentrez&rendertype=abstract>.