

Methods for Assisting the Automation of Dynamic Susceptibility Contrast Magnetic Resonance Imaging Analysis



Rashed Sobhan

Thesis submitted for partial fulfilment of the degree of Doctorate in Philosophy

University of East Anglia.

Norwich Medical School

Submission date: 30 December 2020

Candidate registration number: 100164475

© This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived therefrom must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

Abstract

Purpose

Dynamic susceptibility-contrast magnetic resonance imaging (DSC-MRI) is widely used for cerebral perfusion measurement, but dependence on operator input leads to a time-consuming, subjective, and poorly-reproducible analysis. Although automation can overcome these limitations, investigations are required to further simplify and accelerate the analysis. This research focuses on automating arterial voxel (AV) and brain tissue segmentation, and model-dependent deconvolution steps of DSC-MRI analysis.

Methods

Several features were extracted from DSC-MRI data; their AV- and tissue voxel-discriminatory powers were evaluated by the area-under-the-receiver-operating-characteristic-curve (AUC_{ROC}). Thresholds for discarding non-arterial voxels were identified using ROC cut-offs.

The applicability of DSC-MRI time-series data for brain segmentation was explored. Two segmentation approaches that clustered the dimensionality-reduced raw data were compared with two raw-data-based approaches, and an approach using principal component analysis (PCA) for dimension-reduction. Computation time and Dice coefficients (DCs) were compared.

For model-dependent deconvolution, four parametric transit time distribution (TTD) models were compared in terms of goodness- and stability-of-fit, consistency of perfusion estimates, and computation time.

Results

Four criteria were effective in distinguishing AVs, forming the basis of a framework that can determine optimal thresholds for effective criteria to discard tissue voxels with high sensitivity and specificity.

Compared to raw-data-based approaches, one of the proposed segmentation approaches identified GM with higher (>0.7 , $p<0.005$), and WM with similar DC. The approach outperformed the PCA-based approach for all tissue regions ($p<0.005$), and clustered similar regions faster than other approaches ($p<0.005$).

For model-dependent deconvolution, all TTD models gave similar perfusion estimates and goodness-of-fit. The gamma distribution was most suitable for perfusion analysis, showing significantly higher fit stability and lower computation time.

Conclusion

The proposed methods were able to simplify and accelerate automatic DSC-MRI analysis while maintaining performance. They will particularly help clinicians in rapid diagnosis and characterisation of tumour or stroke lesions, and subsequent treatment planning and monitoring.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Table of Contents

ABSTRACT	2
Purpose	2
Methods	2
Results	2
Conclusion	3
TABLE OF CONTENTS	4
LIST OF TABLES	11
LIST OF FIGURES	12
ABBREVIATIONS	18
COPYRIGHT LICENSES	22
ACKNOWLEDGEMENTS	23
DECLARATION	25
CHAPTER 1: INTRODUCTION	27
1.1. Background	27
1.2. Aims	31
1.3. Thesis outline	32
1.4. Data and computational apparatus	33

CHAPTER 2: BACKGROUND	34
2.1. Perfusion and Perfusion MRI: Brief Introduction	34
2.2. Perfusion MRI: Clinical Applications	34
2.2.1. Brain tumours	34
2.2.2. Glioma grading	35
2.2.3. Tumour histology differentiation	36
2.2.4. Treatment planning	36
2.2.5. Treatment monitoring and survival prediction	36
2.3. Perfusion MRI: Parameters of interest	37
2.4. Perfusion MRI: Types	40
2.4.1. Dynamic susceptibility contrast MRI (DSC-MRI)	40
2.4.2. Dynamic contrast-enhanced MRI (DCE-MRI)	41
2.4.3. Arterial spin labelling (ASL)	42
2.5. DSC MRI: Data acquisition protocol	42
2.6. DSC-MRI: Central convolution equation	44
2.6.1. Residue function and transit time distribution	44
2.6.2. Derivation of the generic equation for perfusion imaging	45
2.7. DSC-MRI: Analysis	48
2.7.1. AIF detection	48
2.7.2. Brain segmentation using DSC-MRI time-series	49
2.7.3. Impulse response function measurement	50
2.7.3.1. Model-independent deconvolution	50
2.7.3.2. Model-dependent deconvolution	53
2.7.4. Perfusion parameter measurement	54
2.7.4.1. CBF	54
2.7.4.2. MTT	55
2.7.4.3. CBV	56
CHAPTER 3: A COMPARISON OF CRITERIA FOR AUTOMATIC ARTERIAL INPUT FUNCTION DETECTION IN DSC-MRI	58
3.1. Introduction	58

3.2. Literature review	60
3.2.1. AIF detection with <i>k</i> -means clustering	60
3.2.2. AIF detection with modified <i>k</i> -means clustering	61
3.2.3. AIF detection with agglomerative hierarchical clustering	62
3.2.4. Comparative analysis of AH, <i>k</i> -means, and <i>c</i> -means clustering	63
3.2.5. Comparative analysis of Ncut, <i>k</i> -means and <i>c</i> -means clustering	64
3.2.6. Summary of threshold-based AIF detection studies	65
3.3. Aims	65
3.4. Materials and Methods	66
3.4.1. DSC-MRI data	66
3.4.2. Data pre-processing	68
3.4.3. Feature extraction	70
3.4.4. Identification of nominally-true AVs	72
3.4.4.1. <i>Thresholding</i>	73
3.4.4.2. <i>Feature space creation</i>	73
3.4.4.3. <i>Clustering</i>	74
3.4.4.4. <i>Centroid Selection</i>	75
3.4.4.5. <i>Manual AV selection</i>	75
3.4.5. Study on the effectiveness of the criteria	77
3.4.6. Investigation of Optimal Threshold for effective criteria	79
3.4.7. Statistical Analysis	80
3.5. Results	80
3.5.1. Nominal AVs	80
3.5.2. Effectiveness of criteria	83
3.5.2.1. <i>Investigation on a single subject</i>	83
3.5.2.2. <i>Investigation on all subjects</i>	85
3.5.3. Optimal threshold for individual criteria	87
3.6. Discussion	89
3.6.1. Nominal AVs	89
3.6.2. Comparison of criteria effectiveness	90
3.6.3. Optimal thresholds	91
3.6.4. Clinical applications	92
3.6.5. Limitations	93
3.6.6. Future scopes	95

3.7. Conclusion	95
-----------------	----

CHAPTER 4: APPLICATION OF FEATURE-BASED CLUSTERING ON DSC-MRI DATA FOR TISSUE SEGMENTATION 97

4.1. Introduction 97

4.1.1. Manual Segmentation	97
4.1.2. Atlas-based segmentation	97
4.1.3. Surface-based segmentation	98
4.1.4. Intensity-based segmentation	98
4.1.5. Hybrid segmentation	99

4.2. Literature review 100

4.2.1. Intensity-based segmentation with modified k -means clustering	100
4.2.2. Intensity-based segmentation of glioma volumes with c -means clustering	101
4.2.3. Segmentation of DSC-MRI using a single feature	102
4.2.4. Clustering a dimensionality-reduced version of the original data space	103
4.2.5. A popular dimension-reduction method and its prospects for brain segmentation	104
4.2.6. Segmentation using PCA-based dimension reduction and clustering	105

4.3. Aims 105

4.4 Materials and Methods 106

4.4.1. Raw data space (R) creation	106
4.4.2. Signal time series from different brain regions	108
4.4.3. Feature extraction	109
4.4.4. Feature space creation	113
4.4.5. Effective feature space (Optimisation of M)	114
4.4.6. Clustering	115
4.4.7. Segmentation approaches	117
4.4.7.1. Feature-based k -medoids clustering ($kmed$ -feature):	117
4.4.7.2. K -medoids clustering on raw data ($kmed$ -raw):	118
4.4.7.3. K -medoids clustering on principal-component-created feature space ($kmed$ -PCA):	118
4.4.7.4. Feature based k -means clustering with POA measure (POA -feature):	119
4.4.7.5. K -means clustering on raw-data with POA measure (POA -raw):	120
4.4.8. Evaluation of segmentation approaches	121
4.4.9. Statistical analysis	122

4.5. Results	123
4.5.1. Effectiveness of features	123
4.5.1.1. <i>Distribution of features</i>	123
4.5.1.2. <i>Individual discriminatory power of features:</i>	124
4.5.2. <i>R</i> space segmentation	125
4.5.2.1. <i>GM</i> segmentation	125
4.5.2.2. <i>WM</i> segmentation	126
4.5.2.3. <i>AIF</i> segmentation	127
4.5.2.4. <i>Lesion</i> segmentation	127
4.5.3. Computation time	129
4.6. Discussion	131
4.6.1. Effectiveness study of the features	132
4.6.2. Comparison of segmentation approaches	132
4.6.3. Clinical applications	134
4.6.4. Limitations	135
4.6.5. Future scopes	136
4.7. Conclusion	136
CHAPTER 5: COMPARING DIFFERENT FORMS OF PHYSIOLOGICALLY PLAUSIBLE TRANSIT TIME DISTRIBUTIONS	138
5.1. Introduction	138
5.2. Literature review on different forms of TTD	140
5.2.1. Gaussian TTD	140
5.2.2. Corrected Normal TTD	140
5.2.3. Skewed-Gaussian TTD	141
5.2.4. Gamma TTD	141
5.2.5. Gamma-variate TTD	142
5.3. Aims	143
5.4. Materials and Methods	144
5.4.1. Data acquisition and analysis	144
5.4.1.1. <i>Data acquisition</i>	144
5.4.1.2. <i>Automatic AIF detection</i>	145
5.4.1.3. <i>Weibull TTD</i>	146

5.4.1.4. Conversion of trial CTC to STC	147
5.4.1.5. Initial parameter guesses	148
5.4.1.6. Curve fitting and parameter estimation	149
5.4.2. Evaluation of Analysis	149
5.4.2.1. Simulation of TTDs for different tissues	149
5.4.2.2. Goodness of fit	150
5.4.2.3. Parameter Estimates	151
5.4.2.4. Success rate	151
5.4.2.5. Computation time	151
5.4.3. Statistical Analysis	152
5.5. Results	153
5.5.1. Functional forms of investigated TTDs	153
5.5.2. Goodness of fit	155
5.5.3. Parameter Estimates	158
5.5.3.1. CBF	159
5.5.3.2. MTT	160
5.5.3.3. CBV	161
5.5.4. Success rate	162
5.5.5. Computation Time	163
5.6. Discussion	164
5.6.1. Comparison between TTD models	164
5.6.2. Clinical applications	166
5.6.3. Limitations	167
5.6.4. Future scopes	168
5.7. Conclusion	169
CHAPTER 6: CONCLUSIONS AND FUTURE DIRECTIONS	171
6.1. Research outcomes	171
6.2. Limitations	173
6.2.1. Global AIF	173
6.2.2. Partial volume effect	174
6.2.3. Size of Patient Cohort	175
6.3. Future directions	176

6.3.1. Relating semi-quantitative parameters to perfusion estimates _____	176
6.3.2. Developing composite parameters _____	177
6.3.3. Local AIF detection _____	177
6.3.4. Validation using digital reference object _____	178
6.3.5. Brain map creation _____	179
6.3.6. End-to-end automated approach _____	180
APPENDICES _____	183
Appendix 1 _____	183
Appendix 2 _____	184
REFERENCES _____	193

Word count: 49,034

Colour pages: 41, 68, 72, 76, 78, 82, 83, 84, 85, 88, 108, 113, 123, 125, 126, 127, 128, 129, 130, 145, 154, 158, 162.

List of Tables

Table 3.1: Number of total brain voxels, N_{BV} (number of voxels left after pre-processing stage, see Figure 3.4), candidate voxels, N_{CV} (number of voxels that were chosen by centroid selection step) and arterial voxels, N_{AV} (number of voxels that were finally selected by visual inspection) for each subject with area under the curve (AUC) and peak concentration (PeakConc) as the centroid selection criteria. In the parenthesis, output voxels of interim steps (i.e. N_{CV} and N_{AV}) are presented as percentages of the voxels input to that step (i.e. N_{BV} and N_{CV} , respectively)	81
Table 3.2: Comparison of the ten features of the resultant AIFs obtained with each of the ten centroid selection criteria.....	84
Table 3.3: AUC_{ROCS} of different criteria for the chosen subject.....	85
Table 3.4: Mean $AUC_{ROC}(SD)$ of different criteria for all nine subjects.....	86
Table 3.5: Mean (SD) of optimal threshold (represented as % above the overall criteria mean) together with the mean (SD) of achieved sensitivity and specificity in distinguishing and discarding non-arterial voxels.....	88
Table 4.1: Mean (SE) of AUC_{ROC} of different features for different regions.....	124
Table 5.1: Shape and scale parameters for gamma TTD with cerebral blood flow (CBF) values for simulating healthy and pathological signals.....	150
Table 5.2: Mean (SD) of root mean square error (RMSE) obtained with each transit time distribution (TTD) for grey matter (GM) and white matter (WM).....	157
Table 5.3: Mean (SD) of parameter estimates obtained with each transit time distribution (TTD), h	159
Table 5.4: Mean (SD) success rate (percentage of successful fits) obtained with each transit time distribution (TTD).....	163
Table 5.5: Mean (SD) of computation time (T_{comp}) obtained with each transit time distribution (TTD).....	164

List of Figures

- Figure 2.1: (a) Segmented DCE-MRI image of brain (b) Typical DCE-MRI signals from different brain regions (c) Segmented DSC-MRI image of brain (d) Typical DSC-MRI signals from different brain regions. Figure reproduced from the work of Artzi et al. (2015), with permission from Springer Nature (Artzi et al., 2015)..... **41**
- Figure 3.1: Similarity between a linear time-invariant (LTI) system and typical parenchymal tissue. Symbols: $H(t)$, the response function of an LTI system; AIF(t), arterial input function; CBF, cerebral blood flow; $R(t)$, residue function. For both cases, the output is a convolution of the input and the response of the system..... **58**
- Figure 3.2: (a) DSC-MRI at different time points during GBCA administration. (b) A typical normalised DSC-MRI signal intensity time course obtained by averaging signals from the ROI in the caudate nucleus (shown in the inset) and then dividing by baseline signal intensity. The three different periods: baseline, first passage and recirculation are marked. **68**
- Figure 3.3: A typical AIF in arbitrary units (A.U.) taken from one of this study's subjects with different features labelled. The AIF is taken from one of the arterial voxels detected through the steps described in Section 2.4.4. Abbreviations: AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; T_{end} , end timepoint of bolus first pass..... **72**
- Figure 3.4: Schematic diagram of semi-automatic nominally-true AV detection. Manual intervention is marked with blue coloured box. Abbreviations: AUC, area under the curve; AUC_{max} , maximum of area under the curve over all voxels; PeakConc, peak concentration; $\text{PeakConc}_{\text{max}}$, maximum of peak concentration over all voxels; N , number of voxels left after thresholding; CTC, concentration time curve; AIF, arterial input function; AV, arterial voxel..... **76**
- Figure 3.5: Schematic diagram for studying the individual effectiveness of each of the ten criteria. Abbreviations: CTC, concentration time curve; AIF, arterial input function; AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak

concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8); MWO, mean wash-out rate. Subscript ‘max’ and ‘min’ refers to the maximum and minimum value of the corresponding criteria for choosing the centroid..78

Figure 3.6: (a) Arterial input functions (AIFs) of finally selected arterial voxels (AVs) for one subject. Inset shows the spatial locations of the AVs for the subject with arrows (the red arrow indicates an ROI that contains four AVs); (b) Spatial locations of AVs (white boxes) at different segments of middle cerebral artery (MCA) with their intensity dynamics. The AVs located specifically at M2 and M3 segments are congruent with those of identified AVs. Figure (b) is reproduced from the work of Zaro-weber et al. (2012), with permission from Wolters Kluwer Health, Inc..... 82

Figure 3.7: Final AIFs achieved with each of the ten criteria for one of the nine subjects (different to the one used for generating Figure 3.6). Abbreviations: a.u., arbitrary units; AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8), MWO, mean wash-out rate..... 83

Figure 3.8: (a) Receiver operating characteristics (ROC) curve for AUC, PeakConc, MWI, M_{peak} , MWO for one subject. (b) ROC curve for FWHM, FM, TTP, BAT, MTD for one subject. Abbreviations: AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8); MWO, mean wash-out rate..... 85

Figure 3.9: Bar chart for mean AUC_{ROC} (SD) for different criteria. The error bars represent SDs. Abbreviations: AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8); MWO, mean wash-out rate.....86

Figure 3.10: Box-plot of optimal thresholds for different criteria shown as percentage above the corresponding criterion mean over all voxels. Median of the optimal thresholds are represented by red lines; interquartile range by blue boxes; the highest and lowest 25% of the threshold values by upper and lower whiskers. Abbreviations: AUC, area

under the curve; PeakConc, peak concentration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8), MWO, mean wash-out rate..... **88**

Figure 4.1: Raw, non-normalised DSC-MRI time-series from different brain regions of one subject. The background and partial volume (PV)-affected signals are created as per the discussion of Section 4.4.1. It can be seen that signals from different regions differ in their features, like the initial drop, recovery, width, and baseline intensity. Abbreviations: AIF, arterial input function; GM, grey matter; WM, white matter; CSF, cerebrospinal fluid.....**108**

Figure 4.2: A sample DSC-MRI signal-time course showing the parameters necessary to calculate nine features. PC3 and PC4 are not included in the figure as they cannot be explicitly shown. Area over the curve (AOC) and L_2 norm were measured from the normalised signal, and L_2 norm measurement uses the signal drop at every time point and is therefore not included in the figure to maintain clarity. Abbreviations: S_{pre} , baseline signal; $SD_{S_{pre}}$, standard deviation of S_{pre} ; BAT, bolus arrival time; MSD, maximum signal drop; VAR, Variance; T_{end} , end time point of first pass; MTD, mean time duration; S_{end} , signal intensity at end of bolus first pass..... **113**

Figure 4.3: Boxplot showing the distribution of normalised feature values for different tissue regions. Colours to represent regions: red, arterial input function (AIF); black, grey matter (GM); blue, white matter (WM); yellow, partial volumes (PV); green, cerebrospinal fluid (CSF); magenta, lesion. For each region: the median of any feature value is represented by the corresponding coloured horizontal line, interquartile range by the corresponding coloured box, 99% confidence bounds by the corresponding coloured vertical lines, and outliers (of each region) by red crosses. Abbreviations: EDP, effective drop percentage; AOC, area over curve; Var, Variance; L_2 , L_2 norm; MTD, mean time duration; PC3, 3rd principal component; PSR, percentage signal recovery; PC4, 4th component; S_1/S_6 , intensity ratio of first image to sixth image..... **123**

Figure 4.4: Bar chart showing the individual performance parameters for GM segmentation. The error bars show the SDs above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-raw); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-raw). Statistical significance ($p < 0.005$,

due to Bonferroni correction) between two segmentation approaches is shown with a line joining the pairs of bar charts. Significance is only shown in the Dice coefficient (DC) to summarise the agreement between the ground truth and segmentation results..... **125**

Figure 4.5: Bar chart showing the individual performance parameters for WM segmentation. The error bars show the SDs above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-raw); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-raw). Statistical significance ($p < 0.005$, due to Bonferroni correction) between two segmentation approaches is shown with a line joining the pairs of bar charts. Significance is only shown in the Dice coefficient (DC) to summarise the agreement between the ground truth and segmentation results..... **126**

Figure 4.6: Bar chart showing the individual performance parameters for AIF segmentation. The error bars show the SDs above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-raw); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-raw). Statistical significance ($p < 0.005$, due to Bonferroni correction) between two segmentation approaches is shown with a line joining the pairs of bar charts. Significance is only shown in the Dice coefficient (DC) to summarise the agreement between the ground truth and segmentation results..... **127**

Figure 4.7: Bar chart showing the individual performance parameters for lesion segmentation. The error bars show the SDs above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-raw); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-raw). Statistical significance ($p < 0.005$, due to Bonferroni correction) between two segmentation approaches is shown with a line joining the pairs of bar charts. Significance is only shown in the Dice coefficient (DC) to summarise the agreement between the ground truth and segmentation results..... **128**

Figure 4.8: Bar chart showing the individual performance parameters for lesion segmentation with a feature-space of dimension 256×9 . The error bars show the SDs

above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-row); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-row). Statistical significance ($p < 0.005$, due to Bonferroni correction) between two segmentation approaches is shown with a line joining the pairs of bar charts. Significance is only shown in the Dice coefficient (DC) to summarise the agreement between the ground truth and segmentation results..... **129**

Figure 4.9: Bar chart showing the mean time elapsed during the clustering stage when five segmentation approaches are applied on (a) R space with no additional noise (b) R space with additional noise (SNR=20). The error bars show the SDs above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-row); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-row). Statistical significance ($p < 0.005$) between two segmentation approaches is shown with a line joining the pairs of bar charts..... **130**

Figure 5.1: (a) Schematic representation of vascular model; arterial input to each tissue is distributed across N tubes. Each fraction of input, h_i , passes with a distinct transit time τ_i , where $i = 1, 2, \dots, N$. (b) Histogram showing discretised approximation to the transit times by different h_i 's (Mouridsen et al., 2006b); when N is a very large number, this histogram converges into a continuous density function, named the transit time distribution (shown by solid curve). Figures reproduced from the work of Mouridsen et al. (2006b), with permission from Elsevier..... **139**

Figure 5.2: (a) Gradient recalled echo (GRE)-DSC-MRI (at the first time point), acquired at 1.5T. 3×3 -ROIs were placed in the caudate nucleus for GM and the frontal lobe for normal appearing WM (shown with yellow squares) (b) Corresponding average signal intensity curves (in arbitrary units, a.u.) for WM (in blue) and GM (in brown) for the selected ROIs..... **145**

Figure 5.3: Functional forms of (a) skewed-Gaussian, (b) gamma, (c) gamma-variate, and (d) Weibull TTDs in healthy (blue), ischaemic (black), and tumour (red) conditions. For each TTD, free parameter values and mean transit time (MTT) in each condition are also

given. Fitting the simulated signals gave unique gamma-variate TTDs, as they were modelled using decay and temporal parameters, unlike the shape and scale parameters of the other TTDs..... **154**

Figure 5.4: (a-d, left) Dot plots showing typical fits to a baseline-normalised GM signal time curve using skewed Gaussian, gamma, gamma-variate, and Weibull distributions, respectively. (e-h, right) Corresponding transit time distributions (TTDs) for each fit along with their free parameters and mean transit time (MTT) values. It can be seen that parametric signals obtained with all four TTDs fit well to the measured data, capturing the rapid signal drop and the recirculation stage accurately. All TTDs, except gamma-variate, have similar shapes and MTTs..... **156**

Figure 5.5: Boxplots of root-mean-square error (RMSE) values averaged over all samples and all subjects for (a) GM and (b) WM. The raw data are shown by black dots, median values by red lines, interquartile range by blue boxes, upper and lower adjacent values by the black whiskers. No significant difference was obtained between RMSEs quantified from four TTDs for both GM and WM..... **158**

Figure 5.6: Boxplot of CBF, MTT and CBV estimates for GM (a-c) and WM (d-f) for four models. The raw data are indicated by black dots, median values by red lines, interquartile range by blue boxes, and 99% confidence bounds by the black whiskers. No significant difference was obtained between the estimates quantified from four TTDs for both GM and WM. Abbreviations: CBF, cerebral blood flow; CBV, cerebral blood volume; MTT, mean transit time..... **162**

Abbreviations

AH: Agglomerative hierarchical

AIF: Arterial input function

AIC: Akaike information criterion

ANOVA: Analysis of variance

AOC: Area over the curve

ASL: Arterial spin labelling

AU: Arbitrary units

AUC: Area under the curve

AV: Arterial voxel

BAT: Bolus arrival time

BBB: Blood brain barrier

BIC: Bayesian information criterion

CBC: Characteristics-based clustering

CBF: Cerebral blood flow

CBV: Cerebral blood volume

CNS: Central nervous system

CSF: Cerebrospinal fluid

CT: Computed tomography

CTC: Concentration time curve

DC: Dice coefficient

DCE-MRI: Dynamic contrast-enhanced magnetic resonance imaging

DP: Discriminatory power

DPP: Diffusion/Perfusion Project

DRO: Digital reference object

DSC-MRI: Dynamic susceptibility-contrast magnetic resonance imaging

DWI: Diffusion weighted imaging

EDP: Effective drop percentage

EES: Extravascular extra-cellular space

EM-PCA: Expectation-maximisation based principal component analysis

EPI: Echo-planar imaging

FCM: Fuzzy *c*-means

FLAIR: Fluid attenuated inversion recovery

FM: First moment

FMRIB: Functional magnetic resonance imaging of the brain

FN: False negative

FP: False positive

FSL: FMRIB software library

FT: Fourier transform

FWHM: Full width at half maximum

GBCA: Gadolinium-based contrast agent

GBM: Glioblastoma multiform

GCV: Generalised cross validation

GM: Grey matter

GRE: Gradient recalled echo

ICA: Internal carotid artery

ITK: Insight Segmentation and Registration Toolkit

LCC: L-curve criterion

LTI: Linear time-invariant

MCA: Middle cerebral artery

MMC: Maximum concentration

MRI: Magnetic resonance imaging

MSD: Maximum signal drop

MTD: Mean time duration

MTT: Mean transit time

MWI: Mean wash-in rate

MWO: Mean wash-out rate

PAM: Partition around medoid

PeakConc: Peak concentration

PET: Positron emission tomography

POA: Probability of assignment

PCA: Principal component analysis

PPCA: Probabilistic principal component analysis

PS: Permeability-surface area product

PSR: Percentage signal recovery

PV: Partial volume

PVE: Partial volume effect PSR: Percentage signal recovery

rCBV: Relative cerebral blood volume

RECIST: Response Evaluation Criteria In Solid Tumours

RI: Roughness index

RMSE: Root-mean-square error

ROC: Receiver operating characteristic

ROI: Region of interest

SD: Standard deviation

SE: Spin echo

SNR: Signal to noise ratio

SPD: Signal peak decrease

SPM: Statistical parametric mapping

SPSS: Statistical package for the social sciences

SS: Post-bolus steady-state value

STC: Signal time curve/course

SVD: Singular value decomposition

SVM: Support vector machine

TE: Echo time

TIKH: Tikhonov regularisation method

TN: True negative

TP: True positive

TR: Repetition time

TTD: Transit time distribution

TTP: Time to peak

VOF: Venous output function

VOI: voxel of Interest

WHO: World Health Organisation

WM: White matter

Copyright Licenses

In this thesis, three figures have been used, with permission, from other published works. Figure 2.1 is copied from the work of Artzi et al. (2015) with permission license: 4962600138009. Figure 3.6 (b) is copied from work of Zaro-Weber et al. (2012) with permission license: 4963000305742. Figure 5.1 is modified from the work of Mouridsen et al. (2006) with permission license: 4962571281148.

Acknowledgements

I would like to first acknowledge *Dr. Glyn Johnson*, the person who gave me this exciting PhD opportunity. His kind advice, feedbacks and patronisation have helped me grow up as a resilient and independent researcher.

Next is my current supervisor, *Dr. Donnie Cameron*, my 'guiding light'. Thank you from the deep of my heart for your understanding of my circumstances, guiding me in my research, thesis, written works, and conference presentations, giving quality feedbacks well ahead of time, and most importantly, helping me revive my lost confidence as a researcher. Throughout my life, I will look up to you as a true researcher, guide, teacher, and motivator. I will really miss the time we spent together in our fortnightly meetings and will cherish any opportunity to work with you in future.

The next person would be *Dr. Gill Price*, my postgraduate training coordinator. During my days of turmoil, she gave me much-needed direction towards getting appropriate administrative support. Her continuous care and appreciation have truly helped me rebuild my confidence, break through all the negativity, and have a fresh start.

Dr. Chris Bishop would be the person who will get the biggest credit for helping me improve my written communication skills. Every fortnight we met, worked on the writing, and exchanged ideas; those sessions will always be the best learning experience for me. The learning journey with him seemed like a 'guided tour' in the ever-interesting world of English language.

I would like to convey my gratitude to *Dr. Penny Powell* and *Dr. Jennifer Whitty* for their administrative support during my troublesome circumstances. Also, I would thank my secondary supervisor *Dr. Pinelopi Gkogkou* for her support and advice during my probationary meeting and thereafter.

Now, to move out of UEA, my heartiest gratitude would go first to all my primary, secondary, undergraduate, and postgraduate teachers. Their selfless effort and consistent care have immensely contributed towards building my expertise and passion for science, teaching, and research.

Some personal, but not trivial, touch now; I want to thank USA Fried Chicken, Norwich and Norwich Pizza and Kebab House for their Kebabs, Burgers, and Fried Chicken—my rewards to myself after achieving any research milestone. Also, an enormous thanks to FIFA Mobile, which gave me a much-needed hiatus during my tiresome working days.

Let me now come to my family. *Amma* (Mom, *Mrs. Momtaz Begum*), **THIS IS FOR YOU!** In my exam days, you used to wait outside my school anxiously until I would come out and nod from far to assure: ‘it went well!’. No nodding today, **TAKE A BOW!** I can hardly give you back even one thousandth of the sacrifice, love, support, and prayer you bestowed on me. Please take this thesis as an attempt to make you content.

Now *Abbu* (Dad, *Mr. Abdus Sobhan*); your bedridden Cancer days are the reason why I chose this path of research. I have grown up seeing you contribute silently to our lives, but express your utmost happiness and pride after our success. I would always look up to you when bringing up my own sons. Your words of wisdom will always resonate in my ear and heart; I can still remember your words on the first day of my PhD journey: “Son, this is the day I have always cherished to see in my life!”

Now my Siblings; to begin with my eldest sister, *Ms. Sabrina Sobhan*, the best mentor for mathematics during my childhood. I would like to convey my heartiest appreciation for all your efforts that later kindled my passion for teaching and developed my verbal communication skills. Next comes the family mastermind, the solution-house: my sister *Mrs. Salma Sobhan*, and her husband—who is no less than a big brother and friend—*Mr. Chowdhury Sujat*. Take my heartiest gratitude, especially for supporting me during my early days in abroad and for the consistent care you guys provided to settle me and my family here in UK. Lastly, my elder brother, *Mr. Rahman Sobhan*, my childhood mentor for English Language. I am ever so grateful to you for those childhood days we played together, watched and shared views on movies, novels, and cricket, discussed grammar on the rooftop, and finally, for your assistance towards my English written communication.

Time to wrap it up, with the typical phrase ‘at last, but definitely not the least’. My utmost appreciation for you, my love, my best friend, my partner, my hard-shoulder, my dear wife: *Mahbuba Ahmed*. **Both this thesis and the associated Degree are as much mine as they are yours.** I know no better way than this to show my heart-felt gratitude for your support through all our ups and downs, your care for our kids during my busy times, your

sacrifice for my career progress, and most importantly, your unconditional love towards our family. Ever since we started our journey together on 2006, I am grateful to *Allah* that I got such a great partner in my life, and now grateful to have such lovely kids from you.

Cometh the two little angels of my life, *Affan Zareef Sobhan* and *Amnan Zareeb Sobhan*. You guys always inspire me to improve myself as a better researcher, a better father, and most importantly, a better human being, so that one day you can take pride in me. If I were not born as Rashed Sobhan, but as someone else, what blessings I would have missed! You guys' smile and love worth all those lovely distractions from work, calls of duty out of the blue, and sleepless nights. One day, when you guys will be old enough to read this, please know that Dad always loves you, takes pride in you, and deeply desires to spend more time with you. 'Loving you guys' would always be the second thing that Dad will ever do—the first being: breathing.

“This work is dedicated to my mom, dad, siblings, wife, and *Affan* and *Amnan*.”

Declaration

I declare that the research contained in this thesis is the original work of the author, unless otherwise formally indicated within the text. The thesis has not been submitted to University of East Anglia or any other universities for a degree and does not incorporate any material study submitted for a degree.

RASHED SOBHAN

Chapter 1: Introduction

1.1. Background

Since its development around the late 1980s as a T_2/T_2^* -weighted perfusion imaging modality, dynamic susceptibility-contrast magnetic resonance imaging (DSC-MRI) has continued to evolve and extend its application into the diagnosis, management, and treatment monitoring of different brain diseases (Jahng et al., 2014, Essig et al., 2012, Calamante, 2012, Cha et al., 2001, Vonken et al., 2000, Mouridsen et al., 2006a, Mouridsen et al., 2014). It is the current standard for quantifying cerebral perfusion, which is defined as the rate of blood delivery to brain tissue. Cerebral perfusion determines the transfer of oxygen and nutrients, as well as the removal of waste through blood in brain capillaries. It is therefore a key parameter for assessing tissue viability and vascularity, brain function and oxygen supply, as well as for identifying and characterising lesions.

Perfusion measurements can provide clinicians with crucial insight about the identification, characterisation, and progress of mass lesions (Welker et al., 2015). Abnormalities in perfusion parameters, such as cerebral blood flow (CBF), cerebral blood volume (CBV), mean transit time (MTT), can indicate pathologies like tumour, ischaemic stroke, moyamoya disease, intracranial neoplasm, haemorrhage, multiple sclerosis, and Alzheimer's disease. For example, malignant tumours show high CBV resulting from the increased metabolic demand following angiogenesis (Essig et al., 2012). In high grade glioma, the immature and leaky vessels lead to an increase in the permeability parameters that measure the leakage from the intravascular to the interstitial space (Sourbron and Buckley, 2013). Regions like the ischaemic penumbra can be characterised by decreased perfusion, but with normal diffusion (i.e. the diffusion-perfusion mismatch) (Calamante, 2012). For subjects with arterial abnormalities, such as stenosis, occlusion, or moyamoya disease, the rise in MTT (and corresponding fall in CBF) of the pathological regions assists diagnosis (Jahng et al., 2014).

Cerebral perfusion information can be captured with several imaging modalities, such as: positron emission tomography (PET), dynamic contrast-enhanced computed tomography

(DCE-CT), and perfusion MRI. PET involves injecting a radio-labelled bio-compound, such as fluorodeoxyglucose, which emits gamma radiation to indicate the accumulation of the bio-compound in a region as a function of time. Its limitations include: low spatial resolution, exposure to ionising radiation, long acquisition time leading to patient discomfort, sensitivity to natural metabolism and certain medications, necessity of co-registration to structural images, high cost of operation, and limited availability of PET centres with a cyclotron (Eugene and Abass, 2019, Grandin et al., 2005). A low-cost alternative is dynamic contrast-enhanced computed tomography (DCE-CT), which involves administering an iodinated contrast agent and dynamically imaging its first pass through a region (Miles, 2004, O'Connor et al., 2011, Grandin et al., 2005). However, this method also has challenges, including exposure to ionising radiation, necessity of patient preparation (e.g. beverage prohibition, bladder evacuation), long acquisition time leading to high chances of patient discomfort and motion artefact, as well as restricted anatomical coverage (Suetens, 2009). Lastly, perfusion MRI, a widely used cerebral perfusion imaging modality, quantifies perfusion and permeability by tracking the passage of an exogenous (i.e. externally injected) or endogenous (i.e. internally available) tracer with high temporal resolution MRI sequences.

Compared to DCE-CT, perfusion MRI exhibits higher soft-tissue contrast, allowing the acquisition of higher resolution brain images. It has no associated risk of ionising radiation and can be used to examine a larger population including pregnant women and younger individuals. It is also more suitable for patients with renal insufficiency, diabetes, and dehydration. Compared to PET, perfusion MRI is more available in clinics and it offers a less expensive, faster imaging modality with better contrast, but no ionising radiation exposure. For these benefits, perfusion MRI is chosen to extract cerebral perfusion information from the brain in this work.

There are three types of perfusion MRI: DSC-MRI, dynamic contrast-enhanced (DCE)-MRI, and arterial spin labelling (ASL). Conventionally, DSC-MRI is used for measuring cerebral perfusion and DCE-MRI, a T₁-weighted perfusion imaging modality, is applied to measure brain-tissue permeability (Sourbron and Buckley, 2013). ASL, on the other hand, extracts absolute CBF utilising labelled blood; the Look-Locker method with 2D or 3D excitation is used for single-time-point or dynamic acquisitions (Jahng et al., 2014).

Among these perfusion MRI modalities, the signal-to-noise ratio (SNR) is lower and acquisition procedures are more complex in ASL. The acquisition time is also longer, especially when larger brain coverage is required. Additionally, there are multiple implementations of ASL with no consensus regarding the ‘best’ choice, which is often decided on the basis of the scanner model, manufacturer, and software available rather than scientific considerations (Borogovac and Asllani, 2012). The need for sequence development and associated technical expertise has further limited its clinical applications. Among DSC- and DCE-MRI, the latter exhibits much lower SNR, lower contrast-to-noise in the first pass for perfusion estimation, and needs longer image acquisition time. It has to maintain a trade-off between spatial and temporal resolution, SNR, and anatomical coverage (O'Connor et al., 2011). Moreover, for DCE-MRI, image acquisition and post-processing with pharmacokinetic modelling are complex (Essig et al., 2013). On the other hand, DSC-MRI is the standard cerebral perfusion imaging modality with its widely available and easy-to-use processing tools (Essig et al., 2013). It allows acquisition and visualisation of images from the whole brain within 1–2 minutes. Therefore, DSC-MRI is currently more suitable than DCE-MRI and ASL for rapid, simple, but effective examination of subjects with brain diseases, such as tumour or stroke (Jahng et al., 2014). Additionally, the wide availability of DSC-MRI allows investigations—like the present work—to further confirm the validity of their inferences (Essig et al., 2013). Considering all these aspects, DSC-MRI was selected as the MRI modality of choice for the glioma patient cohort presented in this work.

In DSC-MRI, dynamic T_2/T_2^* -weighted images are acquired through gradient-recalled-echo or spin-echo echo-planar imaging sequences during the passage of an exogenous, but intravascular, Gadolinium-based contrast agent (GBCA). The paramagnetic GBCA causes each voxel to demonstrate a transient T_2/T_2^* signal, which is then analysed to extract perfusion parameters. The output signal is the convolution of two time-dependent functions: the GBCA concentration input from the artery, referred to as the arterial input function (AIF); and the response of the tissue to that input, referred to as the tissue impulse response, $I(t)$.

Despite its afore-mentioned advantages over DCE-MRI and ASL, DSC-MRI requires expert Radiologists’ input during analysis. For example, AIF detection requires a Radiologist to place a cursor on the candidate voxels and identify arterial voxels (AVs) using visual feedback, such as the shape characteristics of the signal time courses (STCs)

or concentration time courses (CTCs) (Patil et al., 2013, Yin et al., 2014, Mouridsen et al., 2006a). Another intermediate step that often adds a human component to any MRI analysis is the segmentation of brain regions. Human expertise is needed either to manually delineate the desired region, create a manually-labelled atlas to which images are registered, or combine different manual or automatic segmentation methods to segment the brain. Such manual interventions make the analysis more subjective, time-consuming, and less reproducible than automatic approaches, and are thus more prone to producing sub-optimal results. For these reasons, many research groups advocate increased automation of the analysis steps. The present work looks into the automation of three intermediate analysis steps, namely AIF detection, tissue segmentation, and perfusion quantification, as described below.

Conventionally, automatic AIF detection applies an algorithm to identify AVs based on their typical CTC features, such as: a large area under the curve, early bolus arrival, high peak concentration, low bolus width, small first moment, and a short time to peak. To discard the soft-tissue voxels, different thresholds are applied to the criteria characterising these features (Yin et al., 2014, Yin et al., 2015, Mouridsen et al., 2006a). To date, these criteria thresholds have been empirical and there has been no attempt to systemically determine their appropriate ranges. No study has compared the individual AV-discriminatory power of different criteria. Furthermore, many automatic AV detection studies used off-the-shelf clustering algorithms to group the raw CTCs with similar dynamics. Instead of clustering these high-dimensional data, clustering their dimensionality-reduced version can potentially simplify and accelerate AV-detection.

Automatic segmentation can be achieved by clustering similar brain-tissue intensity signatures. However, for DSC-MRI, the signal intensities vary as a function of time for a voxel. Hence, the label assigned to the voxel will also vary as a function of time. To overcome this issue, similar dynamics can be grouped together by clustering the DSC-MRI time-series data. However, the high dimensionality of the raw data increases the complexity and computation time of the overall segmentation. The segmentation can potentially be accelerated and simplified if the dimensionality-reduced data are clustered instead of the raw data.

After AIF determination and tissue class identification for a given region of interest (ROI), an automated analysis quantifies its perfusion parameters. First, a unique response

function, $I(t)$, is determined by removing the contribution of AIF from the tissue output signal with either a model-independent or model-dependent deconvolution method. Instead of estimating $I(t)$ at every time point, like the model-independent variant, the model-dependent methods assume realistic, yet flexible, analytical forms of $I(t)$ with only two or three free parameters. An automated workflow then fits several parametric signals—resulting from a range of initial guesses of free parameters—to the measured data. When the fit converges, the corresponding free parameters are used to generate perfusion estimates. Conventionally, $I(t)$ is derived from parametric models of the transit time distribution (TTD): the density function of the transit times necessary for the GBCA particles to traverse the capillary tubes distributed over a region of interest (ROI). Although many parametric forms of TTD have been proposed, further investigations are necessary to compare their computational benefits—such as rapidity, simplicity, or stability—towards an automated approach.

In this thesis, two main developments to the above-mentioned automations are investigated: methods to *assist* automatic AIF detection and brain segmentation; and an *evaluation* of the utility of several analytical forms of $I(t)$ for use in automated perfusion quantification. These aims are detailed in the following section.

1.2. Aims

The first aim of this research is to develop methods that assist *automatic AIF detection*. The *individual effectiveness* of different AV-detection criteria is evaluated. The optimal *criteria thresholds* for soft-tissue-voxel elimination are then systematically investigated.

The second aim of this research is to explore the applicability of DSC-MRI data for *automatically segmenting* brain regions. The utility of a dimensionality-reduced feature space in brain segmentation is compared to other raw-data-based and dimension-reduction-based segmentation approaches.

The third aim of the research is to *compare different parametric forms of $I(t)$* . Three previously published models of TTD are compared with a proposed model to ascertain whether any model gives at least one computational benefit.

1.3. Thesis outline

Chapter two presents the clinical utility, technical terminologies, basic theory, acquisition protocol, and analysis techniques of perfusion MRI, especially DSC-MRI. The knowledge gaps of different intermediate DSC-MRI analysis steps are also introduced.

Chapter three, the first Methods chapter, details the investigation performed to assist automatic AV detection. Several features are extracted from each CTC and the power of each feature in discriminating the AVs from the available brain voxels is evaluated through receiver operating characteristic curves. Then, a framework is established for the determination of criteria thresholds that can optimally discard tissue voxels.

Chapter four proposes two novel, feature-based segmentation approaches that cluster a dimensionality-reduced version of the original raw data space. First, the individual powers of several signal features to discriminate different tissue regions are evaluated. The most powerful features are then included in the dimensionality-reduced space. The segmentation performance and computation time of these feature-based approaches are compared to two approaches that cluster the raw data for segmentation, and also to an approach that implements principal component analysis for dimension reduction prior to clustering and segmentation.

Chapter five describes and compares three already published and one proposed parametric forms of TTD. For each TTD, non-linear regression is performed to estimate the free parameters of TTDs, which were then used to quantify the perfusion parameters. The TTDs are compared in terms of the goodness and stability of the fits between their resultant parametric signals and the measured data, the consistency of their perfusion estimates, and the overall computation time.

Chapter six concludes the thesis by summarising the major findings and the limitations of the present research before outlining future directions.

1.4. Data and computational apparatus

All the analyses of this thesis were performed on a dataset derived from DSC-MRI scans of 35 low-grade glioma patients (23 male, 12 female) in a previously-published, institution review board approved study (Law *et al.*, 2006). The DSC-MRI data were acquired at 1.5T (Siemens Vision/Symphony; Siemens Healthineers, Erlangen, Germany) with a gradient-recalled-echo echo-planar imaging sequence during the first pass of a standard dose (0.1 mmol/kg) bolus of gadopentetate dimeglumine (Magnevist, Berlex Laboratories, Wayne, NY). Imaging parameters were: TR/TE: 1,000/54 ms, field of view, 230 × 230 mm; section thickness, 5 mm; matrix, 128 × 128; in-plane voxel size, 1.8 × 1.8 mm; interslice gap, 0%–30%; flip angle, 30°; signal bandwidth, 1470 Hz/pixel. Contrast was injected at a rate of 5 ml/sec, followed by a 20 ml bolus of saline at 5 ml/sec. A total of 60 images were acquired at one second intervals, giving a total acquisition time of one minute. The injection coincided with the fifth image, so that the bolus would typically arrive at the fifteenth to twentieth image.

The image processing, and quantitative and statistical analyses were performed using MATLAB 2016-2019a (The MathWorks, Natick, MA, United States). For some analyses, such as generation of the receiver operating characteristic curves and execution of repeated measures analysis of variance (ANOVA), the Statistical Package for the Social Sciences (SPSS) software was used (IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.). All these analyses were executed using a personal laptop workstation, running the 64-bit Windows 10 operating system with a 2.50–2.70 GHz Intel® core™ i5-7200U central processing unit (Microsoft, Redmond, WA, United States).

Chapter 2: Background

2.1. Perfusion and Perfusion MRI: Brief Introduction

Cerebral perfusion describes the rate of blood flow to brain tissue (Calamante, 2012). Blood carries oxygen and nutrients and, hence, cerebral perfusion can characterise the delivery of oxygen and nutrients to living brain tissues (Jahng et al., 2014). Perfusion measurement through magnetic resonance imaging (MRI), first introduced in the mid-1980s, is a cerebral haemodynamic assessment technique that involves acquiring high temporal resolution MR images during the passage of a Gadolinium (Gd)-based contrast agent (GBCA). Perfusion MRI can provide crucial inferences to altered cerebral perfusion for several pathological conditions, such as tumour, acute ischaemic stroke, intracranial neoplasm, haemorrhage, multiple sclerosis, and Alzheimer's disease (Sourbron and Buckley, 2013, Mouridsen et al., 2014, Calamante, 2012).

In this chapter, the applications of perfusion MRI are briefly given to highlight its clinical utility. These will be followed by the description of relevant technical terminologies, different perfusion MRI modalities, standard image acquisition protocol, and analysis techniques.

2.2. Perfusion MRI: Clinical Applications

In this section, the clinical applications of perfusion MRI with regards to brain tumours will be briefly introduced and discussed.

2.2.1. Brain tumours

Cancer cells accumulate rapidly and cause a significant increase in the metabolic demand of affected brain regions (Guzman-de-Villoria et al., 2012). To satisfy this increased demand, new immature vessels are created. This causes the tumour region to exhibit abnormally high perfusion and permeability (i.e. contrast agent leakage from tissue). When tumours become malignant, the neo-vasculature becomes abnormal, demonstrating

high vascular density and disorganised vessel architecture. The ability of perfusion MRI to diagnose these mass lesions makes it highly useful in brain tumour-related applications.

By identifying abnormalities in tissue perfusion or permeability, perfusion MRI can differentiate malignant regions from healthy brain parenchyma (Nagesh et al., 2007, Jahng et al., 2014, Boxerman et al., 2006). For example, the significant rise of cerebral blood volume (CBV, defined in Section 2.3) facilitates the diagnosis of tumours and their differentiation from pseudo-tumoural inflammatory lesions (Essig et al., 2012). Moreover, GBCA takes a long time to pass through the tortuous vasculature of tumours. In such cases, perfusion imaging shows high mean transit time (MTT, discussed in Section 2.3) (Yin et al., 2014, Calamante, 2013). As the grey matter-white matter contrast for MTT is low, it can be used for visual identification of the pathological regions more effectively. For histopathological diagnosis, biopsy guided by perfusion imaging improves the discrimination of malignant tissue (Essig et al., 2012).

2.2.2. Glioma grading

The diagnostic utility of perfusion MRI is further complemented by its ability to characterise the mass lesions. Grading the glioma plays a crucial role in deciding whether the post-surgery treatment plan should include adjuvant chemo- or radiotherapy. Using perfusion-MRI, high-grade glioma—World Health Organisation (WHO) grades III and IV—can be differentiated from its low-grade variants (WHO grades I and II) with sensitivity and specificity ranging from 72.5–100% and 55–96.8%, respectively (Guzman-de-Villoria et al., 2012, Emblem et al., 2008). Relative CBV (rCBV, defined in Section 2.3), a widely used perfusion parameter, has strong correlation with angiographic estimates of vascularity (Sugahara et al., 1998), histological estimates of microvascular density (Cha et al., 2003), and choline—a marker of tumour aggressiveness (Guzman-de-Villoria et al., 2012, L Boxerman et al., 2006, Tzika et al., 2003).

Visual inspection of the perfusion images can give an initial idea about the integrity of blood brain barrier (BBB): a physical barrier that limits the leakage of GBCA from blood to the central nervous system (CNS). High-grade glioma demonstrates a large number of leaky vessels. T₁-weighted perfusion imaging can quantify the GBCA leakage through these vessels and thereby characterise high grade glioma regions.

2.2.3. Tumour histology differentiation

Perfusion MRI facilitates discrimination between the neoplastic aetiologies of masses and also between neo- and non-neo-plastic aetiologies of lesions (Welker et al., 2015). For example, the low rCBV of primary lymphoma can be used to distinguish it from glioblastoma multiforme (GBM) (Hartmann et al., 2003) and certain metastases from high-grade astrocytoma (Kremer et al., 2003, Leu et al., 2016). Peritumoural oedema, a characteristic feature of malignant glioma, can also be identified and distinguished from metastases through CBV. In cases of diagnostic uncertainty in differentiating primary tumours from metastatic lesions, perfusion MRI can improve confidence (Essig et al., 2012, Cha et al., 2001).

2.2.4. Treatment planning

Surgical resection is often the treatment of choice for high-grade glioma (Essig et al., 2012). Perfusion MRI helps surgeons to decide whether to attempt resection and whether to follow it by post-operative radiation therapy, chemotherapy, or both.

The macroscopic tumour regions visible in T₁-weighted perfusion images facilitate the preliminary identification of the gross tumour volume. T₂-weighted perfusion images indicate the clinical target volume through hyper-intense lesions. As a result, a safety margin for the possible microscopic spread is obtained. Additionally, perfusion MRI can guide stereotactic biopsies and other surgical interventions of highest-grade regions of glioma (Welker et al., 2015). It is also used to post-operatively assess the success of partial or total resection by identifying residual tumour (Nagesh et al., 2007, Essig et al., 2012). A set of unenhanced and contrast-enhanced CT and MRI imaging often assists the treatment planning of high-grade glioma patients through radiotherapy.

2.2.5. Treatment monitoring and survival prediction

Perfusion MRI not only assists treatment planning, but also facilitates the evaluation of drug response, and prediction of tumour progression and survival. The performance of anti-angiogenic drugs in decreasing the permeability and vascular density of affected regions can be evaluated, which assists in monitoring the response to drugs (Skinner et

al., 2016b, Zhu et al., 2005, Jahng et al., 2014, Bjornerud and Emblem, 2010, Vonken et al., 2000). The perfusion parameter MTT can provide important information about treatment response by capturing any decline in perfusion pressure (Ibaraki et al., 2007, Nagesh et al., 2007). Additionally, some studies have showed that early CBV changes during radiotherapy can predict response to treatment (Bjornerud and Emblem, 2010, Vonken et al., 2000).

The parameter rCBV assists prediction of the progression-free survival: an indicator of the success of treatment, measured as the time from treatment to death from any cause or to disease progression—based upon an onset of clinical symptoms or follow-up assessments using Response Evaluation Criteria In Solid Tumours (RECIST) criteria (Therasse et al., 2000). For both high- and low-grade glioma, patients with rCBV < 1.75 showed higher mean progression-free survival. For rCBV > 1.75, the progression-free survival is not significantly different for low- and high-grade glioma (Law et al., 2006). rCBV is also a better predictor of disease course than histological analysis (Law et al., 2003).

From the above discussion, the clinical utility of perfusion MRI for the diagnosis, grading, and treatment of brain tumours is evident. These applications require quantification of different perfusion parameters, the most important of which will be introduced in the following section.

2.3. Perfusion MRI: Parameters of interest

Cerebral blood flow (CBF) is defined as the rate at which blood is delivered to the brain tissue. It is measured in millilitres of blood per 100 grammes of brain tissue per minute (ml/100g/min) (Calamante, 2013, Yin et al., 2014). The white matter (WM) contains myelinated axons and transmits impulses between peripheral nervous system and the grey matter (GM), which contains the cell bodies, axon terminals, and dendrites and executes the more demanding task of processing information. Due to their different cellular constituents and metabolic demands, GM and WM exhibits marked contrast in their CBFs (Helenius et al., 2003). Typical values of CBF are 60 ml/100g/min for grey matter (GM) and 20 ml/100g/min for white matter (WM) (Guzman-de-Villoria et al., 2012). CBF evaluates the rate of exchange of oxygen and nutrients, as well as removal of waste from

brain tissue. Therefore, it can inform clinicians about tissue viability, vascularity, brain function, and oxygen supply, and can assist the identification and characterisation of lesions.

Cerebral blood volume (CBV) is defined as the fraction of the tissue volume occupied by blood, with units of millilitres of blood per 100 grammes of brain tissue (ml/100g) (Guzman-de-Villoria et al., 2012, Calamante, 2012). As mentioned above, a larger fraction of blood traverses the GM than the WM to meet the metabolic demand. Hence, GM exhibits higher CBV than WM, with typical values of 4 ml/100g for GM and 2 ml/100g for WM (Calamante et al., 2000, Calamante, 2012). Due to difficulties in measuring absolute CBV, relative CBV (rCBV) is often reported, which is the CBV relative to an internal control, such as contralateral normal WM or arterial concentration time curve (Jahng et al., 2014, Guzman-de-Villoria et al., 2012). As discussed in Section 2.2, CBV and rCBV can serve as important imaging bio-markers for tumour diagnosis, grading, low-grade to high-grade transformation, recurrent tumour discrimination from pseudo-progression, and overall treatment response assessment (Leu et al., 2016).

Mean transit time (MTT) is defined as the average time taken by blood to pass from the arterial inflow to the venous outflow, measured in seconds (Guzman-de-Villoria et al., 2012). Typical MTTs for healthy GM and WM are 4 s and 4.8 s respectively (Calamante et al., 2000, Calamante, 2012). Abnormalities in MTT can indicate pathology as it is sensitive to change in perfusion pressure. In addition to the applications of MTT described in Section 2.2, increases in MTT can suggest perfusion reserve impairment in ischemic stroke and chronic occlusive cerebrovascular disease (Ibaraki et al., 2007). The relationship between the CBV, CBF and MTT is given by the central volume theorem (Perl et al., 1975):

$$MTT = \frac{CBV}{CBF} . \quad (2.1)$$

Several semi-quantitative or heuristic parameters, often referred as ‘summary parameters’ (Jahng et al., 2014, Calamante, 2012), can be easily quantified from perfusion data without intensive DSC-MRI analysis. These summary parameters are effective in representing physiological mechanisms when measuring relative changes in perfusion dynamics for a subject or a group of patients (Gordon et al., 2014). Some of the important summary parameters are discussed below.

Time to peak (TTP) measures the time taken for the perfusion MRI signal time curve (STC) to reach its minimum intensity. When signal is converted to concentration estimates, TTP measures the time at which the concentration time curve (CTC) reaches its maximum. For stroke patients, it can successfully discriminate infarcted from non-infarcted voxels (Christensen et al., 2009).

Bolus arrival time (BAT) is defined as the time taken by the GBCA to arrive at the voxel of interest after its injection. It reflects the sum of all processes that controls the GBCA delivery to the tissue of interest (e.g. injection rate, cardiac output, regional blood flow, etc). Contralateral regions with high contrast in BAT may indicate unilateral carotid artery stenosis (Welker et al., 2015, Jahng et al., 2014).

T_{max} represents the time taken by the tracer bolus to reach the tissue site after arriving at the arterial site (Calamante, 2013). This parameter reflects the delay and dispersion of the bolus and allows the characterisation of macrovascular perfusion (Jahng et al., 2014, Mundiyanapurath et al., 2016, Calamante et al., 2010).

Relative recirculation (rR) is defined as the extent to which the STC recovers back to the baseline from its minimum value (Jackson et al., 2002). It represents the intravascular trapping of GBCA, an attribute observed in malignant capillary beds due to areas of low perfusion pressure (O'Connor et al., 2011).

Thus far, only the perfusion parameters have been discussed. When applied with high temporal resolution and sufficiently long acquisition times, perfusion imaging can additionally quantify parameters related to permeability, which determines the leakage of the GBCA from the intravascular to extravascular extra-cellular space (EES) (Johnson et al., 2004, Sourbron and Buckley, 2013). Detailed discussion on these permeability parameters, their measurements and clinical utility can be found in several published studies (Johnson et al., 2004, Law et al., 2004, Sourbron and Buckley, 2012, Sourbron and Buckley, 2013, Artzi et al., 2015, Essig et al., 2012, O'Connor et al., 2011, Sourbron et al., 2004, Sourbron et al., 2009). The present research prioritises perfusion parameters and, hence, permeability parameters are discussed only briefly, as follows.

The volume of extravascular extracellular space (EES), v_e , is the volume of interstitial space, given as a fraction of total tissue volume. In high grade glioma, leaky BBB leads the blood containing extravasated GBCA to occupy more EES volume, which increases the v_e .

The volume transfer constant, K^{trans} , is the rate at which GBCA is delivered from plasma to the EES, per unit time, tissue volume, and arterial plasma concentration. The permeability surface area product (PS) gives the same rate with respect to capillary plasma concentration (Sourbron and Buckley, 2013, Heye et al., 2016). As an indicator of vascular permeability, K^{trans} has been widely used in tumour grading, identifying disease progression, and assessing treatment response of anti-angiogenic and anti-vascular therapies (Skinner et al., 2016a). A rise in PS indicates angiogenesis; while its decrease confirms response to anti-angiogenic drugs (St Lawrence and Lee, 1998).

For quantifying the above-mentioned perfusion parameters (CBF, CBV, MTT, etc), T_2 - or T_2^* -weighted perfusion imaging is conventionally used, while permeability parameters are quantified with T_1 -weighted perfusion imaging. The variants of perfusion imaging will be introduced in the next section.

2.4. Perfusion MRI: Types

There are two main approaches to perfusion imaging: one that uses exogenous contrast, typically an injected GBCA; and another that uses magnetically-labelled blood as endogenous contrast, without any external injection. Three of the most popular perfusion MRI modalities—two exogenous and one endogenous—will be discussed in this section.

2.4.1. Dynamic susceptibility contrast MRI (DSC-MRI)

Dynamic susceptibility contrast MRI (DSC-MRI) is the standard brain perfusion imaging modality (Sourbron and Buckley, 2013). The process involves injecting an exogenous GBCA and tracking the passage of the bolus with a high temporal resolution T_2/T_2^* -imaging sequence (Jahng et al., 2014). This imaging approach dates back to 1990 when Rosen and colleagues (Rosen et al., 1990) analysed T_2/T_2^* -weighted echo-planar imaging (EPI) along the lines of Axel (Axel, 1980) and others' work on computed tomography (CT).

Due to its paramagnetic properties, GBCA affects the surrounding tissues by altering the local magnetic field. The resultant susceptibility gradient causes a transient signal drop

on T_2/T_2^* images during bolus passage. DSC-MRI assumes an intact BBB and neglects GBCA extravasation to the EES (Jahng et al., 2014, Calamante, 2012, Guzman-de-Villoria et al., 2012). This assumption of intravascular GBCA distinguishes DSC-MRI from the other exogenous perfusion MRI modality, dynamic contrast-enhanced MRI, which is discussed in the next section.

2.4.2. Dynamic contrast-enhanced MRI (DCE-MRI)

Dynamic contrast-enhanced MRI (DCE-MRI) is the standard perfusion imaging modality for regions outside the brain. The process involves injecting an exogenous GBCA and acquiring T_1 -weighted MR images before, during, and after the GBCA injection. This approach of perfusion imaging was proposed by Tofts (Tofts and Kermode, 1991) as an extension to the approaches used in nuclear medicine research. It is routinely applied in perfusion analysis of breast, prostate and muscle tissues (Jahng et al., 2014).

Unlike DSC-MRI, the intactness of the BBB is not a pre-requisite for DCE-MRI. When GBCA extravasates to EES, due to its paramagnetic property, the T_1 -shortening increases the haemodynamic signal. The rate of GBCA extravasation from blood to EES of a tissue is controlled by perfusion, permeability, and the surface area of the capillaries. DCE-MRI can measure these perfusion and permeability-related parameters for regions inside and outside the brain. Figure 2.1 shows typical dynamic data acquired from DCE and DSC-MRI from different brain regions.

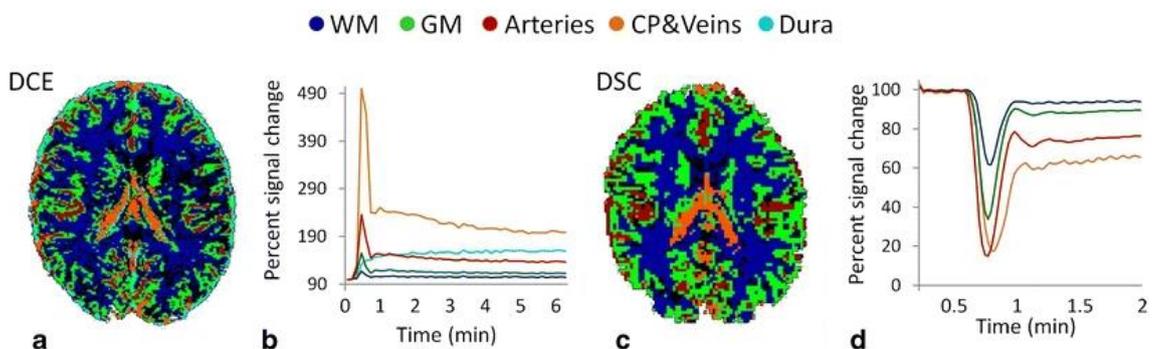


Figure 2.1: (a) Segmented dynamic contrast enhanced MRI (DCE-MRI) image of brain (b) Typical DCE-MRI signals from different brain regions (c) Segmented dynamic susceptibility contrast MRI (DSC-MRI) image of brain (d) Typical DSC-MRI signals from different brain

regions. Figure reproduced from the work of Artzi et al. (2015), with permission from Springer Nature ¹. Abbreviations: GM, grey matter; WM, white matter; CP, choroid plexus.

2.4.3. Arterial spin labelling (ASL)

Arterial spin labelling (ASL) provides a means of measuring absolute cerebral perfusion without any external GBCA injection. The technique, first developed by Williams and colleagues in 1992 (Williams et al., 1992), involves magnetically-labelling the incoming blood and obtaining what is referred to as a ‘label’ image. This ‘label’ image is then subtracted from a ‘control’ or ‘reference’ image obtained with unlabelled blood. The label image shows the passage of protons through capillaries and their diffusion in the tissue water space after a certain delay. When the label is subtracted from the control image, all static effects are cancelled and the resulting difference signal is proportional to blood flow during the delay (Jahng et al., 2014).

Conventionally, DSC-MRI is used for perfusion analysis and DCE-MRI is applied for permeability measurements (Sourbron and Buckley, 2013). However, either of them can be used to obtain both perfusion and permeability information with some modifications to the imaging protocols or analysis (Skinner et al., 2016b, Singh et al., 2007, Sourbron et al., 2009, Law et al., 2004, Johnson et al., 2004). In this thesis, assisting perfusion analysis is the primary focus and, therefore, in the rest of this chapter, the acquisition protocols and analysis of DSC-MRI will be presented.

2.5. DSC MRI: Data acquisition protocol

To track the GBCA bolus, DSC-MRI protocols sample the MRI signal with a high temporal resolution of 1–2 s (Zanderigo et al., 2009, Jahng et al., 2014, Guzman-de-Villoria et al., 2012). As a rule of thumb, the temporal resolution should be lower than

¹ Reprinted by permission from Springer Nature Customer Service Centre GmbH; Springer Nature; *Neuroradiology*; Human cerebral blood volume measurements using dynamic contrast enhancement in comparison to dynamic susceptibility contrast MRI ; Moran Artzi et al, Copyright (2015); Copyright license no: 4962600138009.

the typical MTT value of the tissue to avoid inaccurate STC measurements (Jahng et al., 2014). For image acquisition, an echo-planar imaging (EPI) readout is conventionally performed as it can acquire up to 15–20 slices in a time interval of approximately 1.5 seconds (O'Connor et al., 2011, Calamante, 2012).

Either gradient-recalled-echo (GRE) or spin-echo (SE)-EPI sequences are used as the basis of the DSC-MRI acquisition, depending on the dosage of administered GBCA, investigated vessel size, and required signal-to-noise ratio (SNR). GRE-EPI is sensitive to both micro- and macro-sized vessels, though more to the latter. On the other hand, SE-EPI is more sensitive to micro-vessels (i.e. capillary size vessels) (Jahng et al., 2014). For SE-EPI, signal changes are highest for vessels of 1–2 μm diameter; whereas GRE-EPI exhibits the highest signal for vessels of 3–4 μm diameter (Guzman-de-Villoria et al., 2012). The SNR for GRE-EPI is typically lower than SE-EPI as the large vessels experience static field inhomogeneities, which result in signal losses. Hence, large-vessel contamination is a major disadvantage of GRE-EPI. However, for SE-EPI, the dephasing of spins due to field inhomogeneities is refocused and therefore the signal drop is less apparent and contrast to noise ratio is lower than that of GRE-EPI (Jahng et al., 2014).

Among different GBCAs, gadobenate dimeglumine and gadobutrol are reported to show better performance in distinguishing lesion enhancement and improved diagnostic power (Kuhn et al., 2007, Rowley et al., 2008, Essig et al., 2012). The standard practice is to administer GBCA at a dose of 0.1 mmol/kg or 0.2 mL/kg of the body weight, although double dose can be administered to produce better image quality in cases of diagnostic uncertainty for SE-EPI (Jahng et al., 2014). The GBCA should be injected at a rate of 3–5 ml/s followed by a 20–40 ml saline flush to minimise the bolus spread (Calamante, 2012).

Some of these protocols need adjustments when the integrity of the BBB is compromised, and GBCA can extravasate. GBCA leakage results in 'T₁-shine through': an unwanted T₁-effect in DSC-MRI, which is represented by the rise of the recovery segment of dynamic signal above the baseline. This causes underestimation of CBV as the area above the baseline (contributing as a negative blood volume) is subtracted from the area below the baseline. To reduce this T₁-effect, several adjustments can be undertaken, for example: administering double the conventional GBCA concentration to decrease the injection volume and reduce the extravasated volume; increasing repetition time (TR); reducing the flip angle; injecting a small dose 5–10 min before the scan to pre-saturate

EES and raise the baseline; or administering dysprosium-based GBCAs, which have stronger T_2^* -shortening effects and negligible T_1 -shortening effects (Calamante, 2012, Guzman-de-Villoria et al., 2012, Rempp et al., 1994, O'Connor et al., 2011).

2.6. DSC-MRI: Central convolution equation

The DSC-MRI data obtained from a voxel of interest (VOI) through the above-mentioned acquisition protocol is dependent on two time-dependent functions: the GBCA input to the VOI and the response of the corresponding tissue to that input. This response is dependent on the residue function describing the amount of GBCA remaining in the VOI. In this section, the residue function is introduced prior to the derivation of the central convolution equation, which relates the time dependent output to input and response functions.

2.6.1. Residue function and transit time distribution

The residue function, $R(t)$, describes the probability that a tracer molecule is still in the VOI at time t after entering at $t = 0$ (Jahng et al., 2014). For a tissue region of interest, let us assume that a quantity, Q , of tracer enters at a $t = 0$. The amount of tracer leaving the region between time t and $t + dt$ is therefore: $Q \times h_i(t) \times dt$, where h_i is the transit time distribution: the density function of the transit times necessary for the GBCA particles to traverse capillary tubes of different lengths distributed over the VOI. The amount of tracer remaining in the tube will be given by:

$$Q \times R(t)dt = Qdt - Q \times h_i(t)dt . \quad (2.2)$$

Thus, for the ROI,

$$R(t)dt = dt - h_i(t)dt . \quad (2.3)$$

For the whole period of GBCA passage, equation 2.3 needs to be integrated by parts in the limit of 0 to t . This gives the following two relationships between $R(t)$ and $h(t)$:

$$R(t) = 1 - \int_0^t h(\tau) d\tau . \quad (2.4)$$

$$h(t) = -\frac{dR}{dt} . \quad (2.5)$$

R and h must conform to several constraints to represent real tissue microvasculature. First, $h(0) = 0$, i.e., blood cannot traverse the tissue instantaneously. This means the corresponding GBCA residue should be equal to unity at $t = 0$, i.e. $R(0) = 1$. Second, as $t \Rightarrow \infty$, $h \Rightarrow 0$, $R \Rightarrow 0$ i.e., all of the GBCA must eventually exit the tissue assuming the BBB is intact. This condition also implies that h should be normalised: i.e., its integral should be unity (Sourbron and Buckley, 2013). Third, the tissue residue should decrease as time passes, which is conveyed by an R that is smooth, monotonic, and a decreasing function of time (Mouridsen et al., 2006b). Finally, the form of h must be such as to avoid producing an exponential R , which would imply that there are paths with infinitely small capillary transit times (Schabel, 2012) and also that a fraction of the tracer extravasates instantaneously (Sourbron and Buckley, 2013). This is not possible for DSC-MRI as the indicator is assumed to be intravascular (Leu et al., 2016).

2.6.2. Derivation of the generic equation for perfusion imaging

According to the principle of conservation of mass, the mass of a GBCA within a VOI at time t , $m_{c,voi}(t)$, is the difference between the accumulated masses that have entered via an artery and left the VOI in a time interval $[0, t]$, denoted by $m_{c,voi,in}(t)$ and $m_{c,voi,out}(t)$, respectively:

$$m_{c,voi}(t) = m_{c,voi,in}(t) - m_{c,voi,out}(t) = F \int_0^t C_a(\tau) d\tau - F \int_0^t C_{ven}(\tau) d\tau . \quad (2.6)$$

Here F is the volume flow—assumed to be constant over time— $C_a(\tau)$ is the time course of the GBCA concentration in an artery feeding the tissue of interest (also known as the arterial input function, AIF), and $C_{ven}(\tau)$ is the GBCA concentration time course at the venous outlet.

$C_{ven}(\tau)$ can be calculated from the convolution of $c_{art}(\tau)$ and the transit time distribution, $h(t)$. Therefore:

$$c_{ven}(t) = \int_{-\infty}^{+\infty} C_a(\xi)h(t-\xi)d\xi. \quad (2.7)$$

Substituting equation 2.7 and the delta function, $\delta(t)$, into equation 2.6 gives:

$$m_{c,voi}(t) = F \int_0^t \left[\int_{-\infty}^{+\infty} C_a(\xi)\delta(\tau-\xi)d\xi - \int_{-\infty}^{+\infty} C_a(\xi)h(\tau-\xi)d\xi \right] d\tau. \quad (2.8)$$

Changing the order of the integration and rearranging gives:

$$m_{c,voi}(t) = F \int_{-\infty}^{+\infty} C_a(\xi) \left[\int_0^t (\delta(\tau-\xi) - h(\tau-\xi)) d\tau \right] d\xi, \quad (2.9)$$

and substituting $\tau - \xi$ by τ' inside the square brackets of equation 2.9, leads to:

$$\int_0^t [\delta(\tau-\xi) - h(\tau-\xi)] d\tau = \int_{-\xi}^{t-\xi} [\delta(\tau') - h(\tau')] d\tau', \quad (2.10)$$

Applying this to equation 2.4 and incorporating the delta function:

$$R(t) = 1 - \int_0^t h(\tau) d\tau = \int_0^t [\delta(\tau) - h(\tau)] d\tau. \quad (2.11)$$

Combining equations 2.11 and 2.10 then provides:

$$\int_0^t [\delta(\tau-\xi) - h(\tau-\xi)] d\tau = \int_{-\xi}^{t-\xi} [\delta(\tau') - h(\tau')] d\tau' = R(t-\xi). \quad (2.12)$$

When the blood volume flow F is normalised by the mass of the volume V_{voi} with density

ρ , CBF is obtained. Thus, $CBF = \frac{F}{V_{voi} \times \rho}$ and so $F = CBF \times V_{voi} \times \rho$. Substituting the F

of equation 2.9 for this expression and then substituting the part inside the square brackets of equation 2.9 by equation 2.12 gives:

$$m_{c,voi}(t) = CBF \times V_{voi} \times \rho \int_{-\infty}^{+\infty} C_a(\xi) R(t-\xi) d\xi. \quad (2.13)$$

The GBCA concentration within any VOI, C_t , is defined as the mass of the GBCA per unit volume of the VOI: namely, $C_t = \frac{m_{c,voi}}{V_{voi}}$. So, the above equation can be re-arranged

as:

$$\frac{m_{c,voi}(t)}{V_{voi}} = C_t = \text{CBF} \times \rho \int_{-\infty}^{+\infty} C_a(\xi) R(t - \xi) d\xi. \quad (2.14)$$

The integral of equation 2.14 represents a convolution operation. The tissue concentration C_t is given by the following equation, where \otimes denotes convolution:

$$C_t = \text{CBF} \times \rho \times C_a \otimes R. \quad (2.15)$$

A parameter, k_h , must be included to account for the difference in haematocrit levels between capillaries and large arteries. The following expression gives the value of k_h :

$$k_h = \frac{1 - H(l_v)}{1 - H(s_v)}. \quad (2.16)$$

Here l_v stands for large vessel and s_v stands for small vessel. The typical values of $H(l_v)$ and $H(s_v)$ are 0.45 and 0.25 respectively (Yin et al., 2014, Calamante, 2013). Finally, taking k_h into account, equation 2.15 becomes:

$$C_t = \text{CBF} \times \frac{\rho}{k_h} \times C_a \otimes R. \quad (2.17)$$

Equation 2.17 is the *central convolution equation of perfusion analysis*. It states that the concentration of GBCA in a tissue at time t is the integral of several GBCA contributions that have already entered the tissue at times t' (where $t' < t$), given by $C_a(t')$, and that still remain in the tissue at t , given by $R(t - t')$. The inclusion of CBF in the equation indicates that the tissue concentration is proportional to perfusion. The ratio ρ/k_h serves as another essential scaling factor. The values of ρ and dimensional constant k_h have been modelled as 1.04g/ml and 0.73, respectively (Peruzzo et al., 2011, Rempp et al., 1994, Mouridsen et al., 2006b).

Thus far, different time-dependent functions and their contributions to the signal output of a voxel are introduced. The following section discusses the intermediate steps for analysing this dynamic output signal and thereby quantifying the perfusion parameters.

2.7. DSC-MRI: Analysis

Once the dynamic images have been acquired, background voxels are excluded via a suitable noise threshold. Skull stripping is performed afterwards to extract the brain voxels. After necessary pre-processing, noise reduction, and motion correction, the details of which are outside the scope of this thesis, all dynamic data from the brain voxels are analysed with the steps described in this section.

2.7.1. AIF detection

As evident from equation 2.17, the dynamic signal output of a tissue depends on the convolution of the arterial input (i.e. AIF) with the impulse response function (i.e. $I(t) = \text{CBF} \times R$). Therefore, to estimate the perfusion parameters, the AIF needs to be known along with the measured output data, C_t . Knowledge of the AIF is also essential to isolate the microvascular information about a tissue (i.e. perfusion parameters) from other non-tissue related confounding factors, such as injection protocols, macro-vascular structure, and cardiac output (Calamante, 2013, Calamante, 2012).

Global AIF—an idealised arterial input for every brain voxel—can be detected either manually or automatically. In manual AIF detection, an experienced operator places a cursor on the candidate voxels, typically in the middle cerebral arteries (MCA) or internal carotid arteries (ICA), and uses visual feedback of the STCs or CTCs to identify arterial voxels (AVs) (Patil et al., 2013, Yin et al., 2014, Mouridsen et al., 2006a). For a magnetic field strength of 1.5T, AVs are commonly measured from the M1 segment of the MCA. At 3T, the signal from the M1 region of the MCA can approach the noise floor due to complete de-phasing during the GBCA passage. Consequently, the AIF selected from this segment does not adequately represent its true shape (Yin et al., 2014, Calamante, 2013). To avoid such truncation artefact, the AIF should be measured from the M2 or M3 segments of the MCA. Alternatively, in some brain tumour studies, the venous output function is obtained from the sagittal sinus as a surrogate for AIF, taking the delayed

arrival of GBCA to the vein in account. This alternative to AIF is simple to determine and is robust against distortions caused by partial volume effect (PVE) (O'Connor et al., 2011). Regardless of the detection site, the manual procedure is subjective, time-consuming, poorly reproducible; and often imposes the risk of sub-optimal voxel selection (Mouridsen et al., 2006a, Rempp et al., 1994).

An alternative to manual AV detection is the automatic search for the CTCs with characteristics resembling those of a typical AIF, such as high peak concentration, fast washout of the bolus, high area under the CTC curve, and lower width of the bolus (Yin et al., 2014, Yin et al., 2015, Mouridsen et al., 2006a). It is often assisted by a clustering algorithm, which groups similar dynamics. The cluster that contains the AIFs is identified by applying some of the above-mentioned criteria. The final AIF is then obtained by aligning and averaging the CTCs that belongs to the selected cluster. This automatic AV-detection method discards the human component from the analysis, decreases the computation time, and increases the reproducibility and objectivity (Calamante, 2012).

During automatic AV detection, non-arterial voxels are discarded at first by applying thresholds on criteria that describe different CTC characteristics. To the knowledge of the author, optimal thresholds for these criteria have not been studied systematically, nor has their individual AV-discriminating power been compared. In Chapter 3, these knowledge gaps are addressed.

2.7.2. Brain segmentation using DSC-MRI time-series

After determining the global AIF, regions of interest (ROIs) are placed in GM, WM, or lesion regions, for which perfusion parameters are estimated. Each ROI typically contains several brain voxels whose dynamic signals are averaged to increase the SNR. For an end-to-end automated analysis, this manual placement of ROI can be replaced by more objective, rapid, and reproducible automatic brain region segmentation.

There are several brain segmentation techniques. Manual segmentation techniques use Radiologist expertise to delineate the desired regions; intensity-based methods use the contrast in voxel intensity; atlas-based techniques register the image on an atlas created from a cohort of healthy participants; surface-based methods use the surface feature of a

region; and hybrid methods combine two or more segmentation methods (Despotovi et al., 2015) (further detailed in Chapter 4). Most of these methods are computationally complex and require the data to be transferred to other processing platforms, making the entire analysis time-consuming for clinical setting. In Chapter 4 of this thesis, investigations are performed to overcome these limitations. Brain regions are segmented using the DSC-MRI time-series data; the utility of dimension-reduction in simplifying and accelerating the segmentation is explored.

2.7.3. Impulse response function measurement

After AIF determination and tissue class identification for a given ROI, its perfusion parameters are calculated from $I(t)$: the product of CBF and $R(t)$. To determine I , the contribution of AIF must be removed from the tissue output, C_t , through a process called deconvolution. This deconvolution is ill-posed: a small amount of noise in the data (i.e. C_t and AIF) introduces massive oscillations in the solution (i.e. ringing in R) (Calamante, 2012). To improve the stability of this ill-posed deconvolution, there are two main approaches: model-independent and model-dependent deconvolution, which are detailed as follows.

2.7.3.1. Model-independent deconvolution

Model-independent deconvolution estimates $I(t)$ at every time point, without assuming any model for underlying tissue vasculature. There are two approaches to de-convolve equation 2.17 under this method: Fourier transformation and the algebraic approach.

In the Fourier transformation approach, the convolution theorem of the Fourier transform (FT) is implemented. This theorem states that the convolution of two functions in the time domain is equal to their multiplication in the frequency domain. The impulse response function, $I(t)=F \times R(t)$, can be obtained by Fourier transformation of equation 2.17:

$$\hat{f}\{C_t(t)\} = \text{CBF} \times \frac{\rho}{k_h} \times \hat{f}\{C_a(t)\} \times \hat{f}\{R(t)\},$$

$$I(t) = \text{CBF} \times R(t) = \frac{k_h}{\rho} \times \hat{f}^{-1} \left[\frac{\hat{f}\{C_t(t)\}}{\hat{f}\{C_a(t)\}} \right]. \quad (2.18)$$

Here, $\hat{f}\{ \}$ and $\hat{f}^{-1}\{ \}$ denote the Fourier transform and inverse Fourier transform, respectively. This method of $I(t)$ calculation is highly sensitive to noise and requires application of dampening filters, such as the Wiener filter (Gobbel and Fike, 1994, Rempp et al., 1994).

The algebraic alternative of the model-independent deconvolution method starts by expressing the equation 2.17 in discretised form:

$$C(t_i) = \text{CBF} \times \Delta t \sum_{j=1}^i C_a(t_j) R(t_i - t_j) . \quad (2.19)$$

This equation can be expressed as:

$$\text{CBF} \times \Delta t \times \mathbf{A} \mathbf{r} = \mathbf{c} , \quad (2.20)$$

$$\text{where } \mathbf{A} = \begin{pmatrix} C_a(t_1) & 0 & \dots & \dots & 0 \\ C_a(t_2) & C_a(t_1) & \dots & \dots & 0 \\ \cdot & \cdot & \dots & \dots & \cdot \\ \cdot & \cdot & \dots & \dots & \cdot \\ C_a(t_m) & C_a(t_{m-1}) & \dots & \dots & C_a(t_1) \end{pmatrix} \text{ and } \mathbf{r} = \begin{pmatrix} R(t_1) \\ R(t_2) \\ \cdot \\ \cdot \\ R(t_m) \end{pmatrix}$$

The equation 2.20 was solved with singular value decomposition (SVD) (Knutsson et al., 2010, Ostergaard et al., 1996) with the process described as follows.

The matrix \mathbf{A} can be written as:

$$\mathbf{A} = \mathbf{U} \mathbf{L} \mathbf{V}^T, \quad (2.21)$$

where \mathbf{U} and \mathbf{V} are orthogonal matrices and columns of \mathbf{U} are the eigenvectors of $\mathbf{A} \mathbf{A}^T$, where \mathbf{A}^T is the transpose of \mathbf{A} , and \mathbf{L} is the diagonal matrix of corresponding eigenvalues (Mouridsen et al., 2006b). Since all of the matrices on the right side of equation 2.21 are

invertible, \mathbf{A}^T can be obtained and a solution for \mathbf{r} can be found through the following equation:

$$\text{CBF} \times \Delta t \times \mathbf{r} = \mathbf{V}\mathbf{L}^{-1}\mathbf{U}^T\mathbf{c}. \quad (2.22)$$

Noise in the data introduces high frequency components that result in small, but non-zero singular values of \mathbf{L} , which should otherwise be zero in noiseless conditions. Such small \mathbf{L} values correspond to oscillations in R , making it physiologically unviable by violating the constraint of monotonicity.

To overcome the oscillations in R , small L values are discarded by regularisation, which can be regarded as a noise-filtering approach. Several studies have taken different approaches for selecting how much noise to filter (Calamante, 2012). However, each approach has its own limitations, which are described briefly as follows.

Standard truncated SVD (sSVD) regularises \mathbf{r} by assuming values of \mathbf{L} under a certain threshold, P_{SVD} , to be zero. However, the CBF is highly dependent on the P_{SVD} and a $P_{\text{SVD}} > 20\%$ often results in CBF underestimation (Zanderigo et al., 2009).

Ostergaard et al. proposed P_{SVD} as a function of SNR of the MRI image and found unique thresholds for GM and WM (Ostergaard et al., 1996). To make P_{SVD} independent of the tissue type, Liu et al. proposed it as a function of SNR of the CTC at its peak concentration. However, the optimal number of iterations needed to find such a function was complex to determine (Liu et al., 1999).

Generalised cross validation (GCV) and the L-curve criterion (LCC) were used to find the optimal P_{SVD} for the CTC of each pixel (Sourbron et al., 2004). Despite being comparatively robust methods of threshold selection, they underestimated the CBF in the presence of delay (Zanderigo et al., 2009).

Block circulant SVD (cSVD) allowed local selection of P_{SVD} for each CTC (Wu et al., 2003b). The method could not remove the oscillations and negative values of $R(t)$. Oscillation limited cSVD (oSVD) iteratively repeated cSVD until the oscillation in R came below a certain limit (Bjornerud and Emblem, 2010, Wu et al., 2003b). Ibaraki et al. measured delay with pixel-by-pixel least squares fitting and eliminated its effect by

time-shifting the CTC before performing SVD (Ibaraki et al., 2005). However, the method was more sensitive to noise than SVD (Zanderigo et al., 2009) and, like cSVD and oSVD, was still affected by dispersion.

The Gaussian process deconvolution used Gaussian priors to estimate R for individual time points (Andersen et al., 2002). The resultant R was much smoother than that of other studies and comparable to SVD at high SNR (Zanderigo et al., 2009). The Tikhonov regularisation method (TIKH) applied an oscillation penalty. Compared to SVD, it was less sensitive to dispersion and produced smoother R (Calamante et al., 2003). However, these two methods failed to eliminate the negative values of $R(t)$ (Calamante et al., 2003).

In summary, regularisation through a model-independent approach could not conform to the constraints of monotonicity and non-negativity. Moreover, delay and dispersion caused CBF underestimation and MTT overestimation. For subjects with arterial abnormalities, such as stenosis or occlusion, or moyamoya disease, the model-independent approach may fail to convey explicitly what has caused the overestimation of MTT: the underlying pathological condition, the delay, or the dispersion.

2.7.3.2. Model-dependent deconvolution

The above-mentioned shortcomings of model-independent deconvolution can be overcome by model-dependent deconvolution approach which involves assuming a flexible, yet physiologically plausible, parametric model of the $I(t)$ with several free parameters. The steps involved in basic model-dependent approach can be briefly described as follows:

- To characterise the tissue microvasculature, a trial analytical form of h is assumed with several free parameters q_1, q_2, \dots, q_n ; where n is the number of free parameters. The R is then derived from $h(t; q_1, q_2, \dots, q_n)$ by equation 2.4. To create a realistic h and R , the free parameters are constrained to their physiological limit, as mentioned in Section 2.6.1. Several studies have proposed different parametric models for $h(t)$ (Mouridsen et al., 2006b, Koh et al., 2001, Schabel, 2012, Larsson et al., 2017). However, no study, to the knowledge of the author, has compared the most suitable models. In Chapter 5, three already

published and one proposed model of $h(t)$ are compared to ascertain whether any model provides at least one computational benefit.

- Once the parametric form of residue function, $R(t; q_1, q_2, \dots, q_n)$, is achieved, it is convolved with the manually or automatically determined AIF to create a trial concentration function, \tilde{C}_t , as per equation 2.17. The flow, F , is included as another free parameter in the analytical expression of the \tilde{C}_t .
- Non-linear least square fitting then tries to find values of the free parameters (i.e. F, q_1, q_2, \dots, q_n) that minimise the difference between estimated \tilde{C}_t and measured C_t .
- When the least-squares fitting converges to its global minimum, the optimal values of the free parameters are obtained, which are used to calculate R .
- The perfusion parameters (i.e. CBV, CBF and MTT) are calculated from R with the process described in the next Section 2.7.4.

In summary, the model-dependent approach produces an R that complies with all the constraints imposed on it to be a physically viable representation of tissue microvasculature. Unlike its model-independent variants, which need to estimate R at every time point, the R here is described through two or three parameters only (Mouridsen et al., 2006b). Moreover, Ostergaard et al. showed that for the model-dependent approaches, the flow estimates are independent of vascular delay (Ostergaard et al., 1999). For these reasons, a model-dependent approach is used in the current thesis to estimate perfusion parameters.

2.7.4. Perfusion parameter measurement

Once R is measured by either the model-dependent or model-independent approach, the perfusion parameters are quantified from it by the process described in this section.

2.7.4.1. CBF

The CBF is the maximum value of $I(t)$ (Sourbron and Buckley, 2013). The $R(t)$ is normalised to have a maximum value of unity using CBF as the normalisation factor. If

the arterial input reaches the voxel of interest with no delay, the initial value of the $I(t)$ ($= F \times R(t)$) is the CBF, as $R(0) = 1$. However, if the arterial blood reaches the voxel of interest after a time delay δ , $R(t)$ reaches a maximum at δ , i.e., $R(\delta) = 1$. In such case, the value of $I(t)$ at $t = \delta$ is the CBF (Østergaard, 2005).

2.7.4.2. MTT

The MTT is the expectation value of the transit time distribution, $h(t)$ (Wirestam, 2012). The expectation value, θ , of a density function $f(x)$ can be obtained from the following equation:

$$\theta = \int xf(x)dx \quad . \quad (2.23)$$

Similarly, the expectation value of a TTD is obtained by:

$$\text{MTT} = \int_0^{\infty} th(t)dt \quad . \quad (2.24)$$

Equation 2.24 can be used to establish a relationship between $R(t)$ and MTT. Substituting equation 2.5 into equation 2.24 will result in:

$$\text{MTT} = \int_0^t \tau h(\tau)d\tau = \int_0^t \tau \frac{dR(\tau)}{d\tau} d\tau \quad . \quad (2.25)$$

Integrating by parts will lead to:

$$\text{MTT} = \int_0^t \tau \frac{dR(\tau)}{d\tau} d\tau = tR(t) + \int_0^t R(\tau)d\tau \quad . \quad (2.26)$$

Meier and Zierler showed that, when $t \rightarrow \infty$ then $tR(t) \rightarrow 0$ (Meier and Zierler, 1954).

This leads to:

$$\text{MTT} = \int_0^{\infty} R(t)dt \quad . \quad (2.27)$$

2.7.4.3. CBV

To derive an expression for CBV, first the central convolution equation (equation 2.17) is integrated over time:

$$\int_0^{\infty} C_t dt = \int_0^{\infty} \text{CBF} \times \frac{\rho}{k_h} \times [C_a \otimes R] dt . \quad (2.28)$$

Fubini's theorem (Fubini, 1907) states that

$$\int [f \otimes g] dt = \int f dt \int g dt . \quad (2.29)$$

Applying this to equation 2.29,

$$\int_0^{\infty} C_t dt = \text{CBF} \frac{\rho}{k_h} \int_0^{\infty} C_a(t) dt \int_0^{\infty} R(t) dt . \quad (2.30)$$

From the central volume theorem, $F = \text{CBV}/\text{MTT}$; this and equation 2.27 reduce equation 2.30 to:

$$\int_0^{\infty} C_t dt = \frac{\text{CBV}}{\text{MTT}} \times \frac{\rho}{k_h} \int_0^{\infty} C_a(t) dt \times [\text{MTT}]$$

$$\text{CBV} = \frac{k_h}{\rho} \times \frac{\int_0^{\infty} C_t dt}{\int_0^{\infty} C_a(t) dt} . \quad (2.31)$$

The perfusion quantification steps mentioned above can be executed either for a manually selected ROI placed at GM, WM, or lesion pixels of a subject, or for every brain voxel to create perfusion parameter maps for the brain. In the present study, manual ROIs are placed at GM and WM regions of a pre-selected slice and perfusion estimates are obtained from them through model-dependent deconvolution.

Automation of the above-discussed steps can ensure an objective, rapid, and reproducible analysis. However, further investigations are necessary in order to make the automatic methods more straightforward, systematic, as well as data- or imaging protocol-independent. In the subsequent methods chapters, investigations will focus assisting the

automation of three intermediate steps of perfusion analysis: AIF detection (Chapter 3), brain segmentation (Chapter 4), and the model-dependent deconvolution (Chapter 5).

Chapter 3: A Comparison of Criteria for Automatic Arterial Input Function Detection in DSC-MRI

3.1. Introduction

For the short acquisition duration of a typical dynamic susceptibility-contrast magnetic resonance imaging (DSC-MRI) scan and low dosage of administered Gadolinium-based contrast agent (GBCA), tissues are generally assumed to be linear and stationary (Sourbron and Buckley, 2013, Sourbron and Buckley, 2012, Koh et al., 2011). This means that the link between the tissue inlet and the outlet flux of GBCA, defined through the transit time distribution (described in Section 2.6.1), is independent of the bolus injection time and injected concentration; and dependent on transit time t only. Hence, like any other linear time-invariant system, the tissue concentration output is defined by the time-dependent GBCA concentration input and the impulse response of a tissue, $I(t)$: the product of cerebral blood flow (CBF) and residue function, $R(t)$. This is shown in the central convolution equation 2.17 (Chapter 2) and Figure 3.1.

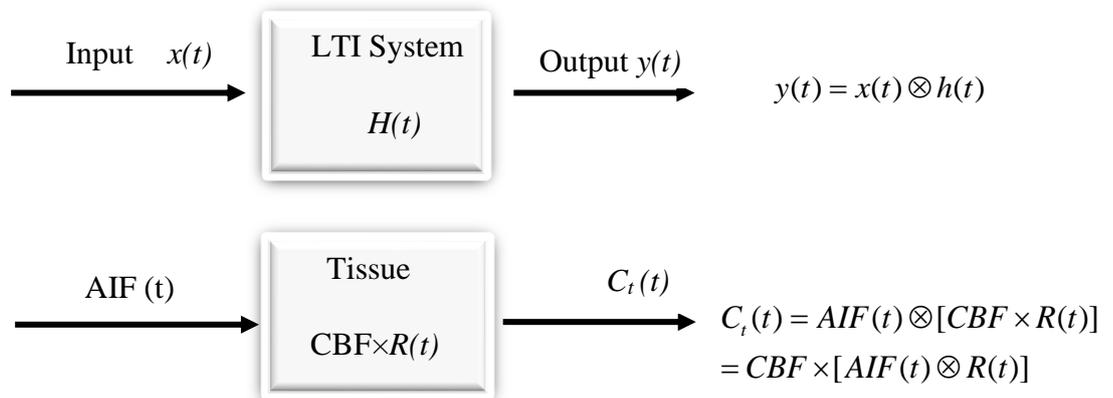


Figure 3.1: Similarity between a linear time-invariant (LTI) system and typical parenchymal tissue. Symbols: $H(t)$, the response function of an LTI system; $AIF(t)$, arterial input function; CBF , cerebral blood flow; $R(t)$, residue function. For both cases, the output is a convolution of the input and the response of the system.

To solve the central convolution equation and quantify the perfusion parameters, it is necessary to have prior knowledge about the time-dependent arterial input. This time

course of the GBCA in an artery feeding the tissue of interest is defined as the arterial input function (AIF). The AIF plays an important role in measuring the cerebral blood volume (CBV), as equation 2.31 requires the area under the AIF to obtain the CBV. Furthermore, an erroneously-quantified, broad AIF would obscure the microvascular contribution towards the broadening of the peak of the tissue concentration time curve (CTC), resulting in underestimation of mean transit time (MTT). Accurate AIF detection is, therefore, an essential preliminary step for perfusion parameter quantification (Calamante, 2013, Calamante, 2012).

AIF can be detected either manually or automatically. As described in Chapter 2 (Section 2.7.1), manual detection involves visual inspection of CTCs in regions containing the middle cerebral artery (MCA) and internal carotid artery (ICA). The AIF is then identified using several shape characteristics, such as a high area under the CTC curve (AUC), high peak concentration (PeakConc), low time to peak (TTP), low bolus width (measured by full width at half maximum, FWHM), and quick washout (indicated by a low first moment, FM). This manual search is subjective, time-consuming, poorly reproducible, and runs the risk of selecting sub-optimal voxels (Patil et al., 2013, Yin et al., 2014, Mouridsen et al., 2006a). For example, in acute stroke patients, a manual process fails to identify the necessary local AIFs (Mouridsen et al., 2006a). Moreover, for some subjects, the CBF estimated from manually-detected AIF misleads about the actual hemispheric location of the pathology (Peruzzo et al., 2011). Therefore, more accurate, objective, rapid, and reproducible automatic alternatives for AIF detection are highly recommended for future routine perfusion analysis (Calamante, 2013, Mouridsen et al., 2006a, Yin et al., 2015, Yin et al., 2014, Peruzzo et al., 2011).

In automatic AIF detection, an algorithm is applied to identify arterial voxels (AVs) based on their afore-mentioned CTC characteristics. Different criteria have been proposed to describe these characteristics; regardless of the criteria used, a threshold must be applied to initially separate the AVs from soft-tissue voxels. For example, if AUC is used as the AIF detection criterion, the mean AUC over all brain voxels is calculated and any voxel with an AUC greater than 60% of the mean brain AUC, for example, is assumed to be arterial.

Detecting AIF solely by applying heuristic thresholds on certain criteria can increase the uncertainty and subjectivity, as the thresholds are often dependent on the datasets and

imaging protocol (Peruzzo et al., 2011, Calamante, 2013). A feasible addition to this thresholding procedure is the time-series clustering of CTC data, which groups arterial CTCs with similar dynamics and distinguishes them from tissue and noisy CTCs with different dynamics. This clustering further refines the threshold-based AIF detection and increases both the objectivity of the method and its robustness against noise (Peruzzo et al., 2011).

In the next section, several previous studies that address the different intermediate steps of automatic AIF detection, such as thresholding, clustering, removal of noise and CTC shape-artefacts, will be reviewed before introducing the aims and research questions of the present study.

3.2. Literature review

Along with tissue-voxel-elimination thresholds, many AV-identification studies have applied clustering algorithms to replace or alleviate the operator bias or manual workload, and at the same time increase the processing speed. Several of these studies are reviewed in the following section.

3.2.1. AIF detection with *k*-means clustering

For rapid, objective, and automatic identification of AIF, Mouridsen et al. used thresholding followed by a standard *k*-means clustering (Mouridsen et al., 2006a). This study describes a successful clustering-based automatic AIF detection method that has motivated many succeeding studies, including the present work.

Mouridsen et al. first computed the area under the curve (AUC) and roughness index (RI) of each CTC. Then, unique thresholds were applied to each of these criteria to discard tissue voxels and voxels with irregular CTCs, respectively. The remaining CTCs were divided into five clusters—based on their different dynamics—using an off-the-shelf *k*-means clustering algorithm (Everitt et al., 2011). The five clusters corresponded to five different tissue regions: grey matter (GM), white matter (WM), arterial blood, venous blood, and ‘other’ regions, for example, ventricle containing cerebrospinal fluid (CSF).

Each cluster was represented by a centroid. The time-series clustering was performed twice. After the first run, the CTCs represented by the centroid with the lowest FM were taken as input to the second run. From the output of the second run, the centroid with the lowest FM was chosen as the one containing the arterial CTCs. The study also analysed the sensitivity of the resultant cerebral blood flow (CBF) to the variation of the AUC and RI thresholds.

The results show that the AIFs obtained with the manual and automatic algorithms had excellent agreement. The agreement between the CBFs of the operator-dependent and this automatic method was similar to that between multiple operator-dependent methods. A later study by Peruzzo et al. (discussed in Section 3.2.3) showed that this method provided CBVs similar to those obtained by simulated AIF for signal-to-noise ratios (SNRs) ranging from 5 to 50. Moreover, for this SNR range, the similarity between automatic and simulated AIF was comparable to other studies (Peruzzo et al., 2011, Rempp et al., 1994, Ibaraki et al., 2005).

A limitation of the study was that the method was applied only to healthy elderly subjects, not patients with brain diseases. Peruzzo et al. (2011) showed that this method produces more false positives (FPs) than other methods for a range of SNRs. Moreover, the method did not search AIF from all the available brain voxels, rather analysed only two manually-selected slices that intersect the MCA. This manual component decreases the objectivity and increases the overall processing time in a clinical setting. To establish a rapid, reproducible, and inclusive AV-detection that searches all brain voxels of all slices, this published method can be modified by either changing the clustering algorithm or reducing its input data dimension (i.e. data mining). In the present work, this k -means clustering is applied to a dimensionality-reduced version of the entire dataspace to identify AIF from all the brain CTCs (see Section 3.4.4 for details).

3.2.2. AIF detection with modified k -means clustering

Like the work of Mouridsen et al. (2006), Bjornerud and Emblem (2010) used k -means clustering for automatic AIF detection, but modified the algorithm in line with the work of Hadjiprocopis *et al.* (2005). Their method was applied to glioblastoma patient data and,

therefore, informs the present study regarding the applicability of *k*-means clustering for AIF selection in the clinical setting.

In this modified *k*-means clustering, the assignment of each CTC (i.e. datapoint) to a cluster-representative centroid was decided through a probability function. The nearer the CTC to a cluster, the higher the probability of its assignment to the cluster centroid. The centroid with the highest PeakConc and lowest FM was considered to contain the candidate CTCs. Out of the candidates, five CTCs with the lowest FM were averaged to create the final AIF. To decrease shape-related errors in AIF, they used a partial volume (PV)-correction based on the steady state concentration of arterial and venous voxels.

A limitation of the work is that the AIFs were not compared to those obtained by other methods, such as those of Rempp et al. (1994), Murase et al. (2001b) or Ibaraki et al. (2005). Another limitation is that the method is only suitable for DSC-MRI imaging protocols that are insensitive to the T_1 -shortening effect of the GBCA (discussed in Section 2.5). Moreover, no thresholding was applied to discard soft-tissue CTCs. As a result, the clustering stage had to deal with a large number of CTC inputs, which may slow down the process and make it unsuitable for analyses that demand rapid processing (such as perfusion analysis of stroke). There is scope to improve the method by investigating the optimal thresholds and implementing data-mining approach to accelerate the clustering (Wang et al., 2006a). These two aspects are covered in the present study.

3.2.3. AIF detection with agglomerative hierarchical clustering

Peruzzo et al. (2011) established a more sophisticated AIF detection approach than those of Mouridsen et al., and Bjornerud and Emblem by improving the thresholding and clustering stage. By comparing different clustering-based AIF detection methods, their work informs the present study about the suitability of different clustering algorithms for AIF detection.

Peruzzo et al. first fitted the first pass of each CTC with a gamma-variate function to remove recirculation (Peruzzo et al., 2011). They then applied thresholds to the AUC and TTP to discard PV-affected and delayed CTCs, respectively. Next, agglomerative hierarchical (AH)-clustering was applied to separate the remaining voxels iteratively into

two clusters represented by two centroids. Of these two, the centroid with highest PeakConc was selected as the one representing AVs. The method was applied to both simulated and clinical datasets.

For the simulated dataset of SNRs ranging from 10–50, the algorithm was able to discriminate AV with fewer FPs than other published methods. For clinical data, the study yielded AIFs with characteristics comparable, if not superior, to manually detected AIF. The spatial locations of the automatically found AVs were validated by expert clinicians.

A limitation of the work is that the removal of tissue voxels depended on a complex gamma-variate fitting, which may increase the time-complexity and negatively impact analyses that require rapid processing, such as acute stroke. Another limitation is that the study applied thresholds only on two criteria: AUC and TTP. There is scope for a systematic investigation of optimal thresholds of different criteria, including those two. The present study explores these aspects.

3.2.4. Comparative analysis of AH, *k*-means, and *c*-means clustering

To overcome the lack of reproducibility and stability of the AIF detection with *k*-means and *c*-means, Yin et al. (2014) proposed an AH clustering method and provided a comprehensive comparison between these three methods.

Although AH was already used for clustering in the work of Peruzzo et al. (Section 3.2.3), Yin et al. applied thresholds to AUC and RI instead of AUC and TTP to discard tissue and motion- or physiological pulsation-affected voxels. An extra criterion—the ratio of the post-bolus steady-state value (SS) to the AUC of gamma-variate fitted CTC first pass (SS: AUC^{1st} ratio)—was used to remove PV-affected CTCs, which was not done by Peruzzo et al. In this study, the feasibility and performance of the method was compared to those achieved with the *k*-means (Mouridsen et al., 2006a) and *c*-means clustering (Murase et al., 2001b). The three methods were applied to both simulated and clinical datasets.

Compared to *k*- and *c*-means clustering, the new method produced AIFs that are more congruent with the simulated AIF—with lower root-mean-square error (RMSE) and percentage of FPs—for a range of SNRs. For the clinical dataset, the method produced

similar AIF to those expected. The final AIF had significantly higher PeakConc and AUC than those obtained by the other two methods. The method was also more reproducible than *k*- and *c*-means.

A limitation of the method was its significantly-higher execution time than the *k*- and *c*-means. Besides this, only healthy subjects were examined, so the clinical efficacy of the method was not confirmed. There is scope to investigate whether a similar AIF-detection can be achieved with a simpler and faster clustering method or whether an optimal AUC threshold can be systematically determined. Both of these aspects are explored in the present work.

3.2.5. Comparative analysis of Ncut, *k*-means and *c*-means clustering

As an extension of their previous work with AH clustering (Section 3.2.4), Yin et al. proposed another clustering method—normalised-cut (Ncut) (Jianbo and Malik, 2000)—for AIF detection to increase its robustness (Yin et al., 2015).

Similar to their previous work, discussed in Section 3.2.4, the method was compared to *k*-means-clustering (Mouridsen et al., 2006a) and the FCM-clustering-based AIF detection (Murase et al., 2001b) for simulated and clinical datasets. Before clustering the clinical dataset, the SS: AUC^{1st} criterion and thresholds on AUC and RI were applied to discard tissue and motion-, PV-, or physiological pulsation-affected voxels.

For the simulated dataset, the Ncut method showed lower FP, higher AUC and RMSE than those of the *k*-means and *c*-means clustering. For the clinical dataset, the Ncut-based detection gave AIF similar to manually detected ones. The AIF had significantly higher PeakConc and lower FWHM than those of other two methods.

A shortcoming of the algorithm was its significantly longer execution time than the *k*-means and *c*-means clustering. Therefore, it is unsuitable for analyses that demand rapid processing (such as, perfusion analysis of stroke). Additionally, similar to their previous work (Yin et al., 2014), only healthy subjects were examined and clinical efficacy was not confirmed. Future work can include investigation of less complex and faster clustering methods and re-assessment of applied AUC thresholds. These aspects are explored in the present study.

3.2.6. Summary of threshold-based AIF detection studies

All of the afore-mentioned studies used some sort of clustering mechanism to aid AV detection. However, many studies have applied thresholding alone, without any clustering to identify AVs. For example, Rempp et al. calculated different features of CTC: PeakConc, moment of maximum concentration (MMC) and FWHM (Rempp et al., 1994). The mean of FWHM and MMC ($FWHM_m$ and MMC_m , respectively) were calculated along with their standard deviations (s_{FWHM} and s_{MMC} , respectively). In the first stage, the CTCs with FWHM and MMC below $FWHM_m - 1.5 \times s_{FWHM}$ and $MMC_m - 1.5 \times s_{MMC}$ were selected as candidate AIFs. Then, 25% of these CTCs with the highest PeakConc were finally selected and averaged to obtain final AIF. Another work used the ratio of PeakConc and MMC as selection criteria; the five voxels with the highest PeakConc to MMC ratios were averaged to produce the desired AIF (Ibaraki et al., 2005).

The performance of these threshold-based methods was highly dependent on the imaging protocol and participating subjects, and thresholds were too stringent in several cases (Peruzzo et al., 2011). Specifically, the method of Rempp et al. produced higher FPs and underestimated CBV for low SNRs. On the other hand, the method of Ibaraki et al. overestimated CBV (Peruzzo et al., 2011). In the present study, clustering is combined with thresholding to avoid these limitations and establish a more robust and operator-independent method.

3.3. Aims

As can be seen from the review above, it is common to apply empirical thresholds to different AIF detection criteria to discard soft-tissue voxels. However, to the author's knowledge, optimal thresholds for these AIF selection criteria have not been studied systematically, nor has the individual effectiveness of each criterion been compared.

The literature review also shows that many AIF detection studies used conventional clustering methods—such as, *k*-means, *c*-means, hierarchical or normalised-cut clustering—to group similar time-series (i.e. CTCs) (Mouridsen et al., 2006a, Emblem et al., 2009, Yin et al., 2014, Yin et al., 2015, Guijo-Rubio et al., 2018). However, most of

the studies have applied their clustering method to just one or two manually-selected slices. When applied to CTCs from all slices during a less-operator-dependent search, the clustering method has to work on a high-dimensional data space. This increases the time complexity of the overall analysis and often provides sub-optimal results (Wang et al., 2006a). A more robust, simpler, and faster alternative to this raw-data-based clustering can be feature-based clustering, where a dimensionality-reduced space containing several CTC features is clustered (Guijo-Rubio et al., 2020, Wang et al., 2006a). The underlying assumption is that CTCs from functionally similar tissue regions will have similar features, and will thus be clustered together in the feature-space. To the knowledge of the author, no study has investigated the feasibility of feature-based clustering for AV detection.

Focusing on the above-mentioned knowledge gaps, the aim of the present work is to assist the automatic detection of a ‘global’ AIF: an idealised representation of dynamic arterial input to every voxel. The following research questions will be addressed in the present chapter:

1. Which of the AIF selection criteria can independently produce plausible AIFs?
2. Can any new criterion perform as well, if not better, than already established criteria?
3. What criteria thresholds can be used to optimally discard non-arterial voxels?
4. How sensitive and specific is the threshold of each effective criterion in discarding tissue voxels?

3.4. Materials and Methods

3.4.1. DSC-MRI data

The data were derived from MRI scans of 35 low-grade glioma patients in a previously published, institution review board approved study (Law *et al.*, 2006) with the acquisition details mentioned in Chapter 1 (Section 1.4). In this thesis, the investigations explore assistive techniques that can contribute to a simplified, fully-automated DSC-MRI analysis. To serve as an exploratory investigation towards that goal, nine subjects were

chosen here; subjects with relatively smooth signal variations, and no noticeable motion-related artefacts were preferred.

Figure 3.2 (a) shows seven brain images obtained at different times, t , during the administration of GBCA in one subject. The figure shows that, as the GBCA enters a voxel, at $t \approx 21$ s, the signal intensity decreases; when the GBCA washes out, at $t \approx 35$ s, the signal intensity starts to recover. The signal intensity does not recover back to the original value due to the recirculation of GBCA. This recirculation does not arise from the return of the traversed GBCA back to the voxel, as the name suggests. Instead, it arises from the simultaneous passage of a fraction of GBCA that has been shunted through the renal and coronary circulations, or organs such as the thyroid, kidneys, or lymph nodes, to reach the voxel via the heart (Calamante, 2013). This recirculation eventually ceases, and the signal intensity returns to the initial baseline value; however, due to the relatively short acquisition time of 60s for the present study, this typically cannot be seen.

Figure 3.2 (b) shows a typical normalised signal intensity time course (S_t / S_{pre}) for an ROI placed in the caudate nucleus, where S_{pre} is the baseline signal and S_t is the raw signal. As per routine MRI perfusion processing, the first six images were discarded, as they did not demonstrate an equilibrium signal for the baseline calculation (Kao et al., 2010); this was also confirmed by visual inspection of STCs. S_{pre} is therefore calculated as the mean of the signal intensity values from the seventh image till the image preceding the bolus arrival.

The STC shown in Figure 3.2 (b) can be divided into three distinct periods: the baseline period showing a constant signal intensity (from 0 – 20s). This represents the native tissue intensity before the arrival of the GBCA. The next period demonstrates a transient intensity variation (from 21 – 37s) representing the effect of GBCA passage through the voxel. The last period is the recirculation (after 37s) showing a second smaller and wider transient intensity variation.

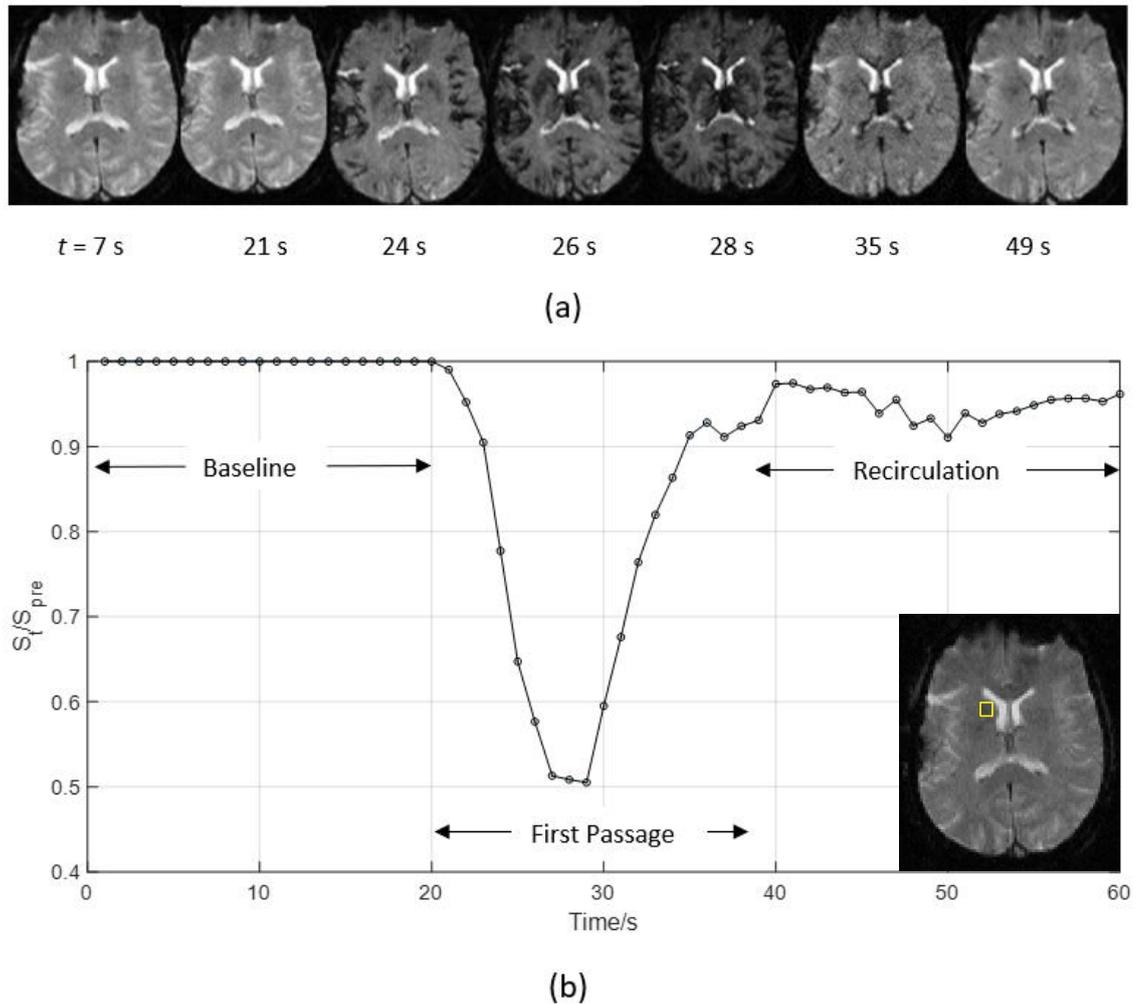


Figure 3.2: (a) DSC-MRI at different time points during GBCA administration. (b) A typical normalised DSC-MRI signal intensity time course obtained by averaging signals from the ROI in the caudate nucleus (shown in the inset) and then dividing by baseline signal intensity. The three different periods: baseline, first passage, and recirculation are marked.

3.4.2. Data pre-processing

Initially a noise threshold was applied to exclude background voxels and isolate the comparatively higher-intensity brain and skull pixels. The ‘regionprops’ function in MATLAB extracted the ‘solidity’ property of the remaining regions. Solidity refers to the area fraction of a region as compared to its convex hull: the smallest convex polygon that can contain the entire region. Skull pixels fill a very small proportion of its convex hull. This low-solidity skull region was then removed applying a threshold; suitable background- and skull-elimination thresholds were empirically determined from a

preliminary study on three subjects. After thresholding, the dynamics only contained typical brain voxels: GM, WM, arteries, veins, ventricles, and lesions (if any), as shown in Figure 3.2(a) and inset of 3.2(b).

The STC was then converted to CTC through the process described as follows. Each DSC-MRI signal was assumed to be related to concentration, C , through the following equation (Patil et al., 2013):

$$S_t = S_{pre} e^{-\Lambda(C)}, \quad (3.1)$$

where $\Lambda(C)$ is a function of concentration that depends on whether GBCA is present in large or small vessels.

For arterial bulk blood (i.e. large vessels), the relaxivity was represented by a quadratic equation that better reflects the relationship between concentration and change in relaxation rate than a linear relationship (Patil et al., 2013, Patil and Johnson, 2013). The form of $\Lambda(C)$ was given by:

$$\Lambda(C) = (qC^2 + pC) \times TE, \quad (3.2)$$

where q and p are constants that depend on the external magnetic field strength, B_0 , and TE is the echo time of the acquisition pulse sequence. At a B_0 of 1.5 T, $q = 0.74 \text{ s}^{-1}\text{mM}^{-2}$ and $p = 7.2 \text{ s}^{-1}\text{mM}^{-1}$; at $B_0 = 3 \text{ T}$, $q = 2.61 \text{ s}^{-1}\text{mM}^{-2}$ and $p = 0.49 \text{ s}^{-1}\text{mM}^{-1}$ (Patil and Johnson, 2013, Patil et al., 2013). The present work used a B_0 of 1.5 T and a TE of 47 ms.

After conversion of signals to CTCs, a noise-filtering step was added to discard the CTCs that suffered from PVE-, noise-, or motion-related artefacts. Such artefactual CTCs demonstrate a high RI, defined as:

$$RI = \int_0^T (C''(t))^2 dt, \quad (3.3)$$

where $C''(t)$ is the second derivative of CTC with respect to time (Guzman-de-Villoria et al., 2012, Yin et al., 2014) and T is the total acquisition time. Twenty five percent of the CTCs with the highest RI were removed as per the practice of several published automatic AIF detection studies (Yin et al., 2014, Mouridsen et al., 2006a, Peruzzo et al., 2011, Yin et al., 2015). Additionally, CTCs with an apparent bolus arrival time (BAT) below 15 s were discarded since the bolus typically arrived between 15 to 20 s in the present imaging

protocol. Any CTC showing a peak at or after the 45th time point (i.e. 45 s) was rejected; this is because CTCs with such late peaks cannot originate from AVs when the acquisition time is 60 s.

3.4.3. Feature extraction

Several criteria (or features) are thought to characterise AIF (Mouridsen et al., 2006a, Yin et al., 2015, Peruzzo et al., 2011). In this work, ten criteria are extracted from each CTC and their effectiveness as well as optimal thresholds are investigated. The criteria are defined as follows.

1. **Area under the CTC (AUC):** The high transient signal drop in the first passage and recirculation for AVs is reflected by a high AUC of the arterial CTCs (Mouridsen et al., 2006a, Yin et al., 2015). The AUC of any CTC, $C_i(t)$, was calculated using the following equation:

$$\text{AUC} = \int_0^T C_i(t) dt \quad . \quad (3.4)$$

2. **Full width at half maximum of the CTC (FWHM):** GBCA in AVs experiences no dispersion and thus arterial CTCs demonstrate lower spread or FWHM in the bolus first pass than those of tissue and venous voxels (Rempp et al., 1994).
3. **Peak concentration (PeakConc):** The high transient signal drop and absence of GBCA dispersion in AVs causes its CTCs to have a higher peak than tissue CTCs.
4. **First moment of the CTC (FM):** FM is defined as the center of gravity of the bolus and calculated using the following formula:

$$\text{FM} = \int_0^T t \times C_i(t) dt \quad , \quad (3.5)$$

In arteries, GBCA washes out quickly; therefore, arterial CTCs show a lower FM (Rempp et al., 1994, Mouridsen et al., 2006a, Peruzzo et al., 2011).

5. **Time to peak (TTP):** TTP is defined as the time elapsed from the injection of GBCA till the time the CTC reaches its peak concentration. For arterial CTCs, the earlier arrival and quick rise of bolus concentration (due to the absence of dispersion) will translate into a shorter TTP (Peruzzo et al., 2011).

6. **Bolus arrival time (BAT):** BAT is the time at which the GBCA CTC first increases beyond the level of the noise (Peruzzo et al., 2011). The tissue voxels receive GBCA after a certain time delay. Consequently, a typical AV will have an earlier bolus arrival and a shorter BAT than tissue voxels.
7. **Mean time duration (MTD):** MTD is proposed in the present study as the mean time duration of the bolus first pass, quantified as the time difference between the BAT and the end of the first pass, T_{end} :

$$MTD = T_{end} - BAT . \quad (3.6)$$

Here, T_{end} is taken as the first time point after the TTP where the concentration value reaches within one standard deviation of the post-bolus concentration (Patil and Johnson, 2011). The post-bolus concentration was calculated by averaging the concentration values of the last ten time points (Yin et al., 2015). Earlier bolus arrival and quick washout should ultimately result in smaller MTDs for AVs.

8. **Mean wash-in rate (MWI):** MWI is defined as the average rate at which the bolus rises to the peak from the BAT. MWI can be calculated from the following equation:

$$MWI = \frac{PeakConc}{TTP - BAT} . \quad (3.7)$$

An arterial CTC quickly rises to the peak and has a high PeakConc; thus, MWI in AVs will be higher than those in other voxels (Newton et al., 2016).

9. **M_{peak} :** Several previous studies included M_{peak} as a parameter to distinguish AVs (Yin et al., 2015, Yin et al., 2014). M_{peak} combines three criteria through the following equation:

$$M_{peak} = \frac{PeakConc}{TTP \times FWHM} . \quad (3.8)$$

The high PeakConc, low TTP, and low FWHM of AVs produce an M_{peak} that is different to other voxels. M_{peak} is included in this study to compare its AV detection ability as a combinational criterion to that of already published and proposed criteria.

10. **Mean washout rate (MWO):** MWO is defined as the rate at which the GBCA concentration drops from $PeakConc$ to the concentration value at the end of the first pass.

$$\text{MWO} = \frac{\text{PeakConc} - C_{\text{end}}}{T_{\text{end}} - \text{TTP}} \quad (3.9)$$

Where, C_{end} is the concentration at the end of the first pass at time T_{end} .

To the author's knowledge, the criteria MWO and MTD have not previously been used but are proposed in this study as two potential 'stand-alone' AIF detection criteria to answer the second research question for this chapter (i.e. can any new criterion perform as well, if not better, than already established criteria?). Figure 3.3 shows a typical AIF with seven of the features listed above.

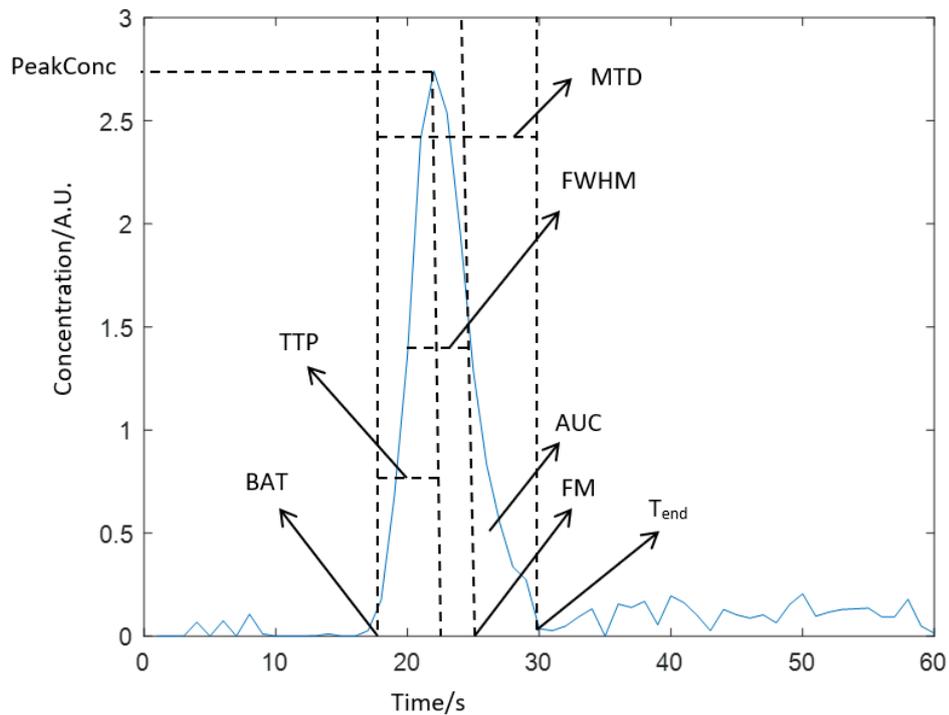


Figure 3.3: A typical arterial input function (AIF) in arbitrary units (A.U.) taken from one of this study's subjects with different features labelled. The AIF is taken from one of the arterial voxels detected through the steps described in Section 3.4.4. Abbreviations: AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; T_{end} , end timepoint of bolus first pass.

3.4.4. Identification of nominally-true AVs

To measure the individual effectiveness of each criterion in distinguishing AV, nominally-true AVs were detected semi-automatically in five stages as follows.

3.4.4.1. Thresholding

The acquisition and processing of the data was performed according to Sections 2.4 and 3.4.2, respectively. Instead of searching for AIF from a manually-drawn ROI in the MCA or ICA like many previous studies, the algorithm searched for nominal AVs from all brain voxels of all slices. Consequently, the automatic algorithm had to search a huge number of voxels. To ease this otherwise extensive and time-consuming search, a thresholding stage was used to eliminate voxels that cannot be considered arterial.

For this ground truth generation step, the thresholds were empirically established based on several previous studies (Peruzzo et al., 2011, Mouridsen et al., 2006a, Yin et al., 2015, Yin et al., 2014). As compared to AVs, tissue voxels show wider boluses (high FWHM), slower washout (high FM), lower PeakConc, and a smaller AUC (Peruzzo et al., 2011, Mouridsen et al., 2006a, Rempp et al., 1994). To exclude such soft-tissue voxels, fifty percent of boluses with the highest FWHM and twenty percent of boluses with the highest FM were rejected. All remaining CTCs with AUCs lower than 40% of the maximum AUC (AUC_{\max}) and with PeakConc lower than 30% of the maximum PeakConc ($PeakConc_{\max}$) were discarded.

The low thresholds identified all the nominal AVs at the cost of also identifying many false AVs. To further refine the algorithm and group the similar concentration dynamics, a subsequent clustering stage (see Section 3.4.4.3) was used. Before clustering, the dimension of the raw data was reduced by the process described in the next section.

3.4.4.2. Feature space creation

From the CTCs of the N voxels that remained after thresholding, ten features (as listed in Section 3.4.3) were extracted. Next, the CTC of a voxel was replaced by a ten-dimensional feature vector, where each dimension contained one of the extracted features. These N feature vectors created a new feature space, F . While all the original CTC data create a raw data space, S , of dimension $N \times 60$ (as there are 60 time points), the F space has a reduced dimension of $N \times 10$. Thus, the F space had more information in a smaller number of dimensions, potentially making the subsequent computation simpler, faster,

and less sensitive to noise, as claimed by Wang et al. (2006a) and Guijo-Rubio et al. (2020). This F space was taken as the input to the subsequent clustering stage.

3.4.4.3. Clustering

The clustering algorithm used in this work involved distance metrics (e.g. Euclidean distance measurement, detailed later in this section), which are sensitive to magnitude variations, different scales, or units of measure of the data. Features with larger magnitude outweigh others and yield inaccurate cluster outcomes (Mohamad and Usman, 2013). To solve this problem, each feature (i.e. dimension) of the F space was given a uniform scale ranging from 0.1 to 0.9 through min-max normalisation (Mohamad and Usman, 2013).

Standard k -means clustering was applied to the normalised F space with the steps as follows (Arthur and Vassilvitskii, 2007, Lloyd, 1982, Raykov et al., 2016, Taxt et al., 1992, Kanungo et al., 2002):

1. **Initialisation step:** The clustering algorithm started with creating k random cluster centroids, w_j (where, $j = 1, 2, \dots, k$), in the F space. The value of k was set as five, like many similar automatic AIF selection studies, to represent five different type of CTCs putatively corresponding to GM, WM, arterial blood, venous blood and ‘other’ regions (e.g. ventricles) (Mouridsen et al., 2006a, Yin et al., 2014, Bjornerud and Emblem, 2010).
2. **Distance calculation step:** Distances between each data point, F_p , to the centroids (i.e. w_j 's) were calculated. There are several proposed distance measurements for k -means (such as, Euclidean, Chebyshev, Manhattan, Mahalanabis, Spearman, Jaccard, etc.), out of which Euclidean distance is most-commonly used (Raykov et al., 2016) and is chosen here for its simplicity. Each data point was assigned to one of the centroids w_j to which it was closest, with the shortest Euclidean distance. This was done by minimising the following objective function J :

$$J(w, F) = \{D(w_{j,i}, F_{p,i})\}^2 = \left(\sqrt{\sum_{i=1}^n (w_{j,i} - F_{p,i})^2} \right)^2, \quad (3.10)$$

where n is the total number of features.

3. **Centroid updating step:** The centroids were updated by taking the mean of all data points included in a specific cluster calculated from step 2.
4. **Iteration step:** Steps 2-3 were repeated until one of the following conditions was met: no more re-locations occurred for a data point—that is, the distance between each data point to their assigned centroids could not be minimised further—or the algorithm had reached its pre-defined maximum number of iterations of 1,000.

3.4.4.4. Centroid Selection

The clustering stage divided the normalised F space into five clusters with high inter- and low intra-cluster variability. Each cluster is represented through a centroid. The CTCs belonging to the centroid(s) that had the highest AUC or highest PeakConc were taken as the candidate AIFs. These two criteria were chosen as they are representative of the most prominent characteristic of an AIF: the large signal drop.

3.4.4.5. Manual AV selection

At the final stage, all candidate AIFs were inspected visually and those that appeared plausible to be considered as AIF were recorded as the nominal AIFs. This five-stage process was applied to the datasets from all nine subjects.

In MATLAB (R2018a, Natick, MA), ten AIFs were plotted at a time and any CTCs that were accepted by the operators were retained. This process continued until 30 AIFs were selected, on which the above process was repeated to find the best candidate (this time five CTCs were plotted at a time). The process continued until the candidate number was below 10, as would be expected from a typical subject. Figure 3.4 provides a schematic diagram for the entire nominal AIF detection process.

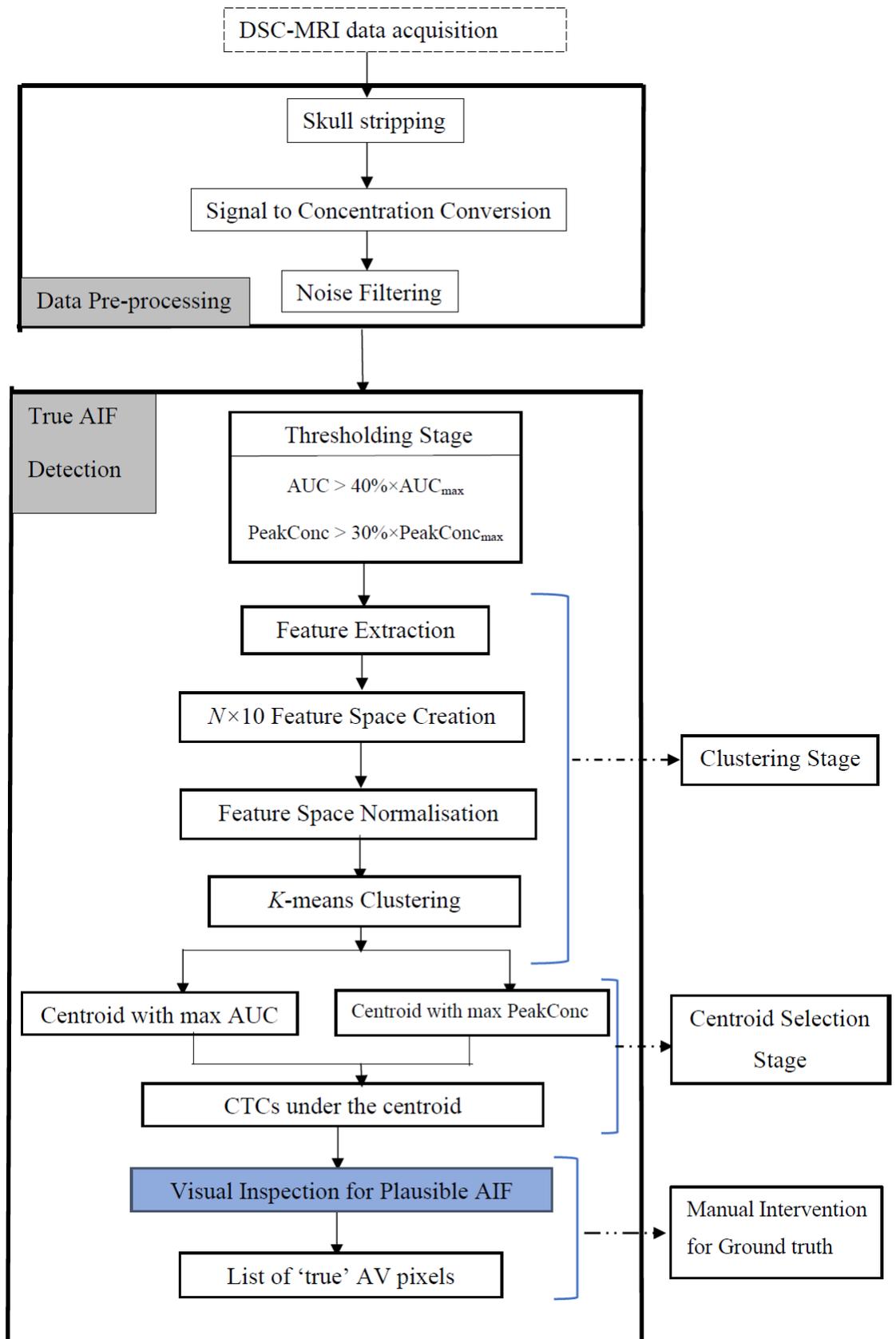


Figure 3.4: Schematic diagram of semi-automatic nominal AV detection. Manual intervention is marked with blue coloured box. Abbreviations: AUC, area under the curve; AUC_{max} , maximum of area under the curve over all voxels; PeakConc, peak concentration; $PeakConc_{max}$, maximum

of peak concentration over all voxels; N , number of voxels left after thresholding; CTC, concentration time curve; AIF, arterial input function; AV, arterial voxel.

3.4.5. Study on the effectiveness of the criteria

To answer the first research question for this chapter regarding the individual AV-detection effectiveness of each criterion, two investigations were performed: initially on a single subject as per the schematic diagram of Figure 3.5 and then on all subjects. One of three subjects that had similar final AIF (detected from Section 3.4.4) was chosen for this investigation. The data acquisition, pre-processing, thresholding, and clustering stages were the same as in sections 1.4, 3.4.2, 3.4.4.1, and 3.4.4.3, respectively.

Like Section 3.4.4.3, the clustering algorithm produced five centroids. The next task was to find which of the ten criteria could effectively identify the centroid that contains the AIFs. Each of the ten criteria was used in turn as a sole centroid selector. For example, when the effectiveness of the AUC was investigated, the centroid with the maximum AUC (AUC_{max}) was selected. On the other hand, to investigate the effectiveness of the FM, the centroid with the minimum FM (FM_{min}) was chosen. The centroids with maximum PeakConc, MWI, M_{peak} , and MWO and minimum BAT, TTP, FWHM, and MTD were selected in their respective turns. Finally, for each criterion, CTCs associated with the selected centroid were aligned and averaged to produce the final AIF. All ten final AIFs were then plotted to visually compare the effectiveness of the criteria to select the correct centroid.

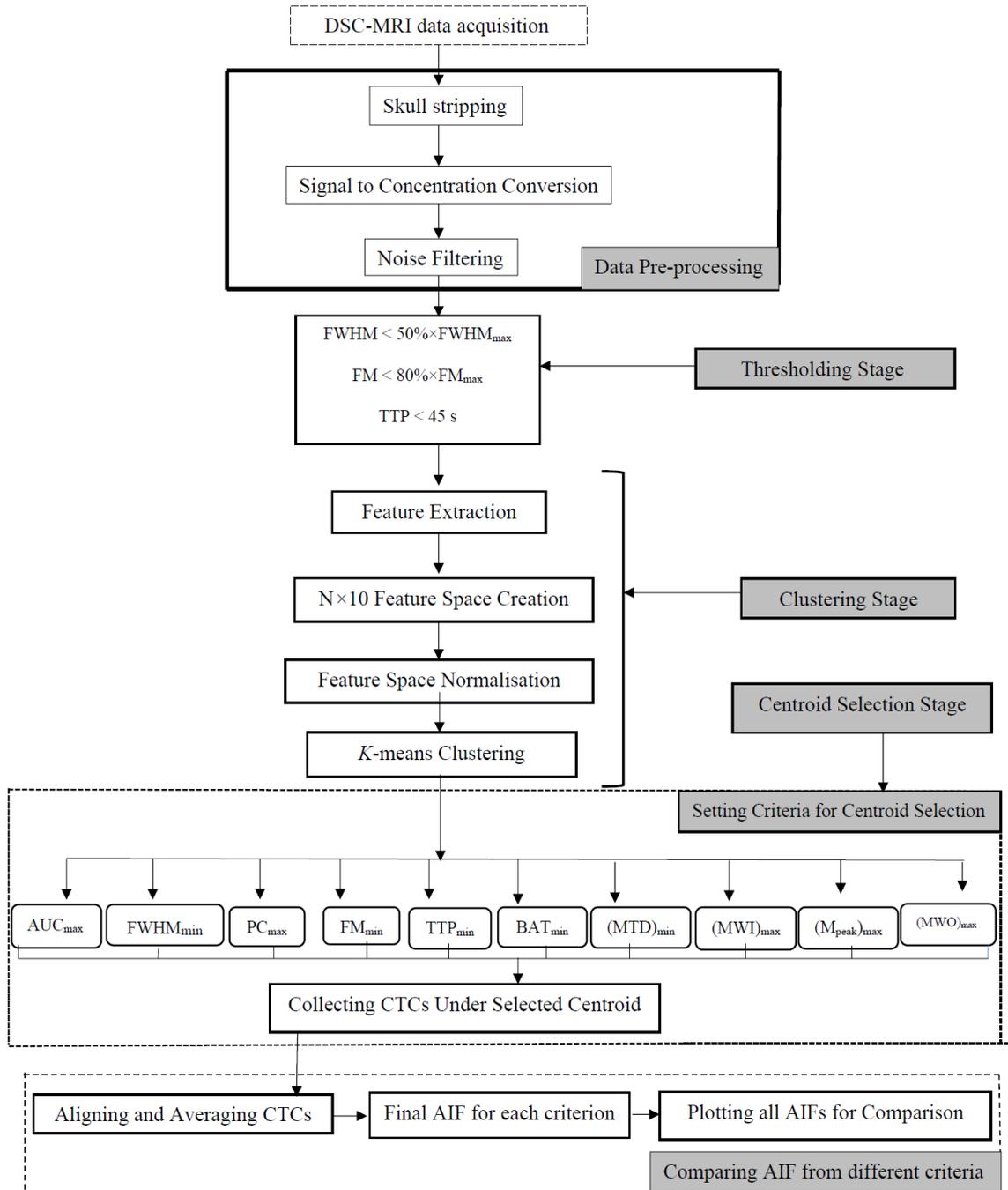


Figure 3.5: Schematic diagram for studying the individual effectiveness of each of the ten criteria. Abbreviations: CTC, concentration time curve; AIF, arterial input function; AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8); MWO, mean wash-out rate. Subscript ‘max’ and ‘min’ refers to the maximum and minimum value of the corresponding criteria for choosing the centroid.

After the above single-subject investigation, the effectiveness of each criterion to distinguish AV was quantified for all subjects. Receiver operating characteristic (ROC) curves were generated for each criterion for all nine subjects using the nominal AVs (identified for each subject in Section 3.4.4). The areas under the ROC curves (AUC_{ROC}) were quantified to evaluate the power of each criterion in distinguishing AVs from tissue voxels, and the mean AUC_{ROC} over all subjects was measured along with the standard deviation (SD). The value of AUC_{ROC} ranges from 0 to 1. The nearer the AUC_{ROC} to 1, the greater the power of a criterion in discriminating AVs from other voxels (Bland, 1995).

The published criteria that produced average AUC_{ROCS} lower than 0.5 were considered unsuccessful in independently selecting realistic AIFs and were excluded from the subsequent investigation that identified the optimal threshold for each effective criterion.

3.4.6. Investigation of Optimal Threshold for effective criteria

This part of the study aims at answering the third and fourth research questions for this chapter (i.e. what criteria thresholds can be used to optimally discard non-arterial voxels; and how sensitive and specific is the threshold of each effective criterion in discarding tissue voxels?).

For the effective criteria, the thresholds to optimally discard the soft tissues were quantified as follows. A cut-off point was identified from the ROC curve of each effective criterion created in Section 3.4.5. The cut-off was set at the point with highest Youden index calculated with the following formula (Krzanowski and Hand, 2009):

$$\text{Youden Index} = \text{Sensitivity} + \text{Specificity} - 1. \quad (3.11)$$

For each subject, the optimal threshold of each criterion was presented as a percentage difference from the mean criterion value over all the brain voxels. The mean and SD of these optimal thresholds were reported along with their mean (SD) sensitivity and specificity in distinguishing and discarding non-arterial voxels.

3.4.7. Statistical Analysis

The AUC_{ROCS} of the criteria were analysed statistically to assess their effectiveness. The present experimental design has one measurement variable— AUC_{ROC} —and two nominal variables. One of the nominal variables is ‘criteria’ with ten categories to represent ten criteria. Another nominal variable is ‘subject’ with nine different categories. Each value of one nominal variable (i.e. criterion) is found in combination with each of other nominal variable (i.e. subject). Due to this design, two-way analysis of variance (ANOVA) was performed to test the hypothesis that the population means of AUC_{ROCS} are equal for different criteria and for different subjects (McDonald and Delaware, 2009). It needs to be noted that there was only one AUC_{ROC} for each combination of the nominal variables. This caused the two-way ANOVA to be performed without replication, and, therefore, it was assumed that there exists no interaction between the nominal variables.

To verify the above-mentioned hypothesis, 45 different hypotheses had to be checked. To reduce the Type-I error of this multiple testing, the p -value was calculated using Bonferroni correction, where the significance level of 0.05 was divided by the number of hypothesis tests. Hence, the present work used $p = 0.001$ for each pairwise comparison (Bland, 1995). The two-way ANOVA analysis was performed in MATLAB (R2018a, Natick, MA).

3.5. Results

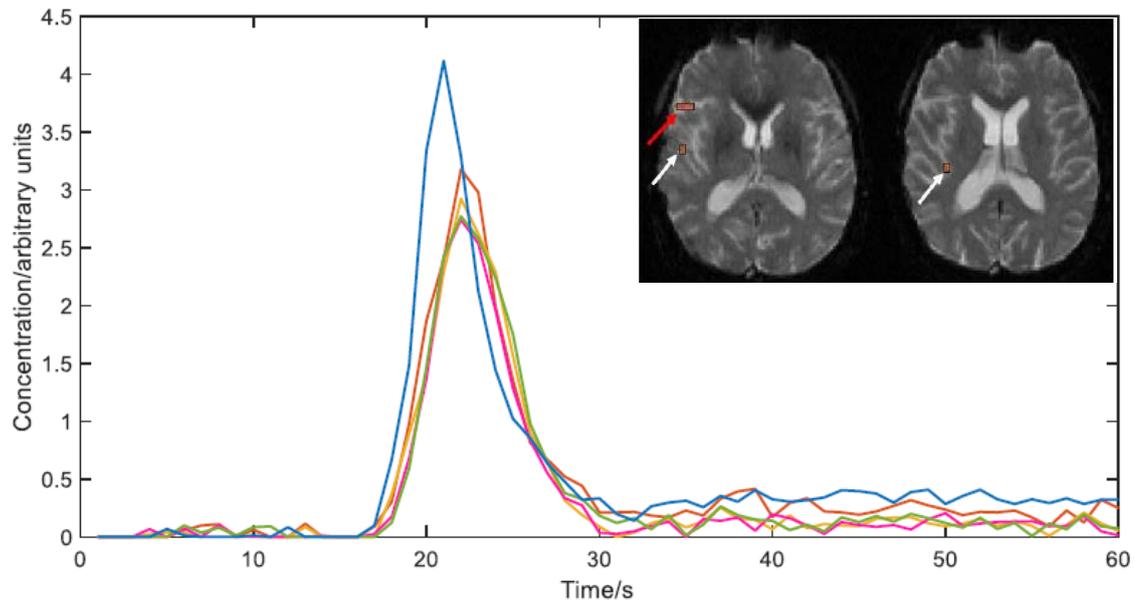
3.5.1. Nominal AVs

Table 3.1 shows the total brain voxels (N_{BV}), number of candidate voxels (N_{CV}) that were obtained after the centroid selection stage of Section 3.4.4.4, and the number of voxels that were finally accepted as AVs after visual inspection (N_{AV}) in Section 3.4.4.5. Comparing the N_{BV} and N_{CV} , it can be stated that the thresholding and clustering steps decreased the number of voxels to be visually inspected by an average of 300-fold with AUC and 292-fold for PeakConc; thus, the need for intensive manual labour is considerably reduced.

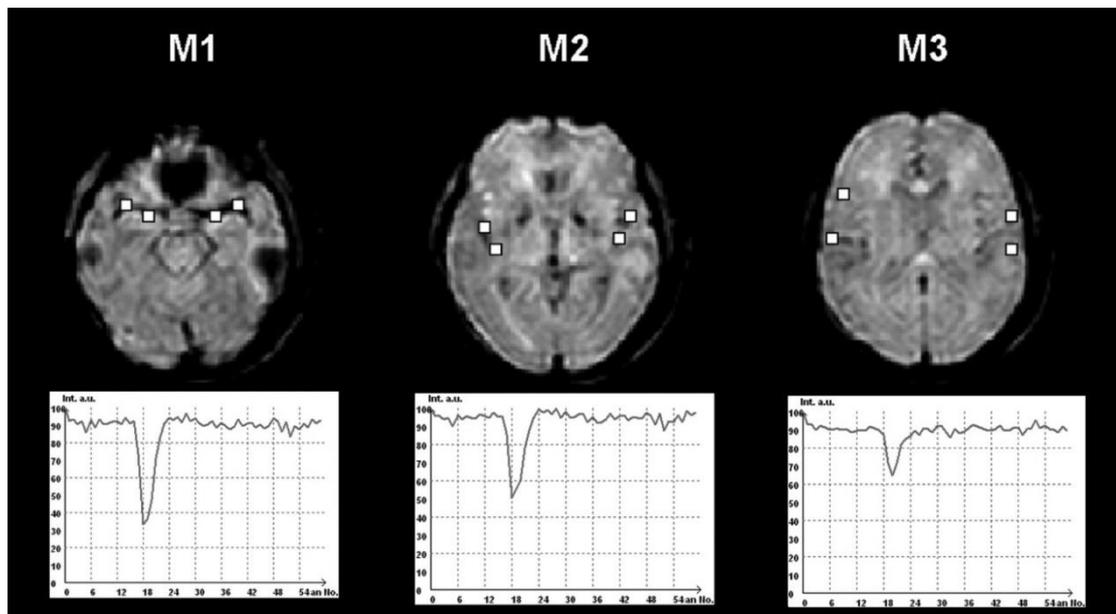
Table 3.1: Number of total brain voxels, N_{BV} (number of voxels left after pre-processing stage, see Figure 3.4), candidate voxels, N_{CV} (number of voxels that were chosen by centroid selection step) and arterial voxels, N_{AV} (number of voxels that were finally selected by visual inspection) for each subject with area under the curve (AUC) and peak concentration (PeakConc) as the centroid selection criteria. In the parenthesis, output voxels of interim steps (i.e. N_{CV} and N_{AV}) are presented as percentages of the voxels input to that step (i.e. N_{BV} and N_{CV} , respectively)

Subject	N_{BV}	Centroid selection criteria: AUC		Centroid selection criteria: PeakConc	
		N_{CV} (% of N_{BV})	N_{AV} (% of N_{CV})	N_{CV} (% of N_{BV})	N_{AV} (% of N_{CV})
Subject 1	37859	192 (0.51)	7 (3.6)	160 (0.42)	7 (4.37)
Subject 2	60390	172 (0.28)	2 (1.16)	74 (0.12)	2 (2.7)
Subject 3	50077	63 (0.13)	7 (11.1)	65 (0.13)	7(0.11)
Subject 4	61626	177 (0.29)	6 (3.4)	219 (0.36)	7 (3.2)
Subject 5	56199	294 (0.52)	10 (3.4)	452 (0.8)	10 (2.2)
Subject 6	50928	340 (0.67)	5 (1.47)	202 (0.4)	5 (2.5)
Subject 7	62150	137 (0.22)	3 (2.19)	180 (0.29)	3 (1.67)
Subject 8	59367	115 (0.19)	6 (5.22)	169 (0.29)	7 (4.14)
Subject 9	57155	140 (0.24)	7 (5.0)	155 (0.27)	7 (4.52)

Figure 3.6 (a) shows the final AIFs for one of the subjects. The spatial locations of the AVs for that subject are shown in the inset. Figure 3.6(b) shows typical spatial locations of AVs at the M1, M2 and M3 segments of the MCA (Zaro-Weber et al., 2012). It can be seen that the locations of AVs for the subject are in good agreement with those found in M2 and M3 segment—middle and right panel of Figure 3.6(b). These spatial locations of AVs were also congruent with those of the large vessels in the sulci, as found by Wismuller et al. (2006).



(a)



(b)

Figure 3.6: (a) Arterial input functions (AIFs) of finally selected arterial voxels (AVs) for one subject, with inset showing the spatial locations of the AVs for the subject with arrows (the red arrow indicates an ROI that contains four AVs); (b) Spatial locations of AVs (white boxes) at different segments of middle cerebral artery (MCA) with their intensity dynamics. The AVs located specifically at M2 and M3 segments are congruent with the identified AVs. Figure (b) is reproduced from the work of Zaro-weber et al. (2012), with permission from Wolters Kluwer Health, Inc ².

²Olivier Zaro-Weber, Walter Moeller-Hartmann, Wolf-Dieter Heiss, et al; Influence of the Arterial Input Function on Absolute and Relative Perfusion-Weighted Imaging Penumbra Flow Detection; *Stroke*, 2012; 43(2): 378-385; <https://doi.org/10.1161/STROKEAHA.111.635458>; Copyright license no: 4963000305742.

3.5.2. Effectiveness of criteria

The results of the effectiveness study for one subject are presented in the first part of this section, whereas the results of the effectiveness study for all the subjects are presented in the second part.

3.5.2.1. Investigation on a single subject

Figure 3.7 shows the finally selected AIFs from each criterion for one subject. Every criterion was used in turn as the cluster centroid selector and the resultant AIF of each was plotted to produce the figure. The criteria AUC and PeakConc yielded AIFs in line with the expected characteristics: high AUC, high PeakConc, early arrival, and quick concentration rise. The criteria M_{peak} and MWI also yielded plausible AIFs, with satisfactory rate of concentration rise and fall, but lower peak values than criteria AUC and PeakConc. All other criteria failed to produce plausible AIFs for the subject.

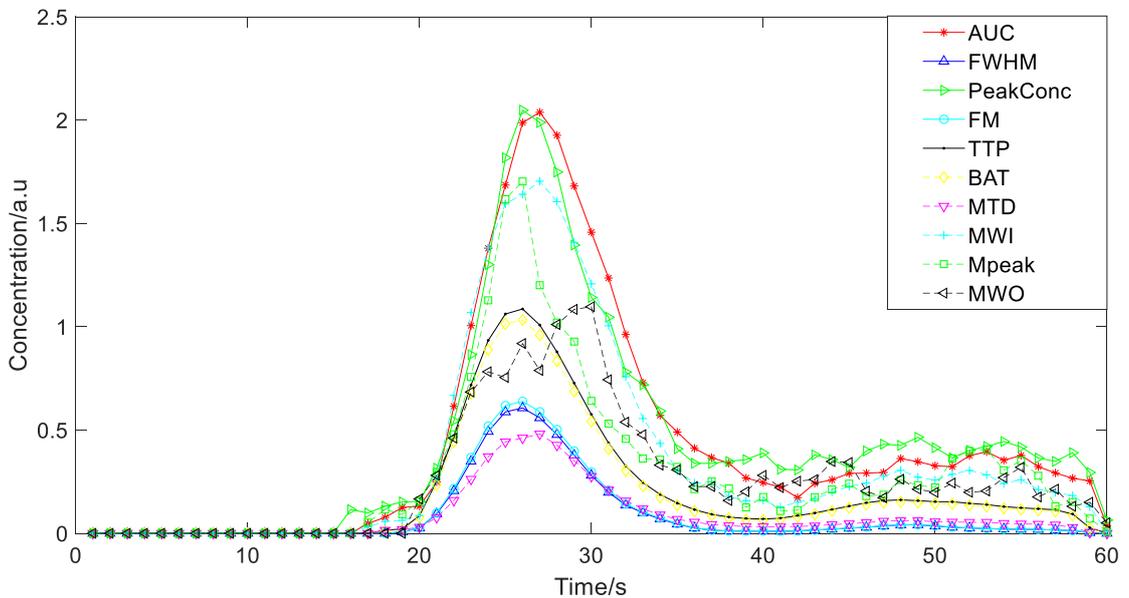


Figure 3.7: Final AIFs achieved with each of the ten criteria for one of the nine subjects (different to the one used for generating Figure 3.6). Abbreviations: a.u., arbitrary units; AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8), MWO, mean wash-out rate.

Table 3.2 gives the feature values for the resultant AIFs obtained from each centroid selection criterion. From the table, it can be noted that when centroid with the maximum PeakConc (PeakConc_{max}) was chosen, the resultant AIF had the highest peak, MWI, and AUC, as well as low FWHM, and high MWO. All the temporal-parameter-based criteria (FWHM, FM, TTP, BAT, and MTD) failed to produce satisfactory AIF.

Table 3.2: Comparison of the ten features of the resultant AIFs obtained with each of the ten centroid selection criteria.

Centroid selection criteria	Feature Values of finally selected AIF									
	AUC	FWHM	PeakConc	FM	TTP	BAT	MTD	MWI	M_{peak}	MWO
AUC _{max}	25.81	8.76	2.04	33.35	27	20	18	0.49	0.0087	0.17
FWHM _{min}	5.12	7.09	0.61	28.77	26	20	14	0.10	0.0033	0.058
PeakConc _{max}	26.41	7.72	2.05	34.3	26	22	16	0.51	0.010	0.18
FM _{min}	5.36	7.09	0.64	28.68	26	20	13	0.11	0.0035	0.061
TTP _{min}	11.93	7.95	1.09	31.64	26	19	14	0.16	0.0053	0.10
BAT _{min}	11.24	7.88	1.03	31.6	26	19	14	0.15	0.0050	0.10
(MTD) _{min}	5.06	7.75	0.48	31.27	27	20	13	0.069	0.0023	0.047
MWI _{max}	21.14	9.16	1.70	32.46	27	20	16	0.24	0.0069	0.16
(M_{peak}) _{max}	17.59	6.01	1.70	33.33	26	22	15	0.43	0.011	0.19
MWO _{max}	15.92	9.56	1.1	34.41	30	31	15	1.1	0.0038	0.15

Note: Feature values obtained from successful criteria are highlighted in bold.

Abbreviations: AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8); MWO, mean wash-out rate. Subscript ‘max’ and ‘min’ refers to the maximum and minimum value of the corresponding criteria for choosing the centroid.

Figure 3.8(a) and (b) show the ROC curves for the shape-parameter-based criteria and temporal-parameter-based criteria, respectively, for the subject. Table 3.3 gives the AUC_{ROC} for each criterion for the subject. Figure 3.8(a) and Table 3.3 illustrate the high effectiveness of AUC, PeakConc, MWI, M_{peak} , and MWO in independently identifying

the AIFs through high AUC_{ROCS} . Figure 3.8(b) and Table 3.3 show the inability of other criteria through low AUC_{ROCS} .

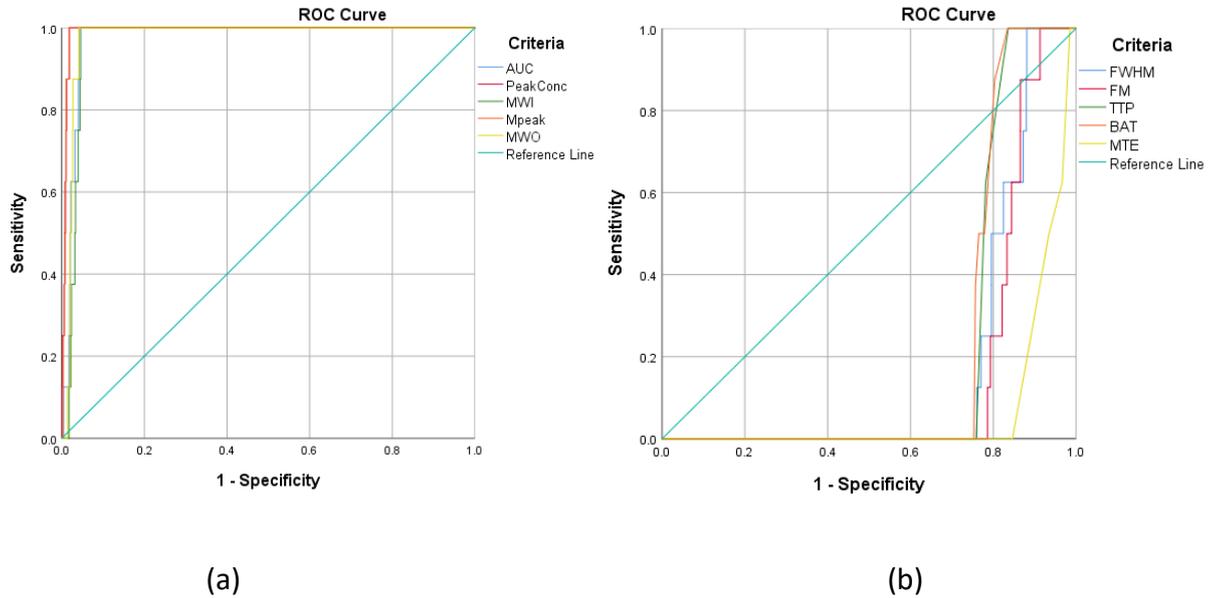


Figure 3.8: (a) Receiver operating characteristic (ROC) curves for AUC, PeakConc, MWI, M_{peak} , MWO for one subject. (b) ROC curve for FWHM, FM, TTP, BAT, MTD for one subject. Abbreviations: AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8); MWO, mean wash-out rate.

Table 3.3: AUC_{ROCS} of different criteria for the chosen subject

Criteria	AUC	FWHM	PeakConc	FM	TTP	BAT	MTD	MWI	M_{peak}	MWO
AUC_{ROC}	0.975	0.178	0.992	0.160	0.215	0.223	0.070	0.968	0.991	0.976

Abbreviations: AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8); MWO, mean wash-out rate.

3.6.2.2. Investigation on all subjects

Table 3.4 shows the mean AUC_{ROC} with the SD for each criterion over all subjects. Figure 3.9 presents these data in a bar-chart. For all nine subjects, the temporal-parameter-based

criteria FWHM, FM, TTP, and BAT failed to identify the AIFs, as can be seen from their low AUC_{ROC} . The shape-parameter-based criteria AUC, PeakConc, M_{peak} , and MWI showed AUC_{ROC} that are significantly higher than those of temporal-parameter-based criteria ($p < 0.001$, exact p -values given in Appendix 1: Table 1). All of the shape-parameter-based criteria were equally effective, with no significant difference between their AUC_{ROC} ($p > 0.001$, exact values given in Appendix 1: Table 1). MWI gave a comparatively lower AUC_{ROC} with a high SD, indicating the uncertainty of AV identification solely with MWI. Out of the proposed two criteria, MTD and MWO, although the former failed in identifying AV (as the mean $AUC_{ROC} < 0.5$), the latter provided AUC_{ROC} similar to those of other effective criteria ($p > 0.001$, exact p -values are given in Appendix 1: Table 1).

Table 3.4: Mean AUC_{ROC} (SD) of different criteria for all nine subjects.

Criteria	AUC	FWHM	PeakConc	FM	TTP	BAT	MTD	MWI	M_{peak}	MWO
AUC_{ROC}	0.98 (0.01)	0.16 (0.05)	0.99 (0.01)	0.15 (0.04)	0.17 (0.06)	0.22 (0.14)	0.067 (0.03)	0.77 (0.21)	0.99 (0.004)	0.97 (0.01)

Abbreviations: AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8); MWO, mean wash-out rate.

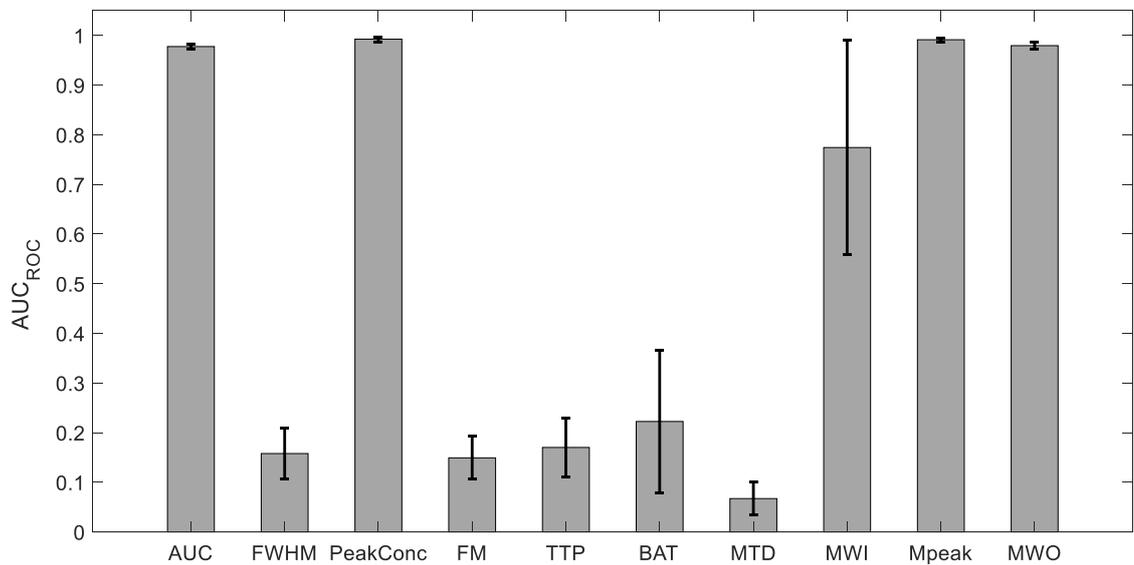


Figure 3.9: Bar chart for mean AUC_{ROC} (SD) for different criteria. The error bars represent SDs. Abbreviations: AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8); MWO, mean wash-out rate.

3.5.3. Optimal threshold for individual criteria

For each criterion, the optimal thresholds were presented as a percentage difference from the mean value of that criterion over all voxels of a subject. Due to the poor performance of FWHM, BAT, FM, TTP, and MTD as individual AV-selection criteria, they were not included in the optimal threshold analysis.

Figure 3.10 shows the optimal threshold ranges in boxplots for the five effective criteria. Table 3.5 gives the mean (SD) of the optimal threshold (first row), and the mean (SD) of sensitivity and specificity in distinguishing and discarding non-arterial voxels (second and third row, respectively) over all subjects. The table shows that a threshold of nearly 60% above the mean AUC can identify an AV candidate with 100% sensitivity and 95.5% specificity. The optimal thresholds for PeakConc and M_{peak} have higher variability and thus higher uncertainties than those of AUC; however, their sensitivity and specificity values are comparable. For MWI and MWO, the optimal thresholds were both above and below the mean value (in Figure 3.10, mean value of a criterion is denoted by percentage threshold of 0%) and SDs were higher (see Table 3.5, first row). This suggests that MWI or MWO cannot individually remove tissue voxels with high certainty. The low sensitivity with high SD of the MWI threshold also indicates its unsuitability for discarding tissue voxels.

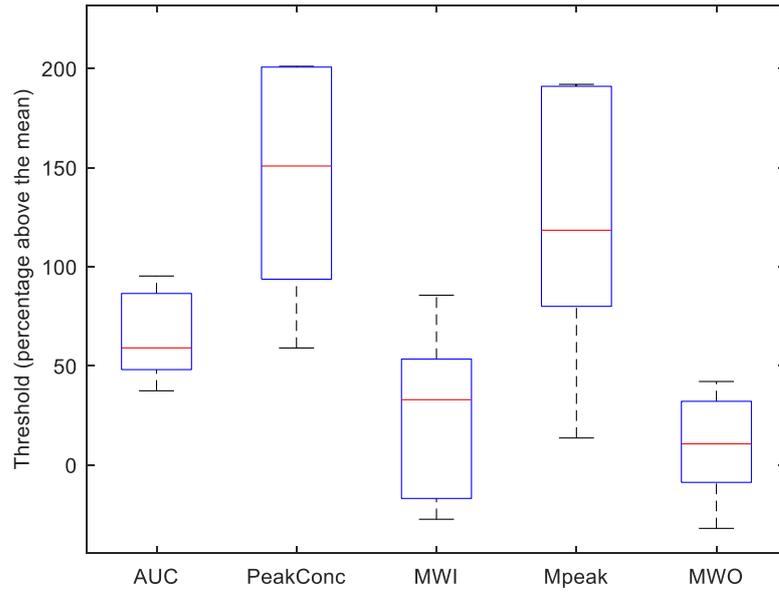


Figure 3.10: Box-plot of optimal thresholds for different criteria shown as percentage above the corresponding criterion mean over all voxels. Median of the optimal thresholds are represented by red lines; interquartile range by blue boxes; the upper and lower quartile of the threshold values by upper and lower whiskers. Abbreviations: AUC, area under the curve; PeakConc, peak concentration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8), MWO, mean wash-out rate.

Table 3.5: Mean (SD) of optimal threshold (represented as % above the overall criteria mean) together with the mean (SD) of achieved sensitivity and specificity in distinguishing and discarding non-arterial voxels

Criteria	AUC	PeakConc	MWI	M_{peak}	MWO
Mean (SD) Threshold (% above overall criteria mean)	64.23 (22.52)	145.82 (63.17)	27.17 (42.29)	118.83 (69.92)	18.30 (33.49)
Mean (SD) Sensitivity	100 (0.00)	100 (0.00)	63.6 (41.2)	100 (0.00)	100 (0.00)
Mean (SD) Specificity	95.5 (0.84)	98.4 (1.16)	95.2 (2.27)	98.0 (1.28)	95.5 (1.68)

Abbreviations: AUC, area under the curve; PeakConc, peak concentration; MWI, mean wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8), MWO, mean wash-out rate.

3.6. Discussion

The primary aim of the study was to investigate the effectiveness of ten criteria in identifying AVs independently. For the effective criteria, the study systematically determined the optimal ranges for tissue-elimination thresholds. The study also served as a preliminary investigation into the feasibility of feature-based clustering for automatic AV selection.

AUC, PeakConc, M_{peak} , and the proposed MWO were found to be the most effective AV detection criteria. The reported optimal thresholds of AUC, PeakConc, and M_{peak} can successfully discard soft-tissue CTCs. Detailed discussion of the findings, potential clinical applications, limitations, as well as future scopes of the present work will be presented in the following sections.

3.6.1. Nominal AVs

The semi-automated AV identification obtained a list of plausible AVs as the ground truth for the subsequent effectiveness analysis. Many published studies searched for AIF from only one slice or a pre-selected region (Mouridsen et al., 2006a, Murase et al., 2001a, Peruzzo et al., 2011); whereas the present work performed a more extensive search of AVs from all brain CTCs of all slices. Additionally, the present work used less stringent thresholds compared to other studies (Mouridsen et al., 2006a, Yin et al., 2015, Yin et al., 2014). This causes selection of all possible AIFs at the cost of selecting many false candidates. However, the subsequent clustering step refined the algorithm and helped reduce the number of FPs by grouping features originating from similar dynamics.

The AIF identification described in this work used feature-based clustering instead of clustering of raw CTCs. Feature-based clustering is reported to increase the robustness of the clustering step and decrease the time-complexity (Wang et al., 2006a). For example, the time-complexity of k -means clustering is $O(IkdN)$ where I is the total number of iterations, d is the dimension of each object (i.e. voxel), N is the number of voxels, and k is the number of clusters. Through the transformation from $N \times 60$ -dimensional raw space to $N \times 10$ -dimensional F space, the value of d reduces from 60 to 10 and, consequently, the time-complexity is decreased.

The clustering step decreased the workload of manually inspecting a large number of CTCs that were left after the thresholding stage. However, the number of candidate AIFs to be visually inspected after the clustering step (i.e. N_{CV}) was still high for some subjects, such as subjects 5 and 6 (see Table 3.1). When PeakConc was used as the centroid selection criterion, the N_{CV} was much higher (except for subjects 2 and 6). This might be due to the erroneous inclusion of PV-affected CTCs in the list of AIFs; these CTCs survived the thresholding stage due to their ‘peaked’ shape. However, they may not necessarily have high AUCs, and therefore including AUC as a supplementary centroid selection criterion could decrease the N_{CV} , and thus, the manual workload. Saying that, even with AUC, many N_{CV} are still left to inspect. This is because of the less stringent thresholds. It is expected that, any future studies will be able to eliminate tissue voxels better if they set their unique data-specific thresholds for AUC and PeakConc using the framework presented here; then, their nominal AV identification step will produce fewer N_{CV} s.

Although this procedure of nominal AIF detection is somewhat subjective, it is no more so than the reference-standard method of manual selection. For the sample subject, the spatial locations of the final AVs were congruent with those found in different studies (Zaro-Weber et al., 2012, Wismuller et al., 2006). However, there is scope to execute a more rigorous and systematic quality control. For example, the spatial locations of finally selected AVs could be assessed by multiple expert Radiologists, then comparing the inter-operator agreements to the agreements between individual operators and the automatic algorithm.

3.6.2. Comparison of criteria effectiveness

Several criteria have been applied collectively in many previously published studies to obtain the candidate and then the final AIF (Mouridsen et al., 2006a, Peruzzo et al., 2011, Rempp et al., 1994). However, no study has reported the effectiveness of each criterion. The present work investigated this aspect and showed that AUC and PeakConc are highly effective in producing plausible AIF. All other previously-published temporal-characteristic-based criteria (such as BAT, TTP, FM, and FWHM) failed as independent AV selectors. Nevertheless, when these temporal parameters are combined with certain

shape parameters, such as for the criterion M_{peak} , they can assist the detection of plausible AIFs. The proposed MWO performed similarly to the already published successful criteria—AUC, PeakConc, and M_{peak} —and can thus be considered in any future automatic AV selection studies.

3.6.3. Optimal thresholds

This study established a general framework for the determination of optimal tissue-elimination thresholds. Many previous studies have applied thresholds that were empirical, based on prior insight and expertise of the researchers regarding the data, or dependent on the underlying imaging protocols. However, those studies lacked complete systematic analysis or explicit justification behind the application of those specific stringent thresholds. Only one study investigated the sensitivity of the CBF to the variation of AUC and RI thresholds (Mouridsen et al., 2006a). In contrast to that study, optimal threshold ranges were determined for five criteria (including AUC) without computationally intensive voxel-by-voxel CBF quantification. Hence, the present study offers substantial information in regard to tissue-elimination thresholds without increasing the computational burden. The threshold-determination framework can be used by other centres and future studies to set the specific thresholds for their dataset.

The prescribed optimal threshold range for AUC is congruent with the range that created CBF maps within an acceptable inter-operator agreement in a previous study (Mouridsen et al., 2006a). For other criteria, no such study was found to validate their reported optimal thresholds. However, as the optimal threshold for AUC is consistent with a previous work, thresholds for other criteria are expected to be in the acceptable range for clinical applications. Saying that, the thresholds for MWO and MWI are highly variable. By using larger sample sizes, future studies can further assess their thresholds and improve the proposed framework.

For discarding non-arterial voxels in future studies, the application of optimal thresholds should be followed by a clustering stage. Otherwise, a vast number of CTCs will be left from which the candidate AIFs would have to be selected by a cumbersome manual inspection. This will ultimately increase the computation time and operator bias. To

simplify and accelerate the clustering stage, future studies can opt for feature-based clustering, whose feasibility can be validated from this study.

3.6.4. Clinical applications

Many aspects of the present study can be applied to routine clinical perfusion analysis. The AIF detection described in Section 3.4.4 can be implemented in the clinical setting for robust, rapid, and less-Radiologist-dependent perfusion analysis of different pathologies, such as glioma, ischaemic stroke, multiple sclerosis, and Alzheimer's disease.

For some pathologies, rapid processing and urgent decision making are of crucial importance. For example, in the typical 'time is brain' situations, such as acute ischaemic stroke, approximately 12 km of myelinated fibre is destroyed in one minute, at a rate of 720 km/h (Calamante, 2013). For perfusion analysis of such pathologies, the proposed feature-based clustering can offer a fast and simple AIF detection by reducing the time-complexity of the clustering stage.

The proposed framework for optimal threshold determination can assist any AIF detection task that applies a thresholding stage. Future perfusion studies can use the presented methods on their data to establish data-specific optimal thresholds. At the very least, the presented thresholds can be used as a reference for other sites, patient studies, or imaging protocols. However, the thresholding stage should not be used as a 'black box' to blindly discard soft-tissue CTCs. The CTCs that remain after thresholding should be further assessed by expert operator, or by including at least some low-level manual quality control, to decide whether higher or lower thresholds would be beneficial for the datasets in hand.

The present study can be used to establish an end-to-end automatic AIF detection algorithm. The proposed systematically-determined optimal thresholds can initially discard the tissue voxels; and then feature-based clustering can rapidly group the CTCs. The feature-based clustering can be made more effective, rapid, and robust by including only the features with high AV-discriminating power—identified by the present study.

The performance of the algorithm can then be evaluated using AIF detected manually or semi-automatically (for example, as per Section 3.4.4).

The findings of the study can also be implemented to assist local-AIF detection. For example, published local-AIF detection methods (Tabbara et al., 2020, Willats et al., 2011) can be accelerated if their raw-data-based clustering step is replaced by the proposed feature-based clustering. Some local-AIF detection studies used criteria, such as effective BAT (different to BAT described in 3.4.3), to initially segment the brain into territories that had similar vascular supply (Willats et al., 2011). These territories were iteratively re-defined until each tissue voxel got assigned to a local AIF. Criteria more suitable than effective BAT can be created by considering the criteria effectiveness information presented here. Moreover, the thresholding stage of these local AIF detection studies can be improved by using the proposed optimal-threshold-determination framework.

3.6.5. Limitations

The limitations of the present work are as follows. Only nine subjects were included in the study. Statistical certainty can be affected by this limited number of samples. A future study with a larger sample size can validate the inferences.

The present method only assists the ‘global’ AIF search. For patients with ischaemic stroke or arterial abnormalities, such as stenosis or occlusion, moyamoya disease, this global AIF can be highly affected by delay and dispersion; therefore, using it may result in inaccurate perfusion estimates (e.g. CBF underestimation) (Tabbara et al., 2020, Calamante, 2013). Instead, a local arterial input to the voxel of interest (i.e. local AIF) provides better perfusion estimates by considering the effects of delay and dispersion (Willats et al., 2011). Ample studies are available on isolating or considering the effects of delay and dispersion (Mehndiratta et al., 2013, Mouannes-Srour et al., 2012, Sourbron and Buckley, 2013, Sourbron and Buckley, 2012, Bjornerud and Emblem, 2010). Nevertheless, to the knowledge of the author, the use of global AIF has not been reported to be inappropriate for glioma patients who do not have any history of arterial abnormalities—the patient cohort studied in this work. Local AIF detection is therefore kept as a future scope of this study.

A further limitation is that the work is not entirely free of human intervention. Manual input was required to create a collection of nominal AIFs for the subsequent evaluation (the coloured box in Figure 3.4). AIFs could have been simulated by a gamma-variate function with known scale and shape parameters. The RMSE between candidate AIFs and ‘true’ AIF could then be used to evaluate the accuracy of the AIF detection approach (Peruzzo et al., 2011, Yin et al., 2014, Yin et al., 2015). However, in this study, *in vivo* patient datasets were used instead of simulated datasets, as the latter often fails to represent CTC variations in patient data.

The applied *k*-means clustering algorithm also possesses limitations. First, it is highly dependent on the randomly selected initial centroid (Rasyid and Andayani, 2018). This randomness can reduce its stability and reproducibility, and create sub-optimal results (Raykov et al., 2016). To avoid these limitations, the number of iterations was kept very high, as per many previous AIF selection studies (Bjornerud and Emblem, 2010, Mouridsen et al., 2006a, Yin et al., 2015, Yin et al., 2014). Second, the Euclidean distance used in *k*-means clustering treats the data space as *isotropic*; that is, all clusters are assumed to be spheres with equal radii around their centroids (Raykov et al., 2016). Moreover, the density of each sphere is assumed to be equal. For real subjects, however, the size of AIF, GM, and WM clusters and the number of elements (i.e. voxels) in them should be different; so, the assumption of equal radius and density fails. Additionally, any outlier in the feature space can significantly affect the linear Euclidean feature space and drastically impair the performance. To overcome this limitation, *k*-medoids or *k*-medians clustering can be used instead of *k*-means, and the *Mahalanobis* distance can be used instead of the Euclidean distance (Raykov et al., 2016). Nonetheless, *k*-means clustering was used only to simplify and accelerate the otherwise time-consuming AV search. So, despite the above-mentioned limitations of *k*-means clustering, their effects are expected to be trivial to the overall inferences made in the present study. Moreover, compared to fuzzy *c*-means clustering (Murase et al., 2001a), another simple clustering algorithm, *k*-means clustering is more successful as the latter gave higher agreement between the CBFs achieved with manually selected AIF and those with automatic method (Mouridsen et al., 2006a).

3.6.6. Future scopes

In future improvements of this work, more patients will be recruited to increase the statistical power of the analysis. Additionally, the spatial locations of the finally identified AVs can be assessed by multiple expert Radiologists.

A new criterion can be established by combining the most favourable AV identification characteristics. The effective criteria identified by this study can be used for such combination. Different analytical expressions of the novel criterion can be investigated to obtain the one with the highest power to discriminate AV from tissue and venous voxels.

The feature extraction step (Section 3.4.3) can be extended to extract features that can identify PV-affected voxels. PV-affected CTCs can be simulated and an approach similar to the effectiveness study (Section 3.4.5) can compare the individual PV-discriminatory power of different features. The study can also be extended to assist PV-correction by finding the venous voxels; the ratio of steady state concentrations of the identified AIF and VOF will give the PV-correction factor (Bjornerud and Emblem, 2010).

Brain segmentation can be a potential extension of the presented strategy of categorising regions on the basis of their dynamics. Future studies can collect different CTC or STC features that are unique for different tissue regions. Feature-based clustering can then group functionally similar tissue regions based on their distinctive shape characteristics. This can potentially develop a simple, fast, but accurate automatic tissue or lesion segmentation technique using DSC-MRI data. This idea is explored in Chapter 4 of the thesis.

3.7. Conclusion

In this study, the individual effectiveness of different arterial voxel selection criteria was compared. Besides this, systematically-obtained optimal threshold ranges for soft-tissue elimination were suggested for the effective criteria. Out of the previously-published criteria, area under the CTC, peak concentration, and M_{peak} were able to independently identify arterial voxels. Thresholds suggested for these criteria can discard tissue voxels

with high sensitivity and specificity. One of the proposed two criteria—mean washout rate—was as effective as these three criteria in identifying AV. The optimal thresholds can work as references for any future tissue voxel elimination work. The criteria effectiveness can inform Radiologists about which criteria to prioritise when assessing the arterial candidacy of any CTC. Finally, the study indicates the feasibility of feature-based clustering, an approach that can accelerate and simplify automatic AV identification methods.

Chapter 4: Application of Feature-based Clustering on DSC-MRI data for tissue segmentation

4.1. Introduction

Image segmentation is the process of assigning each image pixel to a unique class. Brain image segmentation distinguishes the grey matter (GM), white matter (WM), cerebrospinal fluid (CSF), artery, vein, and sinus from one another. It is also used for differentiating pathological regions from normal brain tissues. For dynamic susceptibility-contrast enhanced magnetic resonance imaging (DSC-MRI), accurate segmentation is crucial for all the subsequent analyses that lead to the quantification of perfusion parameters, characterisation of pathological regions, assessment of the success of surgical planning, clinical intervention, and treatment response (Despotovi et al., 2015, Hadjiprocopis et al., 2005). Segmentation methods can be categorised as manual, intensity-based, atlas-based, surface-based, and hybrid (Despotovi et al., 2015). These methods are briefly described below.

4.1.1. Manual Segmentation

Manual segmentation requires Radiologists to delineate the target structure based on their expertise using region-drawing tools (e.g. Insight Segmentation and Registration Toolkit (ITK)-SNAP). These methods are intensive, time consuming, error prone, affected by operator bias, and difficult to reproduce, as reported by several intra and inter-operator variability studies (Vansteenkiste, 2007, Collier et al., 2003, Despotovi et al., 2015).

4.1.2. Atlas-based segmentation

In the atlas-based segmentation approach, a labelled reference atlas or template is created from a specific population of healthy subjects. The brain MR images that need to be

segmented are registered to this atlas, usually through affine registration (Despotovi et al., 2015). To create a satisfactory atlas, the population images need to undergo manual delineation. Although it saves time for Radiologists, this approach still suffers from some of the above-mentioned shortcomings of manual segmentation, such as operator bias and lack of reproducibility. This technique may cause regional inaccuracies due to registration errors (Hadjiprocopis et al., 2005), fail to segment complex structures due to their anatomical variability, and face difficulties segmenting brain MR images that contain lesions, as the atlas is obtained from a healthy population.

4.1.3. Surface-based segmentation

Surface-based segmentation with deformable models, first developed by Kass et al., creates closed parametric surfaces to describe the region boundaries (Kass et al., 1988). These parametric surfaces are affected by the shape of the investigated object, the approximation theories that fit the deformation model to the measured data, and constraints imposed on the temporal and spatial variation of the shape. The surface is deformed by two forces, external and internal, through a complicated iterative relaxation process. The external forces are derived from the local edge-based information and attempt to deform the surface towards the desired feature. The internal forces are computed from within the surface to ensure its smoothness during the deformation. This segmentation method is computationally intensive, and dependent on the initial estimates of the approximation theory and edge information. This dependence on edge information also makes the method sensitive to noise (Despotovi et al., 2015).

4.1.4. Intensity-based segmentation

Intensity-based segmentation distinguishes regions based on their inherent contrast in intensities (Despotovi et al., 2015). This segmentation can be achieved through thresholding, region-growing, clustering, or classification methods. Thresholding is the simplest of intensity-based methods, where a threshold is applied on intensity histograms. This method is sensitive to tissue and radiofrequency inhomogeneities (Hadjiprocopis et al., 2005). The region-growing variant extracts connected regions with similar intensities,

starting from a manually-drawn or automatically-initialised seed point. Its challenges include finding the seed points and setting the homogeneity criterion. Thresholding and region-growing methods are sensitive to noise and partial volume effect (PVE). In classification methods, a large number of labelled MR images is used to train the algorithm that is later used to segment test images. This supervised learning method is manual-input-dependent, complex, and time-consuming, as well as sensitive to noise and operator bias (Despotovi et al., 2015). Lastly, with clustering methods, the similar-intensity pixels putatively originating from similar tissue regions are clustered with an off-the-shelf clustering algorithm, e.g. *k*-means clustering, fuzzy *c*-means (FCM) clustering, or hierarchical clustering (Coleman and Andrews, 1979). Instead of using a separate training dataset, these unsupervised learning methods use the available intensity data to train themselves. They are less time consuming, and less sensitive to noise and operator bias than their supervised variants.

4.1.5. Hybrid segmentation

A hybrid segmentation method is created by combining several segmentation approaches. For example, Xue et al. segmented 3D MRI sections by combining minimum-error global thresholding with spatial-feature-based FCM clustering (Xue et al., 2003). Ortiz et al. combined self-organising maps and entropy-gradient clustering to produce an improved MR images segmentation method (Ortiz et al., 2014). More examples of hybrid segmentation methods can be found in the review of Despotovi et al. (2015). To optimise these hybrid segmentation methods, operators need to adjust many parameters using their prior expertise and understanding of the segmentation mechanisms. The combination of different segmentation methods ultimately increases the complexity of hybrid segmentation approaches.

Out of the above-mentioned techniques, intensity-based segmentation through clustering is preferred in this study as it can offer a computationally-simpler, faster automatic alternative to manual segmentation, while still maintaining accuracy. However, for DSC-MRI, as the intensity of any voxel varies as a function of time, the voxel label assigned by intensity-based segmentation varies across different time points. Hence, clustering the

intensity value at a certain time point cannot produce reliable segmentation. Segmentation can be facilitated by taking the dynamic signals as inputs and clustering the voxels with similar signal dynamics. This method is commonly referred to as time-series clustering (Wang et al., 2006a, Guijo-Rubio et al., 2020).

The following section will review several studies that applied clustering to different types of brain-MRI data for tissue segmentation before presenting the aims and research questions of the present work.

4.2. Literature review

To assist automatic tissue segmentation, intensity-based methods have been applied to different MRI modalities. With minor modifications, some of these techniques can be implemented on DSC-MRI data to generate further-simplified, rapid, and robust segmentation. In this section, a few of these intensity-based brain segmentation methods are reviewed.

4.2.1. Intensity-based segmentation with modified *k*-means clustering

Intensity-based segmentation of diffusion-weighted imaging (DWI) was performed by Hadjiprocopis et al. using clustering methods. The method has already been implemented for automatic arterial input function (AIF) detection in DSC-MRI analysis (Bjornerud and Emblem, 2010) and, with a few modifications, can potentially be applied to brain segmentation.

Hadjiprocopis et al. used a modified *k*-means clustering for segmentation instead of the standard statistical parametric mapping (SPM), which is less successful for DWI segmentation due to the low spatial resolution and inherent geometric distortions of the data. The rule for assigning a datapoint to a cluster was modified by including a probability function that prioritised the datapoints nearer to the cluster centroid. The feature space of a voxel included its own intensity value, as well as the intensities of the voxels that were two co-registered slices above and below the slice containing the voxel of interest. This inclusion of additional intensity information from other slices increased

the dimension of the feature space and maximised the orthogonality between different tissues. The method was compared to SPM for both synthetic and real brain data.

The segmentation method was less affected by radiofrequency and tissue inhomogeneities than SPM. For the synthetic datasets, the segmentation method was comparable to with SPM, with similar sensitivity, specificity, and accuracy, with a lower misclassification percentage. For real brain data, there were large areas of agreement between the method and SPM. The main discrepancies were around the ventricles—the method classified fewer voxels as CSF than SPM02.

The study had three limitations. First, including intensity information from other slices increased the complexity; the greater the number of slices, the more times the clustering algorithms had to run separately. Second, for the synthetic data, partial-volume negatively affected the WM segmentation at the WM-CSF border and gave false positives (FP) at the WM-GM border. GM and CSF voxels were misclassified near the GM-WM and CSF-ventricle borders, respectively. Third, for real brain data, fewer voxels were classified as CSF and more voxels were classified as WM compared to SPM (Hadjiprocopis et al., 2005). Nevertheless, the work presents a reproducible and unbiased fully-automated segmentation technique. There is scope to apply this technique to DSC-MRI data. Further work can assess whether this clustering can be applied to dimensionality-reduced DSC-MRI data in order to both simplify and accelerate segmentation. Both of these aspects are explored in the present research.

4.2.2. Intensity-based segmentation of glioma volumes with *c*-means clustering

Instead of clustering the intensity of the MRI data with *k*-means clustering, Emblem et al. applied FCM-clustering to segment the glioma volumes from T₂-weighted, fluid attenuation inversion recovery (FLAIR), and post-contrast T₁-weighted images (Emblem et al., 2009). They used knowledge-based FCM-clustering to divide T₂-weighted images into three classes—glioma, non-brain, and normal-appearing GM and WM—based on one image feature: pixel intensity value. The cluster with the highest mean pixel intensity represented glioma. For FLAIR images, the clustering divided the tissues, based on pixel intensity, into four classes: glioma, non-brain, normal-appearing GM and WM, and fluid and vessels. For T₁-weighted images, a three-class FCM was applied on all tissues.

The knowledge-based FCM compared well with manual methods in its ability to automatically detect glioma regions. It was both more sensitive and two-times faster; however, it gave more FPs than the manual method and is thus less conservative.

The limitation of the study was that the segmentation approach can be affected by inhomogeneity of the B_0 and B_1 magnetic fields of the MRI scanner. Furthermore, the large vessels were not excluded, which caused misclassification of the pixels adjacent to them due to PVE (Emblem et al., 2009). The solutions to these confounding factors were complex and required operator expertise. Nevertheless, the research showed that FCM-clustering can provide simple and objective segmentation, which can be explored for glioma segmentation in the present DSC-MRI data. The clustering method can be extended to investigate its applicability in distinguishing not only lesion regions, but also GM, WM, CSF, and arteries.

4.2.3. Segmentation of DSC-MRI using a single feature

Features extracted from dynamic perfusion data can be clustered to facilitate segmentation. This idea was explored by Bjornerud and Emblem (2010), who segmented GM and WM automatically—without referring to structural data—by clustering a single feature of the dynamic concentration data.

Bjornerud and Emblem removed non-brain voxels prior to clustering by determining the noise level using Otsu's method (Otsu, 1979). All connected voxels above the noise level were selected by a seed-growing algorithm. To eliminate pathological regions, cluster analysis divided the tissues into two classes—normal and pathological—based on a single feature: the pre-contrast intensity. Voxels with high pre-contrast intensity were assumed to be pathological—either tumour or oedema. After eliminating abnormal and non-brain voxels, the remaining concentration time curves (CTCs) were clustered into three classes—blood, GM, and WM—based on one feature: the area under the CTC curve (AUC). The cluster with the highest AUC corresponded to blood, the lowest AUC to WM, and any AUC in between to GM voxels.

The method reliably identified GM and WM and could avoid large-vessel bias, which produces false positives (FP) in structural image-based GM identification. This approach

also simplified the segmentation process as no co-registration of structural and DSC-MRI images was necessary.

The limitation of the study was that it did not include healthy controls, and the masks for normal-appearing tissues may still have contained abnormal tissues (Bjornerud and Emblem, 2010). This simple segmentation approach is extended in the present study by including more features of the signal time curves (STCs) in the feature space to achieve better orthogonality between tissue regions.

4.2.4. Clustering a dimensionality-reduced version of the original data space

The afore-mentioned idea of clustering a multi-dimensional feature space, instead of raw DSC-MRI data, was inspired by the work of Wang et al. (2006b), who proposed a general framework for grouping similar time-series by clustering their dimensionality-reduced versions.

Wang et al. extracted features from a variety of time-series data by applying different statistical operations, such as trend, skewness, periodicity, kurtosis, etc. The study then clustered a dimensionality-reduced version of the raw data space created from these ‘global’ features. An empirical evaluation was performed to compare this characteristics-based clustering (CBC) to the raw-data-based clustering using the benchmark datasets typically used for data mining.

The results with hierarchical clustering showed that the dimensionality-reduced feature space could efficiently cluster time-series with similar patterns. Experiments on the feature-set with self-organising map clustering could identify same features with stable clusters. The study concluded that more robust and accurate outcomes can be achieved by clustering lower-dimension versions of the original data space. Further, the CBC simplified and accelerated the clustering process and showed lower sensitivity to missing data.

The major limitation of the study was that the extracted feature-set was not consistent for all types of time-series data; for every new application, it was necessary to find the most suitable features. A greedy forward search was suggested as a general mechanism for finding the most suitable feature-set for any application in hand. There is scope to apply

this CBC method to DSC-MRI time series clustering for brain segmentation. In the present study, this CBC method constitutes the base of the two proposed feature-based segmentation approaches (for details, see Sections 4.4.7.1 and 4.4.7.4).

4.2.5. A popular dimension-reduction method and its prospects for brain segmentation

As mentioned above, dimension reduction can offer simple and fast clustering of dynamic DSC-MRI data. One of the popular dimension-reduction techniques is principal component analysis (PCA), whose applicability to DSC-MRI analysis is explored by Akbari et al. (2014).

Akbari et al. performed PCA on raw DSC-MRI signals to investigate whether principal components (PCs) provide any complementary information about peritumoural regions, for which conventional perfusion parameters yield little information. A support vector machine (SVM) used PCs to produce heterogeneity scores within these regions.

The variance of the heterogeneity scores indicated the overall heterogeneity score for a subject, which conventional DSC-MRI cannot quantify without a corresponding pathological region. The research identified highly-infiltrated peritumoural areas that were obscured in conventional perfusion analysis. It also reported that the first six PCs are sufficient for capturing 99% of the variance of the original DSC-MRI data.

While the applicability of this study in clinical routines is limited due to the unavailability of PCA and SVM in clinical workstations (Akbari et al., 2014), it proved the utility of PCs in discriminating different perfusion dynamics. The present research extends this work by extracting PCs from DSC-MRI signals from all regions and investigating the ability of PCA-based dimension reduction to discriminate the perfusion dynamics of different tissue regions.

4.2.6. Segmentation using PCA-based dimension reduction and clustering

A more direct application of PCA in segmentation can be found in the work of Kaya et al. (2017), where the segmentation performance of different PCA-based-dimensionality-reduction methods is compared. This work assisted the present study to choose an appropriate PCA-based approach for segmenting DSC-MRI data.

The study aimed at segmenting T_1 -weighted MR images. Along with conventional PCA, four different PCA methods—probabilistic PCA (PPCA), expectation-maximisation-based PPCA (EM-PPCA), generalised Hebbian algorithm, and adaptive PC extraction—were used to reduce the dimensionality of the data prior to applying k -means and FCM clustering. These PCA variants were compared in terms of their reconstruction errors and the Euclidean distances between original and processed images. The PCA method with the lowest error value was considered to be the most powerful.

PPCA and EM-PPCA were reported as the most powerful dimension reduction tools that could maintain most of the information of the raw data, regardless of the image dimension. They avoided the problem of overfitting and missing data as they used probabilistic methods to find the dimensionality-reduced space for large-variance or large-scale datasets. Effective clustering outcomes can be achieved by incorporating k -means clustering with these two PCAs (Kaya et al., 2017).

The complexity of the otherwise powerful PPCA and EM-PPCA limits their applications; nevertheless, this research indicated that clustering can be amalgamated with PCA-based dimension-reduction methods for identifying lesions. In the present study, the proposed dimensionality-reduction-based segmentation approaches were compared with the conventional PCA-based segmentation approach (see Section 4.4.7.3).

4.3. Aims

The aim of the present study is to investigate the applicability of feature-based clustering to DSC-MRI data for brain segmentation. In feature-based clustering, several features extracted from the raw data are used to create a dimensionality-reduced feature space, which is then clustered. For different types of time-series data, feature-based clustering

approaches were reported to be more efficient, faster and more robust than clustering the raw data (Wang et al., 2006b, Guijo-Rubio et al., 2020).

The primary hypothesis of the present research is that a dimensionality-reduced space not only decreases time-complexity but also improves segmentation. A second hypothesis is that the proposed dimensionality-reduction-based segmentation can outperform at least one of the currently available dimensionality-reduction-based segmentation approaches. To test the above-mentioned hypotheses, this study addresses the following research questions:

1. Which signal features are most effective for the segmentation task?
2. Can feature-based time-series clustering approaches provide better segmentation than their raw-data-based variants?
3. Can the proposed dimensionality reduction provide better segmentation than any available dimensionality-reduction approach?

4.4 Materials and Methods

4.4.1. Raw data space (\mathbf{R}) creation

The DSC-MRI data were acquired with the process described in Chapter 1 (Section 1.4). A raw data space, \mathbf{R} , was populated with 256 DSC-MRI signals from manually-selected pixels of the nine subjects. This \mathbf{R} space was segmented in this study with different approaches. The subjects had visually-comparable arterial supply with similar peak and width. Real-subject dynamic signals were used instead of a simulated data, which are often obtained by convolving an assumed parametric form (e.g. exponential) of residue function with a gamma-variate form of AIF (Peruzzo et al., 2011, Mouridsen et al., 2006a, Yin et al., 2015, Yin et al., 2014). Real-subject signal was preferred as the gamma-variate function fails to replicate the recirculation in real DSC-MRI data and is, therefore, often only a partial representation of a real-subject AIF. Additionally, there is no consensus between researchers in regard to the most inclusive parametric form of residue function. The different signals of the \mathbf{R} space are described as follows.

- **GM and WM signals:** Thirty-six GM and thirty-six WM signals were selected from the normal-appearing caudate nucleus and frontal lobe, respectively. As prescribed by Ostergaard et al., typical GM and WM time-series show a signal peak decrease (SPD) of 40% and 17% from baseline, respectively (Ostergaard et al., 1996). For the present study, GM and WM signals had SPDs between 35-50% and 10-25%, respectively.
- **Lesion signals:** A total of 15 signals were taken from visually-identifiable lesion pixels of two subjects.
- **Arterial input function (AIF) signals:** For each subject, an AIF was created by aligning and averaging the CTCs of arterial voxels (AVs) that were chosen manually with the process described in Section 3.4.4 of Chapter 3. These nine AIFs were then converted to signal time estimates using quadratic equations for relaxivity proposed by Patil and Johnson (2013). The non-linear equation was used as it better reflects the relationship between concentration and change in relaxation rate than a linear equation (Patil et al., 2013, Patil and Johnson, 2013, Calamante et al., 2009, van Osch et al., 2003, Bjornerud et al., 2002).
- **CSF signals:** Twenty CSF pixels were identified from different subjects and their signals were collected.
- **PV-affected signals:** Thirty PV-affected signals were created according to the process described in several simulation studies (Peruzzo et al., 2011, Yin et al., 2014, Mouridsen et al., 2006a, Yin et al., 2015). In this process, each PV-affected signal was taken as a linear combination of arterial, GM, and WM signals with random weights.
- **Background signals:** Lastly, 110 constant intensity signals were included as background signals. To resemble typical background signals having signal to noise ratio (SNR) of 30, noise modelled as a zero mean Gaussian function was added, in line with the work of Peruzzo et al. (2011) and Yin et al. (2015).

These STCs were not converted to CTCs as the noise in the concentration (but not signal) is dependent on amplitude and this conversion can distort useful information (Akbari et al., 2014). The baseline, transient drop, and recovery of the signals, even from functionally similar regions, may vary due to the unique vascular arrangements of different subjects. To account for such inter-subject variation, each signal was normalised

by dividing the signal intensity of each time point by the baseline signal intensity, S_{pre} , determined by the method described in Section 3.4.1 of Chapter 3.

The \mathbf{R} space represents a small-scale version of a typical brain slice. Although, in a typical brain slice, many similar tissue pixels are located closer to each other, the 256 signals were randomly distributed over the \mathbf{R} space. Hence, there is no certainty about the proximity of two identical regions. This was done purposefully to make the segmentation task more challenging for the proposed and available segmentation approaches.

4.4.2. Signal time series from different brain regions

Figure 4.1 shows non-normalised DSC-MRI signals from different tissue regions mentioned in Section 4.4.1. All signals, except those taken from background or affected by PV, are taken from one subject. The PV-affected and background signals are simulated as discussed in Section 4.4.1. Noisier version of these \mathbf{R} -space signals were also created by adding zero-mean Gaussian noise to the \mathbf{R} space using the approach taken by Peruzzo *et al.* (2011).

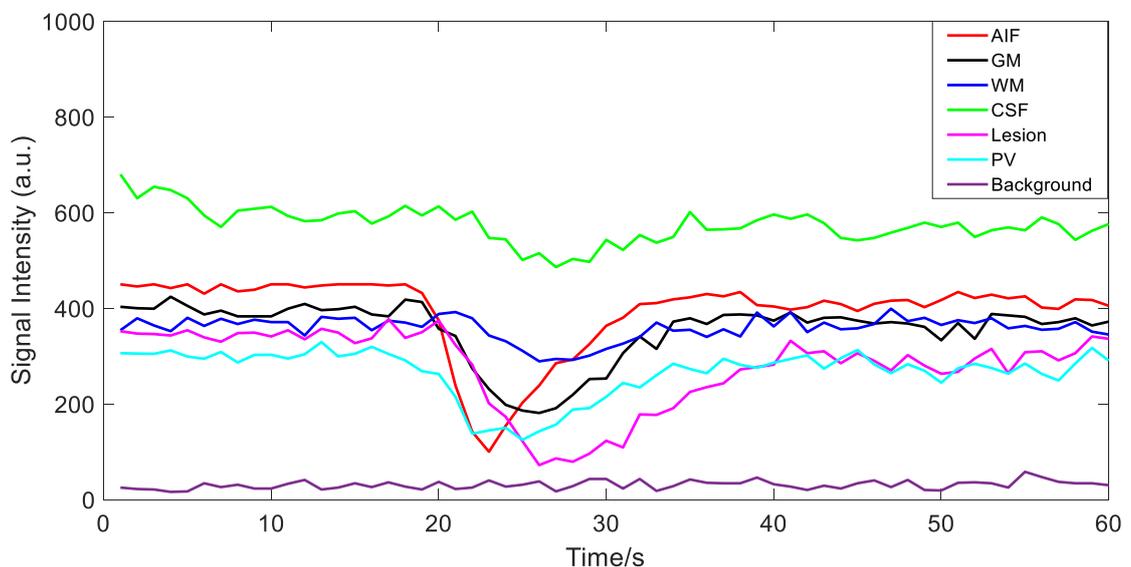


Figure 4.1: Raw, non-normalised DSC-MRI time-series from different brain regions of one subject. The background and partial volume (PV)-affected signals are created as per the discussion of Section 4.4.1. It can be seen that signals from different regions differ in their features, like the initial drop, recovery, width, and baseline intensity. Abbreviations: AIF, arterial input function; GM, grey matter; WM, white matter; CSF, cerebrospinal fluid.

Figure 4.1 shows that the signal time-series from different brain regions are unique and, therefore, they can be categorised using different characteristic features. The next section extracts several such features from the \mathbf{R} -space signals in order to investigate their applicability in distinguishing different brain regions.

4.4.3. Feature extraction

From each element (i.e. dynamic signal) of the above-mentioned \mathbf{R} space, the nine features were extracted. A similar approach was adopted by Wang et al., where several global features (e.g. trend, skewness, kurtosis, seasonality, periodicity, etc.) were extracted from benchmark time-series datasets used for data mining (Wang et al., 2006a) (see review Section 4.2.4). However, in this study, features are extracted specifically for brain segmentation, under the expectation that the distribution of these features will be distinctly different for different tissue classes and similar for identical classes; this is true when each region is supplied with similar arterial input (Calamante, 2013). Although the \mathbf{R} space contained DSC-MRI time-series from different subjects, this assumption is still valid as the chosen subjects had similar AIFs. The extracted features are described as follows.

1. **Effective drop percentage (EDP):** The ratio of maximum signal drop (MSD) to the standard deviation of pre-bolus signal ($SD_{S_{pre}}$) was expressed as a percentage and referred to as the effective drop percentage, EDP:

$$EDP = \frac{MSD}{SD_{S_{pre}}} \times 100\% . \quad (4.1)$$

Signals from each tissue region show unique MSDs. Consequently, EDP will be different for different regions. Furthermore, any voxel with $MSD < 3 \times SD_{S_{pre}}$ is suggestive of either CSF, infarct, or background (Kane et al., 2007). By taking the ratio of MSD to $SD_{S_{pre}}$, EDP allows these regions to be separated from others.

2. **Area over normalised signal time curve (AOC):** The area below the baseline, S_{pre} , and over the normalised STC, $S_{t(norm)}$, was calculated by subtracting the area

under the normalised curve (AUC) from the total area below baseline ($S_{pre} \times T$) with the following formula:

$$AOC = S_{pre} \times T - AUC, \quad (4.2)$$

where T is the total acquisition time, and $AUC = \int_0^T S_{t(norm)} dt$. For normalised STC, $S_{pre} = 1$. Substituting S_{pre} and AUC in equation 4.2 gives:

$$AOC = T - \int_0^T S_{t(norm)} dt. \quad (4.3)$$

Each type of tissue has their unique CBV, which is proportional to the AUC of CTC (equation 2.31 of Chapter 2). Therefore, when analysis is performed on STCs, each tissue region should have a unique AOC.

3. Variance: The variance, σ^2 , of an STC was calculated with the following formula:

$$\sigma^2 = \mu_2 - \mu_1^2, \quad (4.4)$$

where μ_1 and μ_2 are the first and second raw moment of the STC, respectively. This variance indicates the spread of the signal and can be regarded as an alternative measure for the full width at half maximum (FWHM). As tissue and lesion STCs are dispersed differently to arterial STCs, the variance can be used to distinguish them from arterial regions.

4. L₂ Norm: For a time-series $\{S_t\}_{t=1}^T$, the L_p norm is given as follows (Mohajer et al., 2010):

$$L_p = \left(\sum_t |S_t|^p \right)^{1/p}. \quad (4.5)$$

Here, the value of p can be any integer. The value of p is set as 2 for calculation of the L_2 norm. Mohajer *et al.* used the L_2 norm of dynamic-contrast-enhanced MRI time-series data to cluster breast tumour regions (Mohajer et al., 2010). Compared to healthy tissues, signals from lesions are wider, with greater intensity drops and slower recovery towards the baseline. So, for a lesion voxel, the signal intensity drops to values smaller than those of GM and WM voxels. Consequently, the square root of the summation of their squares (i.e. L_2 norm) is also lower.

Furthermore, signals from different regions should have different L_2 norms due to their unique intensity drop and recovery, which can assist in distinguishing them from one another. While EDP only takes the maximum signal drop into account, this feature reflects the shape of the entire time-series and can reveal additional information about the signal drop and recovery pattern.

5. **Mean time duration (MTD):** The mean time duration for the drop of a signal is referred to as MTD, which is quantified as the time difference between the bolus arrival time (BAT) and the end of the first pass, T_{end} , in this study with the formula:

$$MTD = T_{end} - BAT. \quad (4.6)$$

T_{end} was taken as the first time point after the minimum signal intensity time point where the signal was within one standard deviation of the post-bolus signal (Patil and Johnson, 2011). The post-bolus signal was calculated by averaging the signal intensity values of last ten time points (Yin et al., 2015).

Signals from lesions are wider, with high MTDs, and those from AVs are expected to have lower width and MTD. Therefore, MTD can distinguish lesion voxels and AVs from other tissues.

6. **Third principal component (PC3):** PCA captures the variation of the raw data in a dimensionality-reduced space. Each PC is the projection of the data onto a direction that captures a certain amount of the variance of the original data (Akbari et al., 2014). That is, the first principal component (PC1) is the projection of the raw STCs onto a direction that captures the highest raw data variance, the second component (PC2) is the projection towards the direction of second highest variance and so on.

The first four PCs of DSC-MRI are associated with four unique aspects of the shape of STCs (Akbari et al., 2014). PC1 conveys information about the global baseline, and PC2 reflects the drop of the signal with respect to the baseline. EDP was preferred over PC1 and PC2 as it conveys somewhat similar information to the combination of PC1 and PC2 but provides more explicit insight about signal variation with one less dimension. On the other hand, the third component, PC3, relates to the steepness of both the drop and the recovery of signals. As none of the present features captured the rate of fall or recovery, PC3 was included as a potential unique feature.

7. **Percentage signal recovery (PSR):** PSR is defined as the percentage of signal intensity recovered at the end of the bolus first pass, relative to the baseline value. PSR is given by the following equation:

$$\text{PSR} = \frac{S_{\text{end}} - \text{MSD}}{S_{\text{pre}} - \text{MSD}} \times 100\% . \quad (4.7)$$

Here, S_{end} is the signal value at time point T_{end} . PSR is known to be dependent on the rate of blood flow (Mangla et al., 2011). Lesions in glioblastoma multiform show significantly higher PSR than metastatic lesions (Cha et al., 2007), whereas in lymphoma, PSR is more than 100% as the signal rises over the baseline. Hence, PSR gives useful information for characterising lesions. Although the purpose of the present study was not to distinguish between lesions, PSR was included to investigate its applicability in segmentation.

8. **Fourth principal component (PC4):** The fourth principal component (PC4) is controlled by the baseline and its recovery pattern (Akbari et al., 2014). PSR, mentioned above, also quantifies similar signal characteristics. However, the variation in PC4 is dictated not only by the recovery pattern, but also by the intensity drop (Akbari et al., 2014); therefore, PC4 is expected to convey somewhat different information to PSR and was included in the work as another potential unique feature.
9. **S_1 to S_6 ratio (S_1/S_6):** S_1/S_6 is the ratio of the signal intensity at the first time point (S_1) to that of the sixth time point (S_6). Kao et al. and Akbari et al. reported that in the first few time points of DSC-MRI, CSF pixels show a higher signal intensity (Kao et al., 2010, Akbari et al., 2014). The ratio of S_1/S_{pre} to S_6/S_{pre} —equal to S_1/S_6 —was found to provide higher value for CSF and oedema than other regions. This ratio can therefore be used to separate CSF pixels from the others. The present work included this feature to investigate its potential in distinguishing other regions along with CSF.

Figure 4.2 provides a pictorial representation of the signal parameters needed to calculate the above-mentioned nine features.

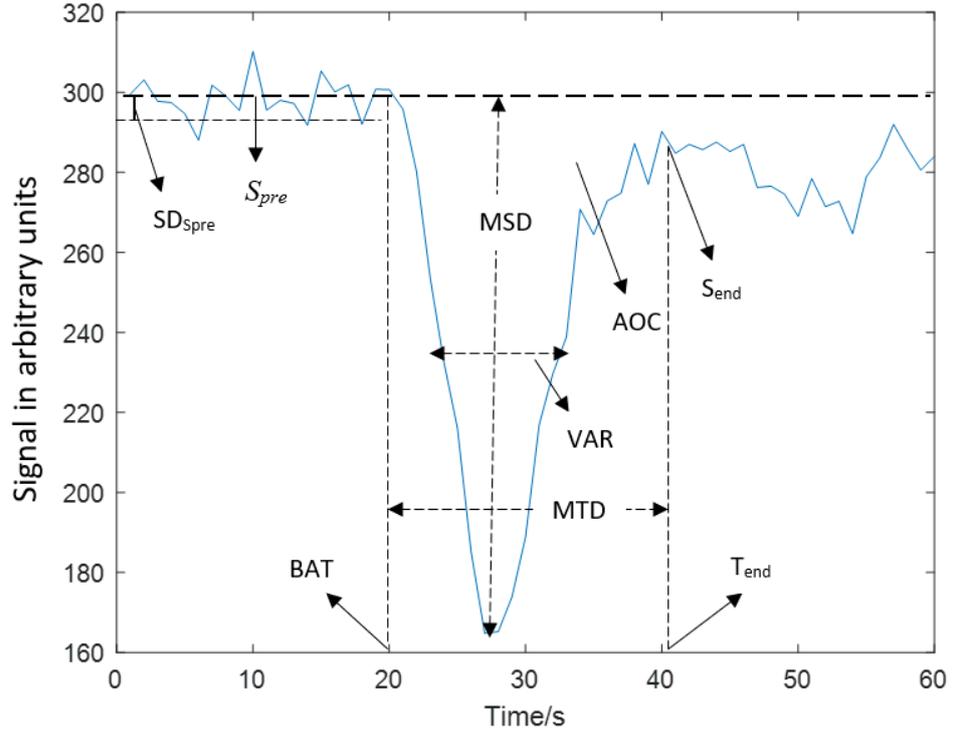


Figure 4.2: A sample DSC-MRI signal-time course showing the parameters necessary to calculate nine features. PC3 and PC4 are not included in the figure as they cannot be explicitly shown. Area over the curve (AOC) and L_2 norm were measured from the normalised signal, where the latter uses the signal drop at every time point and is therefore not included in the figure to maintain clarity. Abbreviations: S_{pre} , baseline signal; SD_{Spre} , standard deviation of S_{pre} ; BAT, bolus arrival time; MSD, maximum signal drop; VAR, Variance; T_{end} , end time point of first pass; MTD, mean time duration; S_{end} , signal intensity at end of bolus first pass.

4.4.4. Feature space creation

In DSC-MRI, each of all N brain voxels is characterised by a dynamic signal with 60 time points. This creates a high-dimensional signal space, S , of dimension $N \times 60$. To reduce the time-complexity and increase the robustness of any computation on the S space, its dimensionality-reduced version is created in this section.

As explained before, with $N = 256$, the R space can be regarded as a small-scale version of the S space. A feature space, F , was created by replacing each element of the R space (i.e. dynamic signal) by a feature vector that contains nine features described in Section 4.4.3. Hence, the R space was converted to a dimensionality-reduced F space of dimension 256×9 . The underlying assumption behind such dimension reduction is that

functionally similar tissue regions have similar perfusion dynamics and therefore similar signal features. So, features from different regions will be distributed in distinct clusters over the multi-dimensional F space, and any standard clustering algorithm can achieve the desired partition between these clusters for the purpose of segmentation.

All nine features that constitute the dimensionality-reduced F space may or may not prove effective for segmentation. A preliminary study was performed to identify the most suitable feature-set, that is, the features with the highest powers for discriminating different regions (further discussed in Section 4.4.5). By finding the optimal number of features, M , the dimension of the F space was further reduced from 256×9 to $256 \times M$.

4.4.5. Effective feature space (Optimisation of M)

The features that convey the most relevant information may vary across applications (Wang et al., 2006a). Before applying the feature-based segmentation approaches, it is therefore necessary to evaluate the individual power of each feature in discriminating different regions and thereby answer the first research question for this chapter (namely, which signal features are most effective for the segmentation task?).

With the nine extracted features, there are $2^9 - 1 = 511$ possible combinations of features for the F space; analysis of the effectiveness of each combination in distinguishing regions (i.e. the discriminatory power, DP) is therefore impractical. Instead, the optimal number of features, M , was determined through the two investigations as follows:

- **Distribution of features:** The first investigation was performed to find how well-clustered the features were. For every well-clustered feature of the F space, the feature values from the same regions should be grouped together and those from different regions should be well-separated. All the feature values were brought to the same scale (0.1 to 0.9) by min-max normalisation and their distributions for different regions were visually assessed. The features for which different tissue regions yielded non-overlapping clusters were regarded as well-clustered. This somewhat qualitative initial investigation was supported by another more-quantitative analysis.

- **Individual discriminatory power of features:** The individual power of each well-clustered feature in discriminating tissue regions was evaluated by generating receiver operating characteristic (ROC) curves. True tissue labels from the \mathbf{R} space were used as ground truth information for calculation of the ROC curve. The area under the ROC curve (AUC_{ROC}) quantified the DP of each feature. The value of AUC_{ROC} ranges from 0 to 1; the nearer the AUC_{ROC} to 1, the higher the DP of a feature was (Bland, 1995).

To generate an ROC curve, it is also necessary to know whether a larger or smaller value of a feature produces a positive segmentation result for a region. For the segmentation of lesions, larger EDP, AOC, variance, and MTD, and smaller L_2 norm should provide positive results. For WM segmentation, smaller EDP, AOC, variance, and MTD, and larger L_2 norm should provide positive results.

4.4.6. Clustering

After determining the optimal dimension of the \mathbf{F} space, it was clustered using unsupervised learning methods, such as standard k -means and k -medoids clustering. Prior to clustering, each feature of the \mathbf{F} space was given a uniform scale ranging from 0.1 to 0.9 through min-max normalisation (Mohamad and Usman, 2013). As discussed in Chapter 3 (Section 3.4.4.3), this prevented any of the dimensions of the \mathbf{F} space from outweighing others, which would have yielded inaccurate cluster outcomes (Mohamad and Usman, 2013). Normalising the \mathbf{R} space eliminated the bias due to PVE or saturation (Mouridsen et al., 2006a). However, normalising can lead to the amplification of noise in low amplitude CTCs (Wismuller et al., 2006). The present study avoided this problem by working with STCs instead of CTCs.

After normalisation, the clustering algorithm was applied on the \mathbf{F} space (for feature-based clustering) and the \mathbf{R} space (for raw-data-based clustering) (Wang et al., 2006a, Guijo-Rubio et al., 2020). During the clustering stage, computational simplicity was prioritised. Although the standard k -means algorithm is simple to implement, scalable for larger datasets, and adaptable to different data types, it has several limitations (as described in Section 3.6.5 of Chapter 3). These include high sensitivity to outliers, low sensitivity to rotation and translation of the dataset, low accuracy for datasets requiring

non-spherical (e.g. elliptical) clusters, and low adaptability to datasets needing different cluster densities (Raykov et al., 2016). Some of these limitations are overcome in this research by improving the centroid selection and the cluster assignment step (steps of k -means clustering is discussed in Section 3.4.4.3).

The centroid selection step was improved by replacing k -means clustering with k -medoids clustering algorithm, whose centroid selection step makes it more robust against outliers (Kaufman and Rousseeuw, 2005). An additional benefit of the k -medoids clustering is its lower time-complexity than that of k -means (Velmurugan, 2010). This k -medoids clustering algorithm was applied on both the F and R spaces separately, which created two segmentation approaches: ‘kmed-feature’ (proposed in this work) and ‘kmed-raw’, respectively.

The cluster assignment step was improved by adding a probability-based cluster assignment method to the distance measurement step of k -means clustering. This was implemented by Hadjiprocopis et al., originally for DWI segmentation (Hadjiprocopis et al., 2005). Again, this modified k -means clustering (discussed in Section 4.4.7.4) was applied on the F and R spaces, creating two more segmentation approaches: ‘POA-feature’ (proposed in this work) and ‘POA-raw’, respectively, where POA stands for probability of assignment.

The performance of above-mentioned feature- and raw-data-based segmentation were compared to answer the second research question for this chapter (namely, can feature-based time-series clustering approaches provide better segmentation than their raw-data-based variants?). In addition to these four approaches, a PCA-based segmentation approach—motivated from the work of Kaya et al. (2017)—was included in this study, where PCA was performed prior to k -medoids clustering (discussed in Section 4.4.7.3). This segmentation approach was included to address the third research question for this chapter: namely, how the performance of proposed and currently-available dimension-reduction-based approaches compare. These five segmentation approaches are discussed in detail in the following sections.

4.4.7. Segmentation approaches

4.4.7.1. Feature-based k -medoids clustering (k med-feature):

The k med-feature approach is proposed in this work as an extension of the work of Bjornerud and Emblem (2010) (see Section 4.2.3). In contrast to segmenting the CTCs using a single feature—AUC—the present study used multiple STC features to classify tissue regions.

The k med-feature approach performs k -medoids clustering with a Euclidean distance measure on the normalised F space. Partitioning around medoid (PAM) is the most common algorithm for finding suitable partition with k -medoids clustering. It is suitable for data with fewer than 3,000 elements and implemented here as the F and R space had 256 elements.

The steps of PAM are given as below (Kaufman and Rousseeuw, 2005):

1. **Initialization step:** Randomly select k number of medoids from the 256 data points. In the present work, k was set as 5 to segment the R space into five regions: GM, WM, AIF, lesion, and ‘other’.
2. **Assignment step (BUILD step):** Assign each data point, F_p , of the F space to the medoid w_j with shortest Euclidean distance where $j = \{1, 2, \dots, k\}$.
3. **Medoid Updating step (SWAP step):** For each cluster, take w_j as the centroid and calculate the total cost of configuration:

$$C_F = \sum_{w_j} \sum_{F_p \in w_j} \left(\sqrt{\sum_{i=1}^n (w_{j,i} - F_{p,i})^2} \right)^2, \quad (4.8)$$

where, n is the number of features in the F space. Then, re-calculate the C_F with F_p as the medoid instead of w_j . If this new C_F is lower than before, take F_p as the new medoid for the cluster. This step essentially computes the average *dissimilarity* of F_p to all other data points in the same cluster with the same medoid, w_j .

4. **Repetition Step:** Repeat steps 2-3 until no change of the assignments (or total cost of configuration) occurred or the maximum number of iterations of 1,000 is reached.

The k -medoids clustering yielded five centroids and assigned a label to each pixel of the \mathbf{R} space. The classification of centroids is described as follows. A label-representative STC was created by aligning and averaging all the signals with the same label. These label-representative STCs were then inspected visually to classify them as GM, WM, AIF, and lesion STCs. The assigned class for a tissue region was the same as that of its label-representative STC. All the segmentation approaches described below followed this manual classification technique.

4.4.7.2. K -medoids clustering on raw data (*kmed-raw*):

The ‘kmed-raw’ segmentation approach was built in line with methods proposed in several published studies for automatically detecting the AIF (Peruzzo et al., 2011, Mouridsen et al., 2006a, Yin et al., 2015, Yin et al., 2014, Murase et al., 2001a). The only difference between the above-mentioned ‘kmed-feature’ and this ‘kmed-raw’ approach is the data space on which the clustering algorithm works. In the kmed-feature approach, the k -medoids clustering worked on the \mathbf{F} space, whereas, in kmed-raw, the \mathbf{R} space was clustered. Therefore, the medoid updating step calculates the total cost of configuration by the following formula:

$$C_R = \sum_{w_j} \sum_{R_p \in w_j} \left(\sqrt{\sum_{t=1}^T (w_{j,t} - R_{p,t})^2} \right)^2 \quad (4.9)$$

Where, R_p is a non-medoid point in the \mathbf{R} space and T is the total number of time points.

4.4.7.3. K -medoids clustering on principal-component-created feature space (*kmed-PCA*):

The third segmentation approach involved PCA and k -medoids clustering to segment the brain regions. As mentioned before, the approach was motivated from the work of Kaya et al. (2017). Although Kaya et al. reported PPCA or EM-PCA as the most effective

dimension-reduction tools (see Section 4.2.6 in review), conventional PCA (explained briefly in Section 4.4.2) was used in this study for simplicity.

PCA converted the \mathbf{R} space to a dimensionality-reduced \mathbf{P} space that is different to the presently-proposed \mathbf{F} space. The PCs of the \mathbf{R} space were obtained as a *score* matrix, where the rows were the observations of the \mathbf{R} space (i.e. signals of the \mathbf{R} space) and the columns were the PCs of each observation, arranged in descending order. Akbari *et al.* suggested that the first six PCs capture more than 99% of the variance in the STC (Akbari *et al.*, 2014); therefore, the \mathbf{P} space was created from the first six columns of the *score* matrix.

The approach then applied k -medoids clustering on the normalised \mathbf{P} space with a Euclidean distance measure. Essentially, this kmed-PCA and the kmed-feature differed only in the data space on which the clustering was applied. For the kmed-feature (in Section 4.4.7.1), the clustering was applied on the \mathbf{F} space, whereas for the kmed-PCA, the clustering was applied on the \mathbf{P} space. The medoid updating step used the following formula to calculate the total cost of configuration:

$$C_p = \sum_{w_j} \sum_{P_p \in w_j} \left(\sqrt{\sum_{c=1}^6 (w_{j,c} - P_{p,c})^2} \right)^2, \quad (4.10)$$

where c is the number of principal components and P_p is a non-medoid datapoint of the \mathbf{P} space.

4.4.7.4. Feature based k -means clustering with POA measure (POA-feature):

The POA-feature approach performed a modified k -means clustering on the \mathbf{F} space. The algorithm, originally proposed by Hadjiprocopis *et al.* (2005), is described as follows:

1. **Initialization Step:** Select k random cluster centroids in the normalised \mathbf{F} space. As before, the number of centroids was set as five to cluster the brain phantom into five regions.
2. **Distance calculation step:** Calculate the distance between a point, F_p , to centroids w_j with the following formula:

$$d(p, j) = \sqrt{\sum_{i=1}^n (w_{j,i} - F_{p,i})^2}, \text{ where } j = \{1, 2, \dots, k\}. \quad (4.11)$$

where n is the number of features of the F space.

3. **Probability of assignment calculation:** Calculate the probability of a point, F_p , belonging to each of the cluster centroids by the formula:

$$P_{p,j} = \frac{e^{-d(p,j)}}{\sum_{j=1}^k e^{-d(p,j)}}. \quad (4.12)$$

The nearer an F_p is to a certain cluster, the higher the value of the numerator and the higher the probability, $P_{p,j}$, that F_p is assigned to the centroid of that cluster.

4. **Cluster assignment step:** Assign each F_p to the cluster for which its $P_{p,j}$ is highest.
5. **Centroid updating step:** The *average* position of all the F_p values in a cluster is chosen as the new cluster centroid. The cluster centroids therefore act as the center of gravity of specific cluster spaces created by their members.
6. **Repetition step:** Repeat steps 2-5 until no more relocation of F_p is necessary or the maximum number of iterations of 1,000 is reached.

4.4.7.5. *K-means clustering on raw-data with POA measure (POA-raw):*

This segmentation approach, named as ‘POA-raw’, clustered raw time-series data (i.e. the R space) with k -means clustering with POA measure. This segmentation approach was used in the work of Hadjiprocopis *et al.* to segment DWI (Hadjiprocopis *et al.*, 2005). POA-raw and POA-feature differed only in the data space on which the clustering was applied. For POA-raw, the clustering was applied on the R space, whereas for POA-feature the clustering was applied on the F space. Therefore, the distance measurement of the clustering method (Step 2 of Section 4.4.7.4) is modified into the following formula:

$$d(p, j) = \sqrt{\sum_{t=1}^T (w_{j,t} - R_{p,t})^2}, \quad (4.13)$$

where R_p is a point in the \mathbf{R} space.

4.4.8. Evaluation of segmentation approaches

Finally, to assess the effectiveness of each segmentation approach, the label assigned by each method was compared to the ground truth label for every pixel of the \mathbf{R} space. The performance of segmentation approaches was reported in terms of their accuracy, precision, sensitivity, and specificity calculated by the following equations:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} , \quad (4.14)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} , \quad (4.15)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} , \quad (4.16)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} , \quad (4.17)$$

Here, TP = number of true positives, i.e. instances when a segmentation approach *correctly* identifies the *positive* tissue class; FP = number of false positives, or instances when the segmentation approach *incorrectly* identifies the *positive* tissue class; TN = number of true negatives, or instances when the segmentation approach *correctly* identifies the *negative* tissue class; and FN = number of false negatives, or instances when the segmentation approach *incorrectly* identifies the *negative* tissue class.

The Dice similarity index or Dice coefficient (DC) was calculated to evaluate the agreement between the ground truth and an assigned label. DC was calculated with the following equation:

$$\text{DC} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} . \quad (4.18)$$

DC ranges from 0 to 1, where 0 indicates a total disagreement and 1 indicates total agreement between an assigned label and the ground truth. A satisfactory segmentation approach should show a $DC > 0.70$ (Zou et al., 2004).

Each segmentation approach was repeated ten times. During each repetition, the clustering step was iterated 1,000 times with a randomly selected initial centroid. These random initialisations resulted in different cluster outcomes and corresponding deviations in DCs. The mean and standard deviation (SD) of the above-mentioned performance parameters were reported. The mean and SD of time duration of the *clustering stages* were measured for all five approaches, for both conventional and noisy R space.

4.4.9. Statistical analysis

Statistical analyses were performed on DC as it summarises the performance of any segmentation approach. Hence, the present experimental design has one measurement variable—DC, with ten observations (representing ten repetitions), and one nominal variable—segmentation method, with five categories representing five approaches. Due to this design, a one-way analysis of variance (ANOVA) was used to test whether there were any significant differences between the DCs obtained from the five segmentation approaches (McDonald and Delaware, 2009).

One-way ANOVA had the null hypothesis that there were no differences between the population means of the DCs of different approaches. Ten hypotheses had to be checked to verify this null hypothesis. To reduce the Type-I error of this multiple testing, the p value was calculated using Bonferroni correction, where the significance level of 0.05 was divided by the number of hypothesis tests. Hence, the present work used $p = 0.005$ (Bland, 1995) for each pairwise comparison. The one-way ANOVA analysis were performed in MATLAB (R2018a, Natick, MA).

For the individual DP analysis of the effectiveness study (Section 4.4.7), the statistical package for the social sciences (SPSS) software was used (IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp).

4.5. Results

4.5.1. Effectiveness of features

4.5.1.1. Distribution of features

Figure 4.3 shows the distribution of the nine features for different brain regions. From the figure, a visual inspection indicates which features were well-clustered and could be suitable for feature-based segmentation. EDP, AOC, Variance, and L_2 norm were well-clustered for different tissue regions. MTDs showed marginally overlapping clusters for GM, WM, and CSF regions. All other features showed severely overlapped clusters, and therefore were not included in the subsequent individual effectiveness analysis.

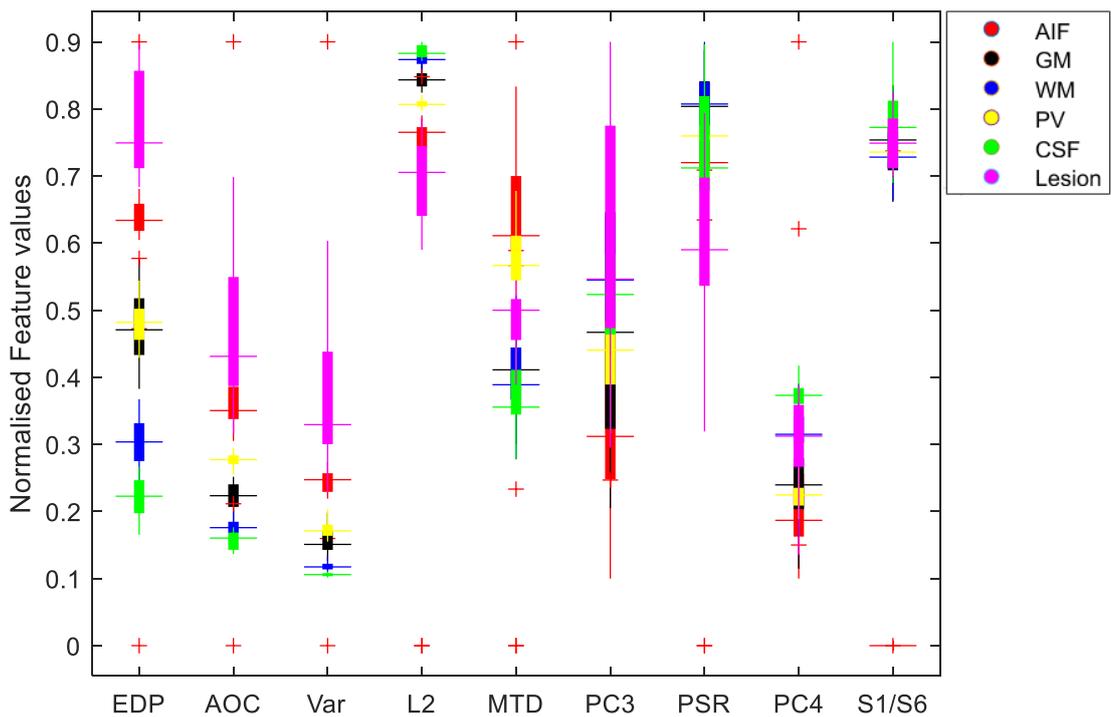


Figure 4.3: Boxplot showing the distribution of normalised feature values for different tissue regions. Colours to represent regions: red, arterial input function (AIF); black, grey matter (GM); blue, white matter (WM); yellow, partial volumes (PV); green, cerebrospinal fluid (CSF); magenta, lesion. For each region: the median of any feature value is represented by the corresponding coloured horizontal line, interquartile range by the corresponding coloured box, 99% confidence bounds by the corresponding coloured vertical lines, and outliers (of each region) by red crosses.

Abbreviations: EDP, effective drop percentage; AOC, area over curve; Var, Variance; L2, L₂ norm; MTD, mean time duration; PC3, 3rd principal component; PSR, percentage signal recovery; PC4, 4th principal component; S_1/S_6 , intensity ratio of first image to sixth image.

4.5.1.2. Individual discriminatory power of features:

Table 4.1 shows the mean AUC_{ROCS}, with their standard errors (SEs), obtained from different well-clustered features for different regions. EDP, AOC, variance, and L₂-norm showed higher effectiveness (AUC_{ROC} > 0.7) for GM, AIF, and lesion segmentation, and therefore were included in the final F space. L₂-norm had an additional advantage of distinguishing WM with higher power than other features. The MTD was included in the F space as its values for different regions were less overlapping than PC3, PSR, PC4, S_1/S_6 and its individual effectiveness for GM and WM segmentation was high (near to 0.7).

Table 4.1: Mean (SE) of AUC_{ROC} of different features for different regions

Brain Region	EDP	AOC	Variance	L ₂ norm	MTD
GM	0.81 (0.03)	0.75 (0.03)	0.80 (0.03)	0.76 (0.03)	0.65 (0.03)
WM	0.59 (0.03)	0.58 (0.03)	0.59 (0.03)	0.92 (0.02)	0.66 (0.03)
AIF	0.92 (0.02)	0.95 (0.01)	0.92 (0.02)	0.97 (0.02)	0.50 (0.03)
Lesion	0.95 (0.05)	0.94 (0.05)	0.94 (0.05)	0.76 (0.05)	0.45 (0.04)

Abbreviations: EDP, effective drop percentage; AOC, area over curve; Var, Variance; L2, L₂ norm; MTD, mean time duration.

This effectiveness study suggested that a feature space containing five features will be suitable for feature-based segmentation. Therefore, both feature-based approaches, kmed-feature, and POA-feature, used this 256×5 -feature space for segmenting the R space.

4.5.2. R space segmentation

4.5.2.1. GM segmentation

Figure 4.4 gives the accuracy, precision, sensitivity, specificity, and DC of the five segmentation approaches in distinguishing the GM region from others. The figure shows that for segmenting GM, the proposed kmed-feature approach provided a significantly higher DC than kmed-raw ($p = 0.0014$) and kmed-PCA ($p < 0.0001$). No statistically-significant difference was found between the DC obtained from POA-feature to those from other segmentation approaches. Moreover, POA-feature showed a higher variance in DC than any other approaches, indicating its lack of robustness and unsuitability for GM segmentation.

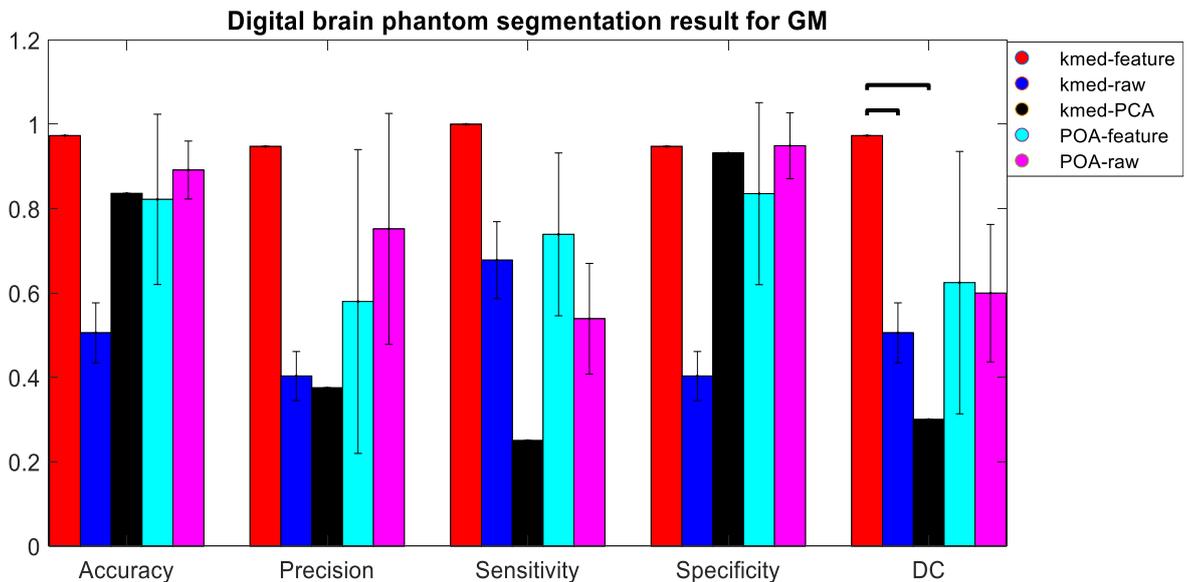


Figure 4.4: Bar chart showing the individual performance parameters for GM segmentation. The error bars show the SDs above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-raw); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-raw). Statistical significance ($p < 0.005$, due to Bonferroni correction) between two segmentation approaches is shown with a line joining the pairs of bar charts. Significance is only shown in the Dice coefficient (DC) to summarise the agreement between the ground truth and segmentation results.

4.5.2.2. WM segmentation

Figure 4.5 gives the accuracy, precision, sensitivity, specificity, and DC of the five segmentation approaches in distinguishing the WM region. The figure shows that for segmenting WM, kmed-feature and kmed-raw performed significantly better than kmed-PCA ($p \leq 0.001$), while no other approaches were significantly different. Clustering the F space produced DCs similar to those produced by clustering the R space, but better than those obtained by clustering the P space.

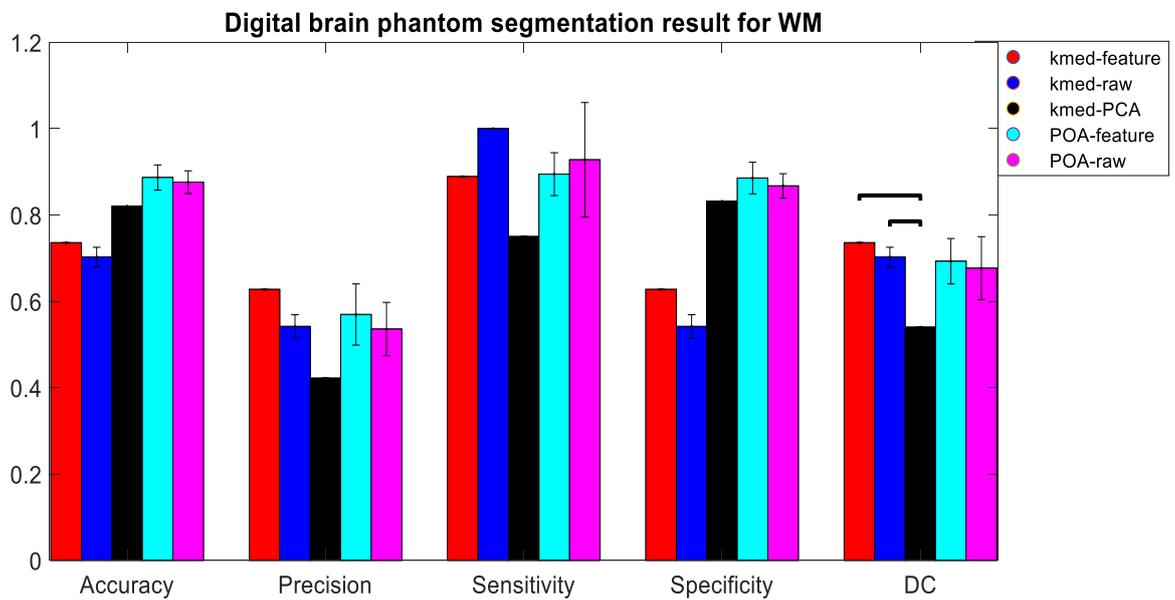


Figure 4.5: Bar chart showing the individual performance parameters for WM segmentation. The error bars show the SDs above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-raw); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-raw). Statistical significance ($p < 0.005$, due to Bonferroni correction) between two segmentation approaches is shown with a line joining the pairs of bar charts. Significance is only shown in the Dice coefficient (DC) to summarise the agreement between the ground truth and segmentation results.

4.5.2.3. AIF segmentation

Figure 4.6 presents the accuracy, precision, sensitivity, specificity, and DC of the five segmentation approaches in distinguishing the arterial region from others. The kmed-feature approach gives significantly higher DC than the POA-raw and kmed-PCA approaches ($p < 0.005$). Out of the proposed approaches, POA-feature showed a high variance of DC, indicating its unsuitability for AIF segmentation.

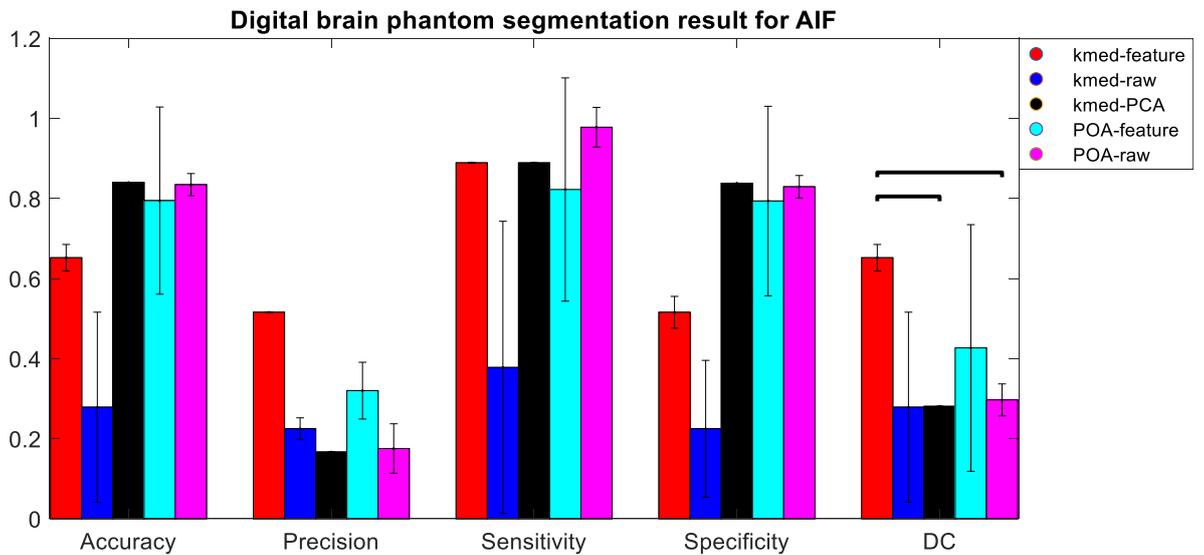


Figure 4.6: Bar chart showing the individual performance parameters for AIF segmentation. The error bars show the SDs above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-raw); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-raw). Statistical significance ($p < 0.005$, due to Bonferroni correction) between two segmentation approaches is shown with a line joining the pairs of bar charts. Significance is only shown in the Dice coefficient (DC) to summarise the agreement between the ground truth and segmentation results.

4.5.2.4. Lesion segmentation

Figure 4.7 gives the accuracy, precision, sensitivity, specificity, and DC of the five segmentation approaches in distinguishing the lesion region. It can be seen that the POA-raw approach yielded significantly higher DCs than the proposed kmed-feature, kmed-PCA, and kmed-raw ($p < 0.0001$). However, the other raw data-based segmentation

approach, kmed-raw, did not give any significant improvement in DC. This suggests that the raw-data-based approaches only segment lesions better when the cluster assignment measure of typical k -means clustering is improved. Like GM and AIF regions, the POA-feature approach showed high variability of DC, indicating its unsuitability.

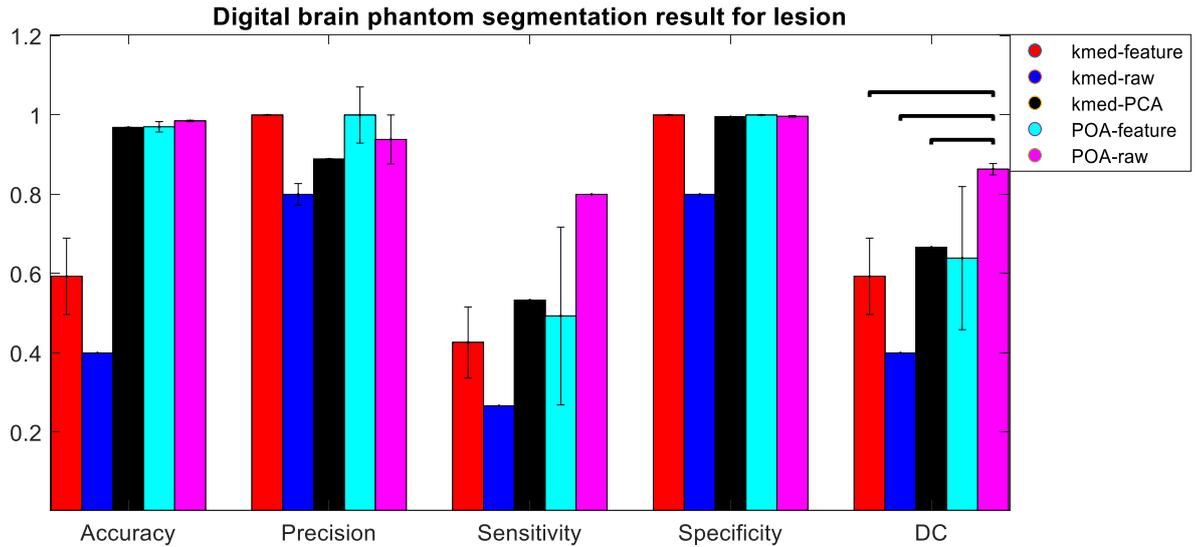


Figure 4.7: Bar chart showing the individual performance parameters for lesion segmentation. The error bars show the SDs above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-raw); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-raw). Statistical significance ($p < 0.005$, due to Bonferroni correction) between two segmentation approaches is shown with a line joining the pairs of bar charts. Significance is only shown in the Dice coefficient (DC) to summarise the agreement between the ground truth and segmentation results.

When all the nine features were included in the F space (see Figure 4.8), kmed-feature showed significantly better performance than kmed-PCA; its performance with a 256×9 -dimensional F space was also comparable to raw-data-based approaches. This indicates that a different F space to the one used for GM and WM segmentation can provide better lesion segmentation.

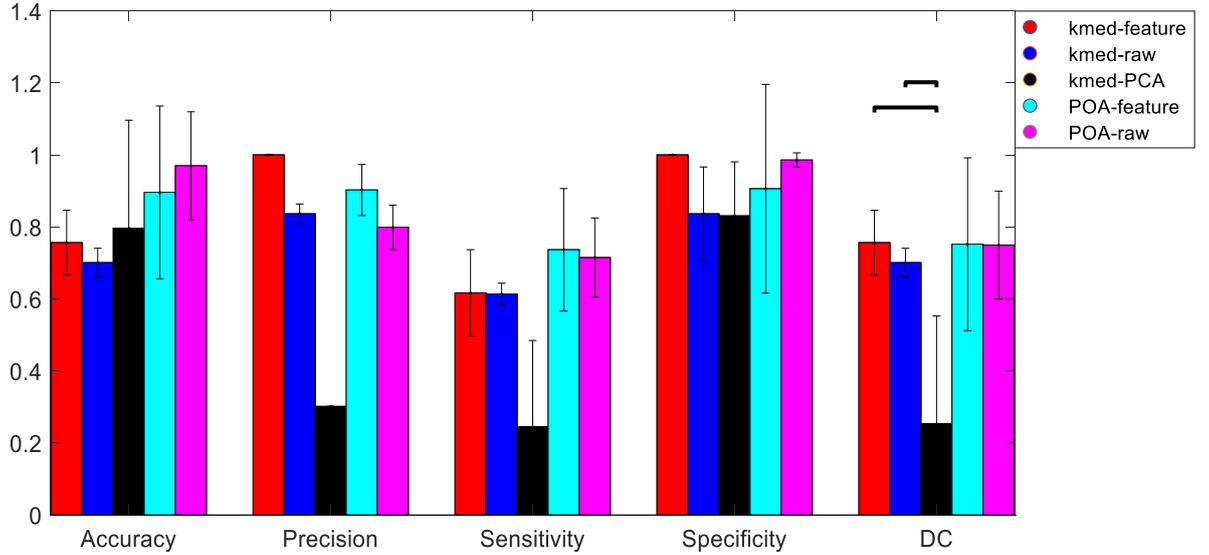
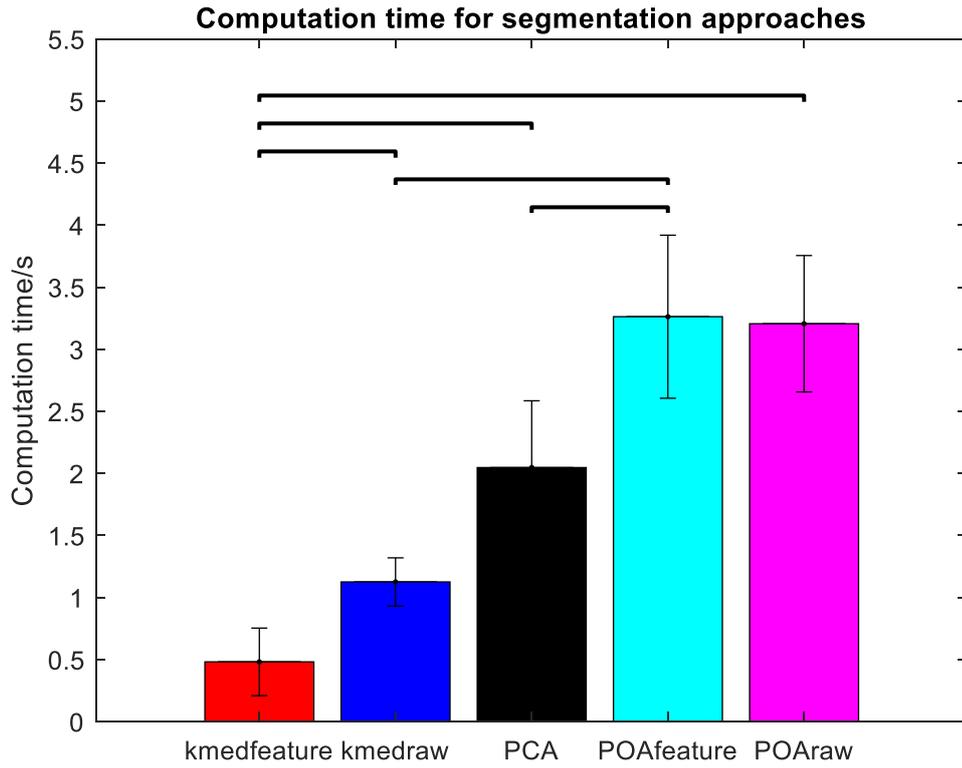


Figure 4.8: Bar chart showing the individual performance parameters for lesion segmentation with a feature-space of dimension 256×9 . The error bars show the SDs above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-raw); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-raw). Statistical significance ($p < 0.005$, due to Bonferroni correction) between two segmentation approaches is shown with a line joining the pairs of bar charts. Significance is only shown in the Dice coefficient (DC) to summarise the agreement between the ground truth and segmentation results.

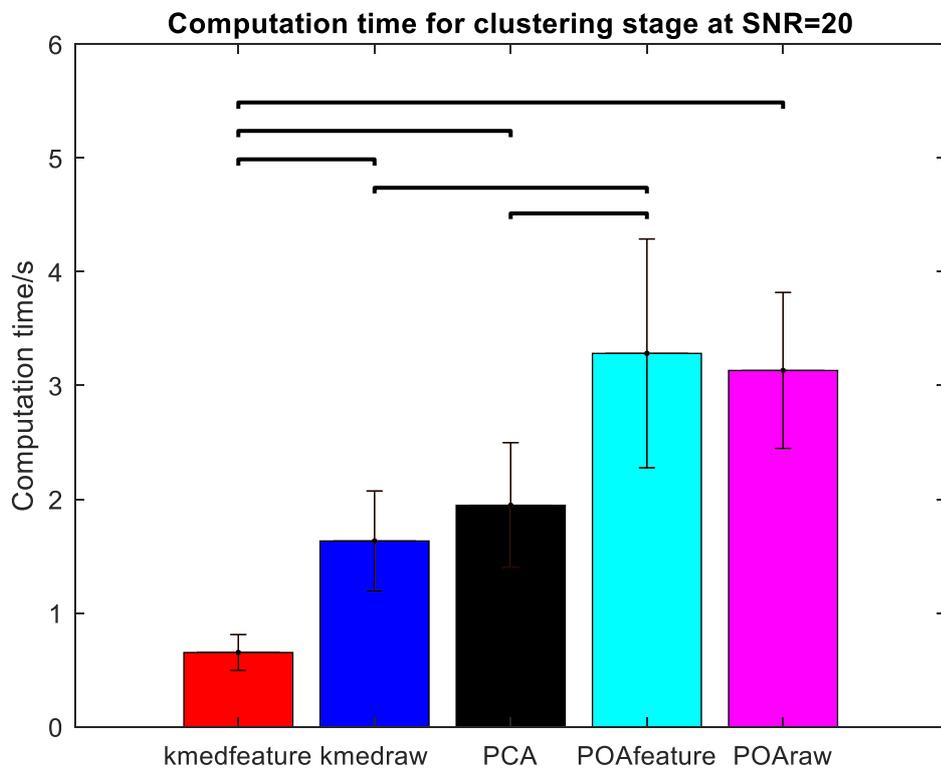
4.5.3. Computation time

Figure 4.9 compares the computation time of clustering stage of the five segmentation approaches for (a) signals with noise typical to clinical setting and (b) signals with SNR = 20.

For the R space signals with noise typical to clinical cases (Figure 4.9a), the feature-based clustering of kmed-feature approach gave significantly shorter computational time ($p \leq 0.001$) than those of any other segmentation approaches. However, feature-based clustering of POA-feature approach showed significantly-higher computation time than that of kmed-raw and kmed-PCA ($p < 0.001$), which can be attributed to the more-convoluted centroid assignment measure of the clustering. For lower SNRs (Figure 4.9b), the kmed-feature clustering was still significantly quicker than others ($p < 0.01$); whereas POA-feature had the slowest clustering method.



(a)



(b)

Figure 4.9: Bar chart showing the mean time elapsed during the clustering stage when five segmentation approaches are applied on (a) \mathbf{R} space with no additional noise (b) \mathbf{R} space with

additional noise (SNR=20). The error bars show the SDs above and below the mean values. Bar Colours: red, k -medoids clustering on the F space (kmed-feature); blue, k -medoids clustering on the R space (kmed-raw); black, k -medoids on P space (kmed-PCA); cyan, k -means clustering with POA measure on F space (POA-feature); magenta, k -means clustering with POA measure on R space (POA-raw). Statistical significance ($p < 0.005$) between two segmentation approaches is shown with a line joining the pairs of bar charts.

It should be noted that Figure 4.9 gives the computation time of the clustering stage only, not the overall segmentation time. The overall segmentation times can vary as the approaches involved manual classification of cluster centroids. However, if the duration of the classification stage can be kept similar, the kmed-feature approach would segment faster than all other approaches.

4.6. Discussion

The primary aim of the study was to investigate the feasibility of feature-based time-series clustering in tissue segmentation. To create a dimensionality-reduced feature space, the raw time-series of each pixel was replaced by a vector containing several features. A preliminary effectiveness study compared the individual discriminatory powers of these features and identified the most effective features to include in the final feature space. Two proposed segmentation approaches clustered this feature space with different clustering mechanisms before classifying the brain regions. The performances of these proposed approaches were compared to those of two segmentation approaches that clustered raw dynamic data and to another approach that used PCA for dimension reduction of DSC-MRI data prior to clustering.

The results showed that feature-based clustering provides simple, fast, but effective segmentation approaches. The dimensionality-reduced feature space facilitates better segmentation as compared to PCA-based dimension-reduction methods. Detailed discussion on the findings, potential clinical applications, limitations, as well as the future scope of the present work will be presented in the following sections.

4.6.1. Effectiveness study of the features

The effectiveness study identified the five features with higher discriminatory power than others. It also suggested that, although features like PSR and S_1/S_6 can be suitable for discriminating specific regions such as lesions or CSF respectively, they cannot be regarded as effective features for segmenting all the regions. The other two features—PC3 and PC4—were not well-clustered and were already used in the kmed-PCA approach; therefore, including them in the F space would not add any further discriminatory power. MTD individually showed a DC close to 0.7 for GM and WM segmentation, but failed to distinguish AIF and lesion regions. This may have contributed to the low DCs for AIF and lesion segmentation (Figure 4.6 and 4.7), as compared to those obtained for GM and WM segmentation for the proposed kmed-feature. However, considering its GM- and WM-discriminatory power, MTD was included in the F space.

The inter-subject variation affected the distribution of some features. For example, signals from CSF and WM showed lower variance and MTD in Figure 4.3, but this should be the case for AIF pixels. The reason for this discrepancy is that the signals were normalised to account for any inter-subject variation in their intensity (y-axis), but no normalisation was done to account for differences in transit time (i.e. x-axis). This suggests that only visual inspection of the distribution of features (from Figure 4.3) may not confirm the effectiveness of each feature; therefore, it was followed by another more-quantitative individual DP analysis (Section 4.4.5).

4.6.2. Comparison of segmentation approaches

The analysis of the DCs of the five segmentation approaches (Section 4.5.2) showed that the proposed kmed-feature segments GM significantly better than the two raw-data-based approaches. For WM segmentation, it performs similar to the raw-data-based methods. It also performs better GM, WM, and AIF segmentation than the PCA-based segmentation approach. The kmed-feature, with a 256×9 -dimensional F space, distinguishes the lesion regions equally well to other raw-data-based approaches, with DCs comparable to those obtained by the 256×5 -dimensional F space for GM and WM

segmentation. The DCs obtained with another proposed approach, POA-feature, show high variance, probably due to the sensitivity of its k -means clustering to outliers and also to random initialisation step that converges the clustering into one of the local minima of the objective function (Raykov et al., 2016). The computation times of POA-feature were longer as well, due to the complex cluster assignment step. Therefore, out of the two proposed feature-based approaches, the present study advocates the kmed-feature for segmentation.

In the \mathbf{R} space, different regions were scattered randomly; whereas for real patients, similar voxels are more in proximity, which will assist the clustering. Moreover, for clinical perfusion analysis, all brain voxels of a subject are often assumed to have an idealised AIF. So, the current assumption of similarly-perfused regions having similar-shape time-series can be sustained with more confidence. However, the clinical DSC-MRI data are of large dimension, due to a vast number of brain voxels and many time-points in their dynamic signals. This problem is addressed here by the proposed dimensionality-reduction, which has the following advantages. First, by decreasing the number of time-points, the dimension of each element is now decreased from 60 to 5. Consequently, the time-complexity of the clustering stage is decreased (Wang et al., 2006a), as described in Section 3.6.1 in Chapter 3. This is demonstrated in Section 4.5.3, where kmed-feature is shown to cluster faster than other clustering approaches. So, it can also segment faster than other raw-data-based or dimensionality-reduction-based approaches; especially when clustering can be followed by a similar-duration, if not entirely-automatic, classification stage instead of the present manual classification of cluster centroids (described in Section 4.4.7.1). Second, using a few signal features instead of the entire dynamic intensity variation allows more information to be stored in a smaller number of dimensions (Wang et al., 2006a). This offers a more effective segmentation approach by preventing it from being overburdened by less informative parts of the dynamic signal, such as a long baseline and recirculation. Nevertheless, in future research, the segmentation performance of the proposed approaches on a noisy \mathbf{R} space needs to be investigated.

The kmed-feature approach produced low DCs for AIF and lesion segmentation (Figure 4.6 and 4.7). This can be due to the lower power of MTD in discriminating the AIF and lesions. When segmenting these two regions with this approach, a higher DC can be achieved if the present \mathbf{F} space contains different or additional features. For example, for

AIF identification, a feature space created with the effective AIF detection features (as suggested by Chapter 3) may improve the DC. On the other hand, for lesion segmentation, when a 256×9 -dimensional feature space was used instead of the 256×5 -dimensional F space, the resultant DC was comparable to kmed-raw and POA-raw and significantly different to kmed-PCA (see Figure 4.8). Hence, for lesion identification, the proposed kmed-feature approach sustains the segmentation performance, while providing the additional advantage of rapid computation.

4.6.3. Clinical applications

The proposed feature-based segmentation approach can be useful for many clinical applications. For example, the kmed-feature approach can facilitate rapid diagnosis of lesions in acute stroke patients, without the necessity to transfer the data into another processing platform or to execute intensive voxel-by-voxel analysis. Clinicians can make quick decisions regarding a suitable treatment plan and thus prevent rapid, progressive damage to myelinated fibres (Calamante, 2013).

Combining different semi-quantitative features (e.g. EDP, AOC, L_2 -norm, MTD) can create novel parameters with high region-discriminating power; such parameters can offer both the researchers and operators in the clinical settings an initial insight for the identification of normal appearing WM, lesion, penumbra, or arterial supply for a certain tissue neighbourhood.

Brain maps created from different semi-quantitative features can allow future studies to investigate their association with glioma grade, progression-free and overall survival, as well as their utility as biomarkers of glioma or other pathologies. Then, feature-based analysis would allow hospitals to circumvent the time-consuming and computationally-intensive voxel-by-voxel perfusion analysis and pave the way to a fast and objective characterisation of lesions, monitoring of angiogenesis, or detection of arterial and venous voxels for PVE-correction.

4.6.4. Limitations

The limitations of the present study are as follows. The applied k -medoids clustering is not suitable for data that require non-spherical clusters. However, an initial visual inspection of the distribution of the features in the F space indicated that the data can be partitioned with spherical clusters. The satisfactory results of the kmed-feature approach also suggest the suitability of k -medoids clustering for the present datasets. Another limitation with regard to all five clustering methods is that their centroids are randomly initialised. Different iterations, therefore, may produce different cluster outcomes. To ensure reproducibility, the clustering step of every segmentation approach was iterated 1,000 times (Yin et al., 2014). Nevertheless, for future research, different clustering algorithms should be explored to achieve either more robust and reproducible outcomes or better trade-offs between the segmentation performance and computational complexity than k -medoids.

The limitation of the applied PAM algorithm (Sections 4.4.7.1, 4.4.7.2, and 4.4.7.3) is that it is only suitable for data with fewer than 3,000 elements. For typical DSC-MRI with a large number of elements (i.e. voxels), an alternative to the PAM algorithm will have to be found. For more than 3,000 but fewer than 10,000 brain voxels, clustering should use a variant of Lloyd's iterations, based on the work of Park and Jun (2009). For a much larger number of brain voxels, another variation of the k -medoids algorithm should be used, where a random sample of cluster members is examined during each iteration.

The kmed-PCA approach (Section 4.4.7.3) used a conventional PCA-based dimension-reduction, whereas there exist more powerful methods like EM-PCA or PPCA (Kaya et al., 2017) (see Section 4.2.6). However, their application would have further increased the complexity and the computation time. Therefore, even though they might have marginally increased the DC, their application would contradict the primary purpose of this study: establishing a simplified and rapid automatic segmentation approach.

Lastly, the feature space is application dependent. Although the extracted features are suitable for segmenting GM, WM, and with some modification for AIF or lesion, they may not be useful for other tasks or processing steps, such as venous voxel detection, or further characterisation or grading of lesions. To find the most suitable task-specific features for any such processing step, some prior knowledge about the task and the

corresponding dataset will be required. Otherwise, researchers are advised to design preliminary investigations similar to the presented work, to finalise the task-specific feature-sets.

4.6.5. Future scopes

For future research, instead of using a common F space for all regions, the kmed-feature approach can be improved by applying target-region-specific F spaces. For example, for arterial voxel detection, the F space can be populated by features that were found to be effective for AIF detection in Chapter 3. For lesion detection, a 256×9 -dimensional F space can be useful as suggested in Section 4.5.2.4. After masking the arterial and lesion voxels, the GM and WM voxels can be segmented with the 256×5 -dimensional F space of the present study.

Future studies can also develop novel features for segmentation by combining different task-specific features of STCs. From different parametric forms of this composite feature, the most effective one can be identified by measuring the area under the ROC curve. Additionally, the threshold for segmenting different regions with this novel combinational feature can be investigated.

Relevant task-specific features extracted from different co-registered perfusion images (such as dynamic contrast enhanced-MRI, arterial spin labelling) can be merged into a single feature space. This will accommodate more information about a voxel into a lower-dimensional F space and increase the orthogonality between voxels, enabling better region discrimination. This can enable clinicians to characterise or grade the pathological regions more conclusively by analysing the F space containing multiple cross-modality features.

4.7. Conclusion

The conclusion of the study is that one of the proposed feature-based segmentations, namely kmed-feature, outperforms two raw-data-based and one conventional dimensionality-reduction-based approaches. When followed by an automatic or

controlled-duration classification stage, the kmed-feature can potentially offer the fastest segmentation, even for noisy data. The research suggests using five features for GM and WM segmentation: effective drop percentage, area over the normalised signal time curve, variance, L_2 norm, and mean time duration. Identification of arterial voxels may be further improved by including the features prescribed in Chapter 3. Lastly, the kmed-feature provides a simple, fast, but effective approach for distinguishing lesions, without transferring the data to other processing platforms, which can be beneficial for clinical cases that demand rapid processing and urgent clinical decision making (e.g. ischaemic stroke).

Chapter 5: Comparing Different Forms of Physiologically Plausible Transit Time Distributions

5.1. Introduction

As mentioned in Chapter 2, the present research solves the central convolution equation of dynamic susceptibility-contrast magnetic resonance imaging (DSC-MRI) (equation 2.17) through model-dependent deconvolution, which involves assuming an analytical form for the tissue impulse response function, $I(t)$. Since $I(t)$ is the product of cerebral blood flow (CBF) and residue function $R(t)$, this deconvolution can be achieved by assuming a physically realistic model for $R(t)$ with several free parameters. This parameterisation of $R(t)$ requires knowledge about the physiological model of tissue microvasculature.

The brain capillary system can be modelled as a combination of many parallel tubes, as shown in Figure 5.1 (a) (Mouridsen et al., 2006b, Ostergaard et al., 1999). Each tube receives a unique fraction of the arterial blood, h_i , that passes through the tube with a unique transit time, τ_i . Each voxel contains several capillary-tissue units with a large number of parallel tubes. Consequently, the discrete h_i 's convert to a continuous density function of transit time, referred to as the transit time distribution (TTD), shown in Figure 5.1 (b). As discussed in Section 2.6.1, the TTD can be used to derive R through the following equation:

$$R(t) = 1 - \int_0^t h(\tau) d\tau, \quad (5.1)$$

As the number of free parameters of R is proportional to the number of tubes, a large number of parallel tubes leads to a large number of free parameters in the model of R . This increases the computational complexity. A feasible model of R requires a compromise between mathematical complexity and the practical limits set by temporal resolution, signal-to-noise ratio (SNR), and tissue coverage (St Lawrence and Lee, 1998). To maintain this trade-off, the parametric model should be as biologically specific as possible with the lowest number of free parameters (O'Connor et al., 2011). The effective number

of model parameters can be reduced by characterising the vasculature with a parametric form of $h(t)$ that contains two or three free parameters to describe its shape. Variation in the free parameters of $h(t)$ can then vary its shape to resemble the different transit times taken by Gadolinium-based contrast agents (GBCAs) while passing through vasculatures of unique structure.

Several published studies have investigated different physiologically plausible models of $h(t)$ for the brain capillary system, which are reviewed in the following section before introducing the research questions.

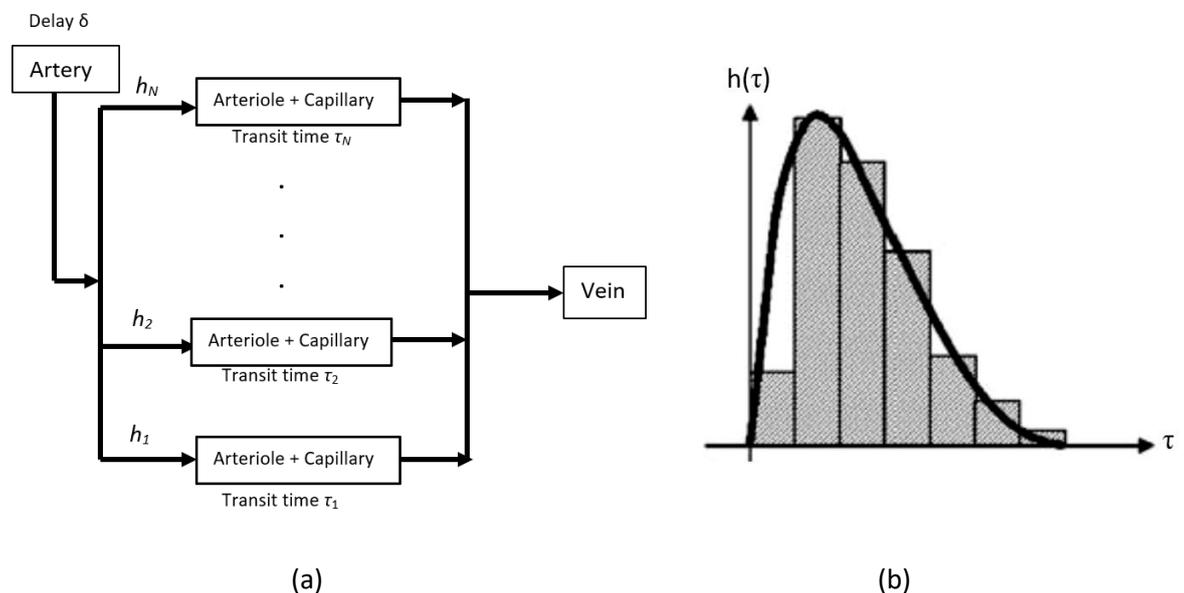


Figure 5.1: (a) Schematic representation of vascular model; arterial input to each tissue is distributed across N tubes. Each fraction of input, h_i , passes with a distinct transit time τ_i , where $i = 1, 2, \dots, N$. (b) Histogram showing discretised approximation to the transit times by different h_i 's (Mouridsen et al., 2006b); when N is a very large number, this histogram converges into a continuous density function, named the transit time distribution (shown by solid curve). Figures reproduced from the work of Mouridsen et al. (2006b), with permission from Elsevier³.

³ Reprinted from NeuroImage; 33(2); Kim Mouridsen, Karl Friston, Niels Hjort, Louise Gyldensted, Leif Østergaard, Stefan Kiebel; Bayesian estimation of cerebral perfusion using a physiological model of microvasculature; 570-579, Copyright (2006), with permission from Elsevier. Copyright license no: 4962571281148.

5.2. Literature review on different forms of TTD

5.2.1. Gaussian TTD

Koh *et al.* modelled h as a Gaussian distribution with the following form:

$$h_{\text{Gaussian}}(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp[-(t-u)^2 / (2\sigma^2)] , \quad (5.2)$$

where μ is the mean and σ is the standard deviation of the TTD (Koh et al., 2001). No constraints were imposed on the transit times and, hence, the TTD was assumed to be completely symmetric. This Gaussian h allows the transit time to be negative, violating the $h(0) = 0$ constraint discussed in Section 2.6.1; therefore, this TTD cannot be a realistic representation of a physical microvasculature.

5.2.2. Corrected Normal TTD

To correct the above-mentioned limitation of the Gaussian h , Koh *et al.* proposed a corrected normal distribution for h , which has the following form:

$$h^{\text{corr}} = \frac{1}{N} h_{\text{Gaussian}}(t) ; t > 0, \quad (5.3)$$

where N is the normalisation factor given by:

$$N = \int_0^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp[-(t-u)^2 / (2\sigma^2)] dt . \quad (5.4)$$

For the values of σ and μ of h^{corr} to be associated with the mean and standard deviation of the TTD, the μ needs to be greater than 3σ . Despite being computationally simpler, this corrected normal TTD starts with a positive value and thus violates the basic TTD constraint of $h(0) = 0$.

To better represent the underlying asymmetry in the physical system and the transit times, a positively skewed TTD is desirable (Koh et al., 2001). This can be achieved if the exponent $-(t-\mu)^2 / (2\sigma^2)$ of equation 5.4 is replaced by a quadratic function $-(At^2 + Bt +$

C). In such case, there would be three free parameters— A , B , and C —instead of two, μ and σ ; the extra parameter would increase the computational complexity.

5.2.3. Skewed-Gaussian TTD

To skew a TTD without adding extra parameters, Koh *et al.* proposed a naturally-asymmetric skewed-Gaussian TTD with the following form:

$$h_{sg}(t) = \frac{2\alpha^{\rho/2}}{\Gamma(\frac{\rho}{2})} t^{\rho-1} \exp[-\alpha t^2]; t, \alpha > 0, \rho > 1, \quad (5.5)$$

where α is the scale parameter and ρ is the shape parameter (Koh et al., 2001). The gamma function of equation 5.5 is defined as follows:

$$\Gamma(z) = \int_0^{\infty} x^{z-1} \exp[-x] dx ; z > 0. \quad (5.6)$$

This skewed-Gaussian TTD follows the constraints: $h(0) = 0$ and $h(\infty) = 0$. Moreover, the resultant R is a smooth and monotonic decreasing function of time. Therefore, this asymmetric curve is a better representation of a realistic capillary system than h^{corr} . As discussed in Section 2.6.1, an exponential $R(t)$ should be avoided in DSC-MRI analysis since it implies instantaneous extravasation of a fraction of the GBCA (Sourbron and Buckley, 2013). For the skewed-Gaussian distribution, an exponential R was avoided by setting the shape parameter $\rho > 1$. A limitation of this distribution is that it cannot be solved analytically, and therefore, needs to be solved numerically (Schabel, 2012).

5.2.4. Gamma TTD

A continuous, closed-form solution for the response function can be obtained by choosing a family of gamma distributions as TTD. Provided certain constraints (as discussed in Section 2.6.1) are imposed, this family of distributions can plausibly describe the tissue vasculature (Mouridsen *et al.*, 2006)(Schabel, 2012). The gamma TTD is given by:

$$h_{\text{gamma}}(t) = \frac{1}{\beta^\alpha} \frac{t^{\alpha-1}}{\Gamma(\alpha)} \exp[-t / \beta] ; \alpha > 1, \beta > 0. \quad (5.7)$$

Here, β is the scale parameter and α is the shape parameter. The gamma function $\Gamma(\alpha)$ is defined by equation 5.6. Due to the reasons explained in Section 2.6.1, an exponential $R(t)$ was avoided by setting α greater than unity, and an infinite value of gamma h was avoided by setting β greater than 0.

The gamma distribution can cover a wide range of shapes of R , such as an exponential for normal tissue and a boxcar for ischaemic tissue. Schabel et al. modelled dynamic contrast-enhanced MRI (DCE-MRI) data using a gamma TTD and showed that the determination of α is dependent on the quality of the datasets, such as high signal-to-noise ratio (SNR), pronounced first pass, and limited contrast extravasation. The present DSC-MRI data meet all these criteria, and therefore the gamma function can potentially represent tissue microvasculature for the current patient cohort.

5.2.5. Gamma-variate TTD

The gamma-variate function has been used in perfusion MRI studies to model the bolus shape function (Patil and Johnson, 2011), to remove tracer recirculation from the GBCA bolus to analyse the first pass (Peruzzo *et al.*, 2011), and to simulate AIF (Bjornerud and Emblem, 2010, Yin et al., 2015, Yin et al., 2014). In a recent work, the gamma-variate distribution was used as a TTD to model DCE-MRI data (Larsson et al., 2017).

The usual form for the gamma variate function is given by:

$$h_{\text{gv}}(t) = At^\alpha \exp[-t / \beta], \quad (5.8)$$

where A , α , and β are free parameters (Madsen, 1992).

The free parameters are coupled in equation 5.8; any change in one parameter affects both the magnitude and shape of the distribution. Madsen proposed a normalised gamma variate function by decoupling the parameters, thereby making it more robust for least-squares fitting (Madsen, 1992). The form suggested by Madsen is as follows:

$$h_{\text{gv}}(t) = y_{\text{max}} (t')^\alpha \exp[\alpha(1-t')], \quad (5.9)$$

where y_{max} is the maximum value of the distribution, α is the decay parameter, and $t' = t/t_{max}$, with t_{max} being the time at which h_{gv} is maximum. As discussed in Section 2.6.1, the integral of h should equal unity. This unit integral constraint yields:

$$y_{max} = \left(\frac{\alpha}{t_{max}}\right)^{\alpha+1} \frac{1}{\Gamma(\alpha+1)} \quad (5.10)$$

where Γ is the gamma function defined by equation 5.6 (Patil et al., 2013, Patil and Johnson, 2013). Substituting y_{max} from equation 5.10 into equation 5.9 and expanding t' gives:

$$h_{gv}(t) = \left(\frac{\alpha}{t_{max}}\right)^{\alpha+1} \frac{1}{\Gamma(\alpha+1)} \left(\frac{t}{t_{max}}\right)^{\alpha} \exp\left[\alpha\left(1 - \frac{t}{t_{max}}\right)\right]. \quad (5.11)$$

Like the skewed-Gaussian and gamma functions above, h_{gv} conforms to all the constraints of section 2.6.1 to be a physically viable representation of tissue microvasculature. Additionally, visual inspections of the time differences between AIF peak and STC minimum can provide a rough estimate for the upper limit of initial guess of t_{max} , making the initialisation stage of the least-squares fitting more intuitive for the gamma-variate. In the above-mentioned study of Larsson et al. (2017), the use of the gamma-variate TTD offered a simple approach for characterising the capillary transit time heterogeneity between healthy controls and patients with carotid stenosis and with brain tumours. This recent success motivated the inclusion of the gamma-variate function in this work's potential models of TTD.

5.3. Aims

It is evident from the discussion above that a variety of functions can be used to model the tissue microvasculature, as long as they conform to the constraints described in Section 2.6.1. However, to date, no study, to the knowledge of the author, has compared the available functions in terms of their computational benefits.

The aim of this chapter is to evaluate the clinical utility of the skewed-Gaussian (equation 5.5), gamma (equation 5.7), and gamma-variate (equation 5.11) functions as TTDs; and

to compare their utility to a novel form of TTD, the Weibull function (discussed in Section 5.4.1.3). The following research questions are addressed in this chapter:

1. Can modelling TTD with a Weibull function provide additional benefits over the available models?
2. Which model gives better goodness and stability of fit for model-dependent deconvolution?
3. How do the perfusion parameters vary with different models of h ?
4. How does the total computation time vary with different h ?

5.4. Materials and Methods

For each TTD, the perfusion parameters, namely cerebral blood flow (CBF), cerebral blood volume (CBV), and mean transit time (MTT), were quantified through the steps described in this section.

5.4.1. Data acquisition and analysis

5.4.1.1. Data acquisition

The data were derived from DSC-MRI scans of nine glioma patients as described in Section 1.4. STCs were obtained from 3×3 -pixel regions of interest (ROIs) placed manually in four areas of normal-appearing white matter (WM) in the frontal and parietal lobes and in two areas of grey matter (GM) in the caudate nucleus, as marked in Figure 5.2(a). Typical average signals from the marked GM and WM ROIs are shown in Figure 5.2(b). All processing and analysis of the acquired signals was performed offline with code written in MATLAB (R2018a, Natick, MA).

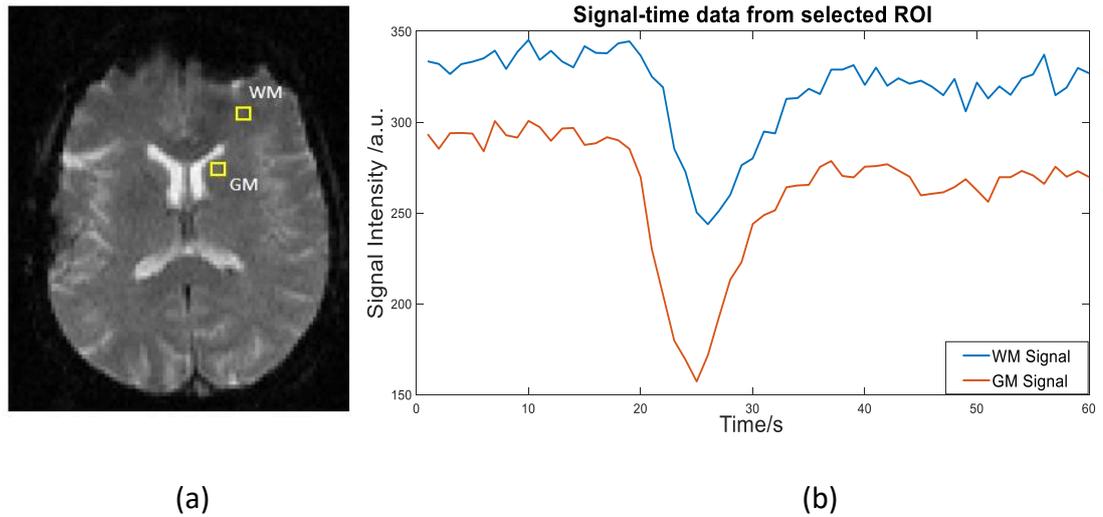


Figure 5.2: (a) Gradient-recalled-echo (GRE)-DSC-MRI (at the first time point), acquired at 1.5T. 3×3 -ROIs were placed in the caudate nucleus for GM and the frontal lobe for normal appearing WM (shown with yellow squares) (b) Corresponding average signal intensity curves (in arbitrary units, a.u.) for white matter, WM (in blue), and grey matter, GM (in brown), for the selected ROIs.

5.4.1.2. Automatic AIF detection

The arterial voxels were automatically identified following the processes used in several published studies (Yin et al., 2015, Mouridsen et al., 2006a, Emblem et al., 2009, Yin et al., 2014, Peruzzo et al., 2011). However, instead of searching AIFs from a manually-drawn ROI at the MCA or ICA like in those studies, every brain voxel from all slices was included in the search.

After background voxel removal and skull stripping, each brain STC was converted into a CTC using equations 3.1 and 3.2. Each CTC then went through noise filtering and thresholding stages to eliminate motion- or partial volume-affected voxels and tissue voxels, respectively. The stages are described as follows:

1. A roughness index, given as $RI = \int_0^T (C''(t))^2 dt$, where $C''(t)$ is the second derivative of CTC with respect to time, was calculated and 25% of the CTCs with the highest RI were discarded, as these are expected to be the results of motion or partial volume effects (Yin et al., 2014, Yin et al., 2015).

2. Out of the remaining CTCs, 50% with the highest full width at half maximum (FWHM) and 20% with the highest first moment (FM) were removed, since these are thought to represent soft tissue CTCs (Mouridsen et al., 2006a, Yin et al., 2015, Rempp et al., 1994) .

3. Out of the remaining CTCs, 30% with the lowest peak concentration were discarded, because their concentration peaks were not high enough to resemble those of AIF.

4. The areas under the remaining CTCs were calculated and 40% of the CTCs with the lowest area under the curve (AUC) were rejected, as these correspond to tissue CTCs where concentration changes are much smaller than those in arterial voxels.

The above-mentioned less-stringent thresholds allowed identification of all the true AVs, but at the cost of also identifying many false AVs. To identify and group the true AVs, the above-mentioned thresholding step was followed by a standard *k*-means clustering (Lloyd, 1982, Raykov et al., 2016) (described in Section 3.4.4.3 of Chapter 3). The remaining CTCs were divided into five clusters, putatively corresponding to GM, WM, arterial blood, venous blood, and ‘other’, such as ventricles containing cerebrospinal fluid (CSF) (Bjornerud and Emblem, 2010, Bleeker et al., 2011, Mouridsen et al., 2006a, Yin et al., 2014). Each cluster was represented by a centroid. The cluster whose centroid had the highest peak and lowest FM was chosen as the representative of AVs (Peruzzo et al., 2011) (Mouridsen et al., 2006a) (Yin et al., 2015). The final AIF was obtained by aligning and averaging the CTCs within that cluster.

5.4.1.3. Weibull TTD

The present study investigates the applicability of the Weibull distribution as a possible candidate for TTD to answer the first research question for this chapter (i.e. can modelling TTD with a Weibull function provide additional benefits over the available models?). The Weibull distribution is popular for modelling reliability data, hazard function, failure times, and analysing the lifetime of electrical as well as mechanical components (Lai et al., 2006). The normalised form of the Weibull distribution is given by:

$$h_{weibull}(t) = \alpha\beta^\alpha t^{\alpha-1} \exp[-(\beta t)^\alpha]; t > 0, \alpha > 1, \quad (5.12)$$

where α is the shape parameter and β is the reciprocal of the scale parameter, η (i.e. $\beta = 1/\eta$). The Weibull distribution was included as a trial TTD function for the following reasons. First, it conforms to all the constraints (described in Section 2.6.1) for a TTD to be physiologically realistic. Moreover, this distribution can transform into several different distributions by varying α . When $0 < \alpha < 1$, the TTD decreases exponentially from an infinite initial value. These values of α were avoided as the constraint $h(0) = 0$ is violated. With $\alpha = 1$, the distribution is a simple exponential, which is avoided here, as per the discussion in Section 2.6.1. With $1 < \alpha < 2$, the TTD rises sharply with slow rate of fall (positively skewed). With $\alpha = 2$, it turns into a Rayleigh distribution, and $3 < \alpha < 4$ turns it into a symmetrical, bell-shaped curve, resembling a Gaussian distribution, starting at $t = 0$. With higher α (> 10), it takes the shape of an extreme value distribution, which is negatively skewed (Abernethy et al., 1983). Therefore, by varying its shape and scale parameters, the Weibull distribution can potentially characterise a variety of microvascular environments, observed in both normal and pathological tissue.

5.4.1.4. Conversion of trial CTC to STC

The automatically-determined AIF was convolved with trial R s derived from each of the four functions (as per equation 5.1) and scaled by CBF to produce trial concentration functions (in accordance with equation 2.17). All trial concentration functions were then converted to trial signal time curves (STCs), which were later fitted to the measured signal. Signal was assumed to be related to concentration through the following equation (Patil et al., 2013):

$$S_t = S_{pre} \exp[-\Lambda(c)], \quad (5.13)$$

where S_{pre} is the baseline signal. Following the suggestion of Kao et al. and visual inspection of the signals, the first six images were discarded as they failed to display an equilibrium signal for the baseline calculation (Kao et al., 2010). S_{pre} was therefore quantified as the mean of the signal intensity values from the seventh image till the image preceding the bolus arrival.

$\Lambda(C)$ is a function of concentration that depends on whether contrast is present in large or small vessels. The form of this function also depends on the assumption of a linear or non-linear relationship between the GBCA concentration and the relaxation rate. Several researchers recommend using a non-linear relationship to reduce relaxivity-dependent systematic errors and increase the accuracy of absolute and relative perfusion measurements for a variety of imaging protocols (Patil et al., 2013, Patil and Johnson, 2013, Calamante et al., 2009, van Osch et al., 2003, Bjornerud et al., 2002). A non-linear relationship is therefore adopted in this work. In Chapter 3, a non-linear relationship was used for arterial $\Lambda(C)$; likewise, a non-linear $\Lambda(C)$ is used in this chapter for tissue regions. For tissue microvasculature (i.e. small vessels), Patil et al. (2013) used the interpolation formula of Yablonskiy and Haacke for $\Lambda(C)$, given by the following expression:

$$\Lambda(C) = \zeta \frac{1}{3} \int_0^1 (2+u) \sqrt{1-u} \frac{1 - J_0\left(\frac{3}{2} \omega \times TE \times u\right)}{u^2} du, \quad (5.14)$$

where J_0 is the zeroth order Bessel function, angular frequency ω is assumed to have the form $\omega = (a + bC) \times B_0$, where $a \approx 0$ and $b = 114.4 \text{ s}^{-1}\text{T}^{-1}\text{mM}^{-1}$ as per Patil *et al.*, and ζ is the dimensionless tissue vascular fraction, with a value of 3.8 for GM and 2.5 for WM (Yablonskiy and Haacke, 1994, Patil et al., 2013).

5.4.1.5. Initial parameter guesses

The trial signal function obtained from the above-mentioned step was then fitted to the normalised measured signal, S_t/S_{pre} . The STCs were fitted instead of CTCs because noise in signal (but not concentration) is independent of amplitude. For finding the optimal value of free parameters, this multi-dimensional non-linear optimisation problem requires a suitable initial guess, without which it can repeatedly get stuck in the local minima of the error surface. In the present work, fits were repeated for one hundred random combinations of initial guesses uniformly distributed over their likely physiological limits.

For each TTD to be physically realistic, the lower limits of its free parameters were controlled by the constraints imposed on them in Sections 5.2.3, 5.2.4, 5.2.5 and 5.4.1.3.

The upper limits were decided through either visual inspection of the input (i.e. AIF) and output (i.e. STC), or a trial-and-error-based preliminary study on two subjects.

5.4.1.6. Curve fitting and parameter estimation

For curve fitting, the 'lsqcurvefit' least-squares algorithm was used to find the values of the free parameters (i.e. F and TTD model parameters) that minimised the difference between estimated and normalised measured signals. Optimisation settings were as follows: algorithm: trust-region-reflective (Coleman and Li, 1996); step tolerance: 1×10^{-20} ; function tolerance: 1×10^{-20} ; maximum number of function evaluations: 5,000; maximum number of iterations: 2,000. All other settings were left to their default values, as described in MATLAB's documentation⁴.

The F (i.e. CBF) and the TTD model parameters obtained from the best fit with the lowest RMSE. MTT and CBV were calculated with equations 2.24 and 2.1, respectively.

5.4.2. Evaluation of Analysis

Plausibility of all four TTDs were first visually assessed by simulating their functional forms with different literature values corresponding to normal and pathological conditions. Afterwards, the goodness and stability of fit, consistency of perfusion estimates (CBF, MTT, CBV), and the computation time with all four h 's were compared through processes described below.

5.4.2.1. Simulation of TTDs for different tissues

To visualise each of the four TTDs in healthy and pathological conditions, a literature search was performed to obtain representative free parameter values. For healthy GM, ischaemic, and tumour tissues, the shape and scale parameters for gamma TTDs were obtained from Schabel (2012). With these parameters, three gamma TTDs and their

⁴ <https://uk.mathworks.com/help/optim/ug/lsqcurvefit.html>

corresponding R_s were simulated using equation 5.7 and 2.4, respectively. Three different CTCs were then created by convolving these R_s with AIF and scaling them by three different literature CBF values to represent healthy GM, ischaemic, and tumour tissue (Larsson et al., 2008, Larsson et al., 2017, Schabel, 2012, Vonken et al., 2000). The AIF was taken from one of the three subjects that had visually similar AIFs. Table 5.1 gives the shape and scale parameter along with the assumed CBF values used for simulating the healthy GM, ischaemic, and tumour CTCs.

Table 5.1: Shape and scale parameters for gamma TTD with cerebral blood flow (CBF) values for simulating healthy and pathological signals.

Tissue	Shape parameter, α	Scale parameter, β	CBF, ml/100g/min
Healthy GM	10 ^a	0.6 ^a	36 ^b
Ischaemic	90 ^a	0.1020 ^a	62 ^c
Tumour	1.5 ^a	7.86 ^a	87 ^d

- a. (Schabel, 2012); b.(Larsson et al., 2017); c. (Larsson et al., 2008);
d. (Vonken et al., 2000).

Simulated CTCs were converted to STCs via the process described in Section 5.4.1.4, and were later fitted to the four parametric forms of signal created from four different TTDs (as per the process described in Sections 5.4.1.5 and 5.4.1.6). After convergence, the free parameters were used to generate and plot each TTD to illustrate their functional forms in normal and pathological conditions.

5.4.2.2. Goodness of fit

To answer the second research question for this chapter (i.e. which model gives better goodness of fit for model-dependent deconvolution?), goodness of fit was assessed by calculating RMSE with the following equation:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n [S_t(t_i) - \tilde{S}_t(t_i)]^2}, \quad (5.15)$$

where n is the total number of images; $S_i(t_i)$ is the measured normalised DSC-MRI signal and $\tilde{S}_i(t_i)$ is the estimated normalised signal at i^{th} time point. Mean RMSE gives an estimate of the accuracy of the fits. The standard deviation (SD) of the RMSEs over multiple fits provides an indication of the precision of fits. Range, median, and upper and lower quartiles of RMSEs were calculated for both the GM and WM samples over all subjects for all four h 's.

5.4.2.3. Parameter Estimates

The CBV, CBF, and MTT estimates obtained with each TTD were calculated over all subjects in both GM and WM. They were compared with published literature values to assess the consistency of the perfusion estimates. The GM to WM ratio for each perfusion parameter was measured and compared with published values to further validate the consistency. Then, the estimates were compared with each other to answer the third research question for this chapter (i.e. how do the perfusion parameters vary with different models of h ?).

5.4.2.4. Success rate

For both GM and WM signal fitting, the success rate was quantified as the percentage of total fits converging to the global minimum. It evaluated the stability of fit: a model with a higher success rate evidently gave more fits that converged into the global minimum and gave the lowest RMSE, regardless of the initial guess. Success rates from four TTDs were compared to answer the second research question for this chapter regarding the comparison between the stability of fits of the four TTDs.

5.4.2.5. Computation time

Computation time (T_{comp}) represents the total time elapsed during the entire process: from fitting the trial signal to measured data for all hundred initial guesses, to finding the fit with the lowest RMSE, and finally calculating the perfusion parameters from the value

of the free parameters. Means and SDs of T_{comp} were measured and compared for all four h 's over all subjects for both GM and WM to answer the fourth research question for this chapter (i.e. how does the total computation time vary with different h ?)

5.4.3. Statistical Analysis

All statistical analyses were performed with the Statistical package for the Social Sciences (SPSS) software (IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.). In the present study, a set of dependent variables (RMSE, CBF, MTT, CBV, success rate, and T_{comp}) from a group of subjects were compared with different TTDs (repeated samples). For GM and WM regions, averages of these dependent variables were taken for each subject so they could be considered independent. Repeated measures within-subjects analysis of variance (ANOVA) was therefore used to assess statistically-significant differences between dependent variables obtained from different TTDs.

Repeated measure ANOVA assumes the dependent variables are normally distributed for every TTD (i.e. assumption of normality) and the variances of the differences between all combinations of TTDs are equal (i.e. the assumption of sphericity) (Field, 2013). The assumption of normality was tested with the Shapiro-Wilk test, which is suitable for small sample sizes, like the present study (Ghasemi and Zahediasl, 2012, Field, 2013). If the Shapiro-Wilk test gives a $p > 0.05$, the data are normally distributed.

The assumption of sphericity was tested using Mauchly's test. In cases of non-sphericity, degrees of freedom were corrected with either Greenhouse-Geisser, Huynh-Feldt or lower bound correction to decrease the Type-I error rate (Singh et al., 2013, Field, 2013). A correction factor, epsilon (ϵ), dictated the correction choice. The convention is to take the significance value of the Huynh-Feldt correction when $\epsilon > 0.75$ and Greenhouse-Geisser correction when $\epsilon < 0.75$ (Field, 2013).

Post-hoc tests were performed to indicate pairwise significance between TTDs. Repeated measures ANOVA has the null hypothesis that there are no differences between the population means of the dependent variables achieved with different conditions (i.e. different TTDs). Six separate hypotheses had to be checked to verify this null hypothesis.

To reduce the Type-I error of this multiple testing, the p -value for each comparison was calculated using Bonferroni correction, where the overall significance level of 0.05 was divided by the number of pairwise tests. Hence, for each pairwise comparison, the present work used $p = 0.008$ (Bland, 1995).

When the assumption of normality failed, the non-parametric Friedman's test was performed. In these cases, to identify pairwise significance, the Wilcoxon signed rank test was used with Bonferroni-corrected significance values.

5.5. Results

5.5.1. Functional forms of investigated TTDs

Figure 5.3 shows the functional forms of TTD investigated in the present study for healthy, ischaemic, and tumour conditions. For each condition, all TTDs, except gamma-variate, show similar functional forms with marginal variation in MTT.

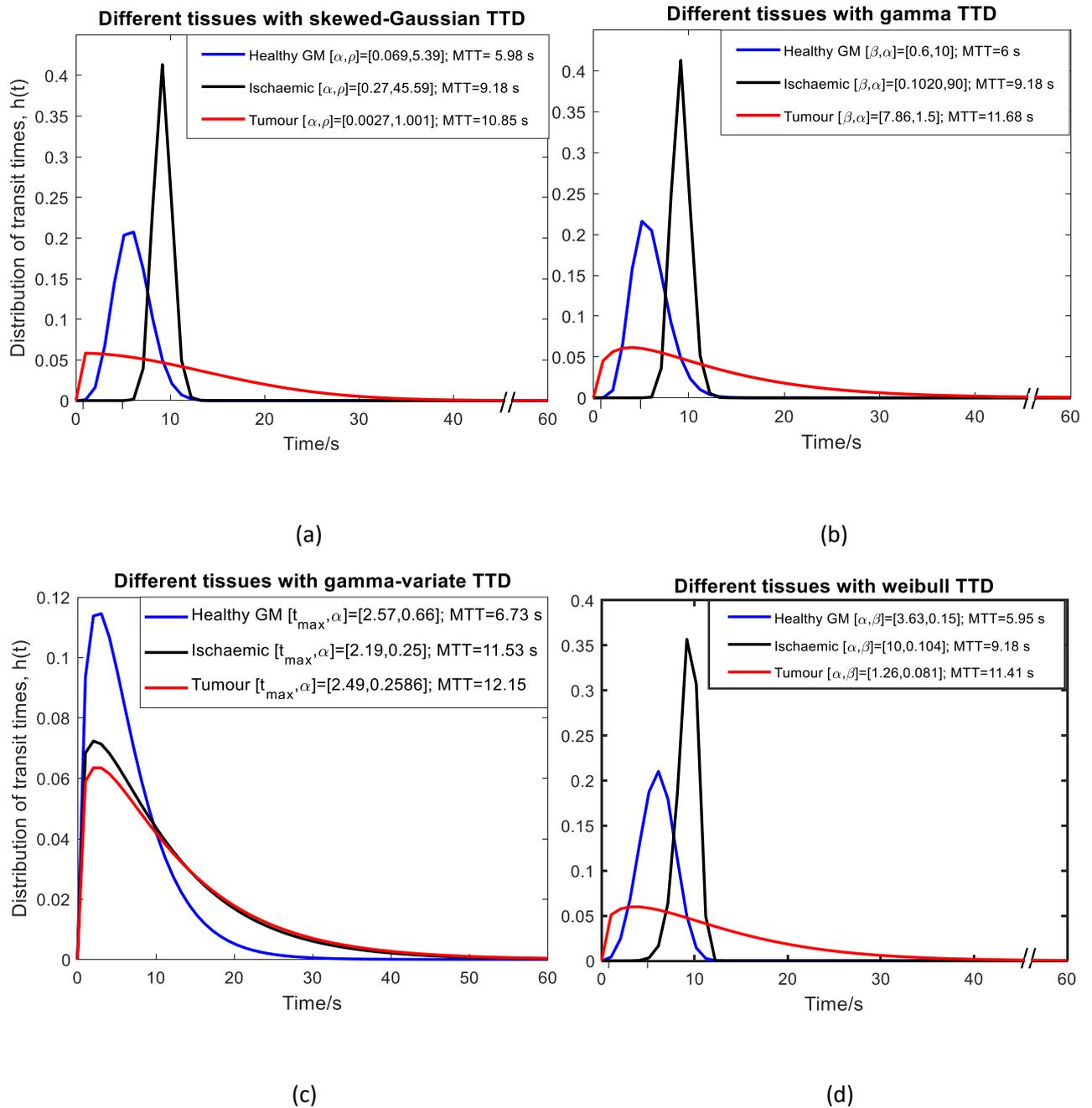


Figure 5.3: Functional forms of (a) skewed-Gaussian, (b) gamma, (c) gamma-variate, and (d) Weibull TTDs in healthy (blue), ischaemic (black), and tumour (red) conditions. For each TTD, free parameter values and mean transit time (MTT) in each condition are also given. Fitting the simulated signals gave unique gamma-variate TTDs, as they were modelled using decay and temporal parameters, unlike the shape and scale parameters of the other TTDs.

For gamma-variate TTD, one of the free parameters, t_{max} , is neither a shape or scale parameter; it is a temporal parameter that depends on the time difference between the AIF peak and the STC minimum. A time difference of 2–2.5s was observed between the AIF

and the simulated signals, which produced the shown t_{max} values and ultimately the unique functional forms of the gamma-variate TTDs. However, the obtained MTT values do not differ substantially between models. Furthermore, the gamma-variate functional forms follow all the constraints discussed in Section 2.6.1 for a TTD to be physiologically plausible. Therefore, the gamma-variate function can be considered as a possible form of TTD for obtaining perfusion estimates from different subjects.

5.5.2. Goodness of fit

Figure 5.4 (a-d) gives typical fits of estimated signal to a measured GM STC with each TTD. All fits capture the rapid initial drop in signal and accurately follow the recirculation bump. Figure 5.4 (e-h) also shows the corresponding TTDs for each fit. Although the shape of the gamma-variate TTD differs from other TTDs, the deviation is marginal and the resulting perfusion parameters are similar to those achieved with other TTDs.

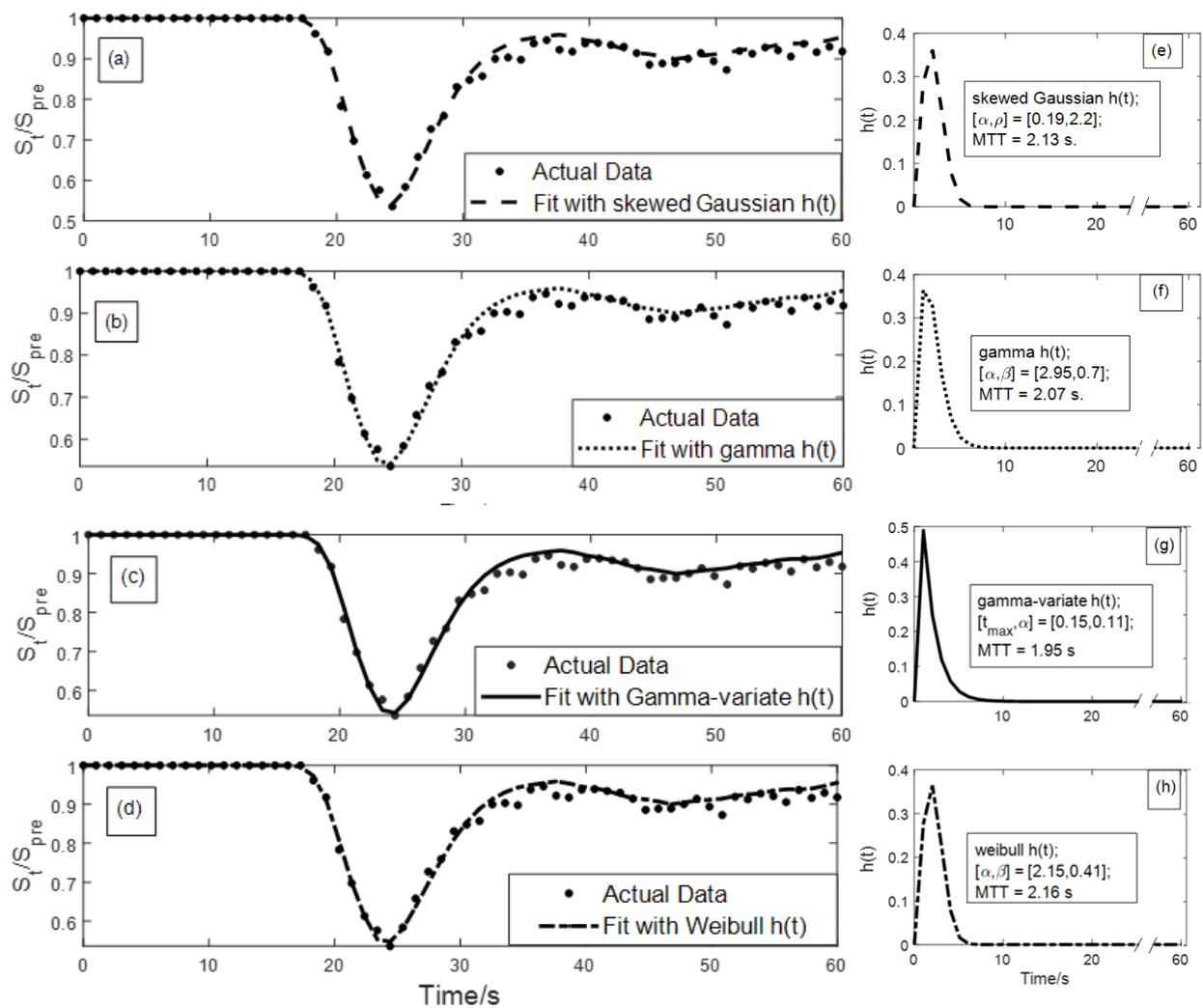


Figure 5.4: (a-d, left) Dot plots showing typical fits to a baseline-normalised GM signal time curve using skewed Gaussian, gamma, gamma-variate, and Weibull distributions, respectively. (e-h, right) Corresponding transit time distributions (TTDs) for each fit along with their free parameters and mean transit time (MTT) values. It can be seen that parametric signals obtained with all four TTDs fit well to the measured data, capturing the rapid signal drop and the recirculation stages accurately. All TTDs, except gamma-variate, have similar shapes and MTTs.

Table 5.2 gives the means and SDs of RMSE of the fits with four different TTDs averaged over all subjects for GM and WM.

Table 5.2: Mean (SD) of root mean square error (RMSE) obtained with each transit time distribution (TTD) for grey matter (GM) and white matter (WM)

Function		Skewed Gaussian	Gamma	Gamma-variate	Weibull
RMSE	GM	0.024 (0.0058)	0.024 (0.0058)	0.024 (0.0061)	0.023 (0.0055)
	WM	0.018 (0.0073)	0.016 (0.0050)	0.016 (0.0050)	0.016 (0.0060)

For the RMSE of GM fitting, the Shapiro-Wilk test indicated normality ($p = 0.257, 0.158, 0.374, 0.335$, for the skewed Gaussian, gamma, gamma-variate, and Weibull distributions, respectively; see Appendix 2: Table 1) and Mauchly's test indicated non-sphericity ($p < 0.001, \chi^2(5) = 186.606$, Appendix 2: Table 2a) for every TTD. After Greenhouse-Geisser correction of the degrees of freedom, ANOVA revealed no statistically-significant differences in the RMSEs of the fits obtained using different TTDs ($F(1.082, 37.856) = 2.595; p = 0.113$, Appendix 2: Table 2b).

For WM fitting, the Shapiro-Wilk test indicated that the RMSE values significantly deviated from a normal distribution ($p < 0.001$ for all four TTD; see Appendix 2: Table 1) and thus the non-parametric Friedman's test was performed. Friedman's test indicated that there were no statistically-significant differences between the RMSEs of the fits for WM signals with different TTDs ($p = 0.021, \chi^2(3) = 9.733$, Appendix 2: Table 3a).

Figure 5.5 gives boxplots of the RMSE for GM and WM obtained with the different TTDs, respectively. The range, quartiles, and median value of the RMSE of fits were similar for GM and also for WM signal fitting as shown in Figure 5.5(a) and 5.5(b), respectively.

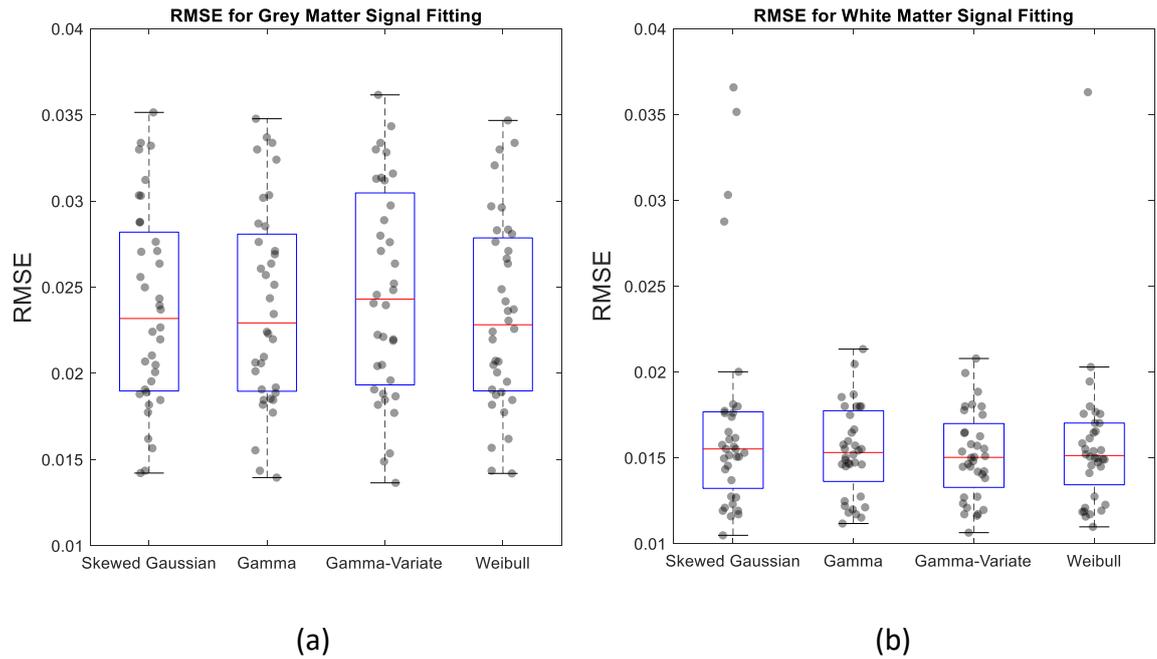


Figure 5.5: Boxplots of root-mean-square error (RMSE) values averaged over all samples and all subjects for (a) grey matter (GM) and (b) white matter (WM). The raw data are shown by black dots, median values by red lines, interquartile range by blue boxes, and upper and lower adjacent values by the black whiskers. No significant difference was obtained between RMSEs quantified from four TTDs for both GM and WM.

5.5.3. Parameter Estimates

Table 5.3 gives estimated mean values and SDs of CBF, MTT, and CBV with each TTD averaged over all subjects. The perfusion estimates obtained with all four TTDs are comparable with published literature values, as shown in Table 5.3. CBF values are in good agreement with the reported ranges for WM (20–40 ml/min/100g) and GM (60–80 ml/min/100g) (Bjornerud and Emblem, 2010, Ostergaard et al., 1998, Larsson et al., 2008). The MTT estimates are in line with the range suggested in a previous study (Ibaraki et al., 2007). The relevant range for CBV estimates was calculated by applying the central volume principle (i.e. $CBV = CBF \times MTT$) to the published CBF and MTT values and was found to be concordant with CBVs obtained in published work (Sourbron et al., 2009, Rausch et al., 2000).

The GM to WM ratios for CBF and MTT were in line with published values (Lieberman et al. 2015, Ostergaard et al. 1998) (Bjornerud and Emblem, 2010, Larsson et al., 2008,

Hakyemez et al., 2005, Ibaraki et al., 2007, Zhu et al., 2005). For CBV, the GM to WM ratios were also in the range reported by several previously published studies (Patil et al., 2013, Law et al., 2003, Zhu et al., 2005, Hakyemez et al., 2005).

Table 5.3: Mean (SD) of parameter estimates obtained with each transit time distribution (TTD), h

Function		Skewed Gaussian	Gamma	Gamma-variate	Weibull	Reported literature values
CBF, ml/min/100g	GM	68.99 (15.24)	69.36 (14.59)	71.33 (13.84)	66.21 (17.58)	50–80 ^a
	WM	26.51 (9.74)	24.46 (8.01)	25.11 (9.41)	23.61 (9.57)	20–40 ^a
CBF ratio	GM:WM	2.93 (1.06)	3.09 (1.00)	3.17 (1.17)	3.10 (1.08)	1.83–2.72 ^a 0.82–3.42 ^c 2.9–3.5 ^d 2.38 (0.18) ^f
CBV, ml/100g	GM	2.07 (0.53)	2.06 (0.55)	2.08 (0.46)	2.07 (0.52)	1.1–4.4 ^{b,g}
	WM	1.14 (0.30)	1.15 (0.30)	1.12 (0.30)	1.13 (0.28)	0.5–2.33 ^{b,g}
CBV ratio	GM:WM	1.84 (0.31)	1.82 (0.29)	1.90 (0.46)	1.85 (0.30)	0.72–2.01 ^c 0.91–3.05 ^e 2.04 (0.17) ^f
MTT, s	GM	1.93 (0.78)	1.91 (0.75)	1.92 (0.76)	2.10 (0.96)	1.2–3.3 ^{d,g}
	WM	2.99 (1.12)	3.21 (1.15)	3.11 (1.16)	3.36 (1.18)	1.5–3.5 ^d
MTT ratio	GM:WM	0.72 (0.31)	0.65 (0.22)	0.73 (0.51)	0.68 (0.28)	0.71–0.90 ^d 0.88 (0.04) ^f

Abbreviations: CBF, cerebral blood flow; CBV, cerebral blood volume; MTT, mean transit time; GM, grey matter; WM, white matter; TTD, transit time distribution; SD, standard deviation.

^a(Bjornerud and Emblem, 2010, Larsson et al., 2008) (Ostergaard et al., 1998)

^bCBV range calculated from concordant CBF and MTT ranges from published literature with formula $CBV = CBF \times MTT$ and found consistent to (Sourbron et al., 2009, Rausch et al., 2000)

^c(Hakyemez et al., 2005) ^d(Ibaraki et al., 2007) ^e(Zhu et al., 2005) ^f(Chou et al., 2007) ^g(Helenius et al., 2003)

5.5.3.1. CBF

In GM estimates of CBF, Shapiro-Wilk test indicated normality ($p = 0.674, 0.483, 0.631, 0.713$ for skewed-Gaussian, gamma, gamma-variate, and Weibull distributions, respectively; see Appendix 2: Table 1). Mauchly's test indicated that the assumption of

sphericity was violated as $p = 0.004$, $\chi^2(5) = 17.759$ (See Appendix 2: Table 4a). After Greenhouse-Geisser correction, ANOVA revealed no significant differences in estimates using different TTDs on the CBF value in GM ($F(1.314, 10.510) = 2.409$; $p = 0.147$) (Appendix 2: Table 4b).

In WM estimates of CBF, the Shapiro-Wilk test gave p -values of 0.300, 0.314, 0.078, 0.046, for skewed-Gaussian, gamma, gamma-variate, and Weibull distributions, respectively (see Appendix 2: Table 1)—confirming the normality of the first three. The marginal deviation from normality for the CBFs of Weibull TTD was considered trivial after visually inspecting the CBF distribution and comparing it to those of other TTDs. Non-sphericity was confirmed by Mauchly's test ($p < 0.001$, $\chi^2(5) = 23.045$, Appendix 2: Table 5a). ANOVA followed by Greenhouse-Geisser correction revealed no significant differences between estimates using different TTDs ($F(1.196, 9.569) = 1.617$; $p = 0.239$, Appendix 2: Table 5b).

Figure 5.6(a) and 5.6(d) give boxplots of the CBF estimates for GM and WM obtained with the different TTDs, respectively.

5.5.3.2. MTT

In GM estimates of MTT, the Shapiro-Wilk test indicated normality ($p = 0.374, 0.327, 0.239, 0.429$, for skewed-Gaussian, gamma, gamma-variate, and Weibull distributions, respectively; see Appendix 2: Table 1) and Mauchly's test indicated non-sphericity ($p = 0.001$, $\chi^2(5) = 21.109$, Appendix 2: Table 6a). After Greenhouse-Geisser correction of degrees of freedom, ANOVA revealed no significant differences between estimates using different TTDs ($F(1.324, 10.595) = 1.601$; $p = 0.241$, Appendix 2: Table 6b).

In WM estimates of MTT, normality and non-sphericity were confirmed respectively by the Shapiro-Wilk test ($p = 0.960, 0.957, 0.984, 0.687$, for skewed-Gaussian, gamma, gamma-variate, and Weibull distributions, respectively; see Appendix 2: Table 1) and Mauchly's test ($p = 0.013$, $\chi^2(5) = 14.723$, Appendix 2: Table 7a). ANOVA followed by Greenhouse-Geisser correction revealed no significant differences between estimates using different TTDs ($F(1.714, 13.711) = 3.864$; $p = 0.052$, Appendix 2: Table 7b).

Figure 5.6(b) and 5.6(e) give boxplots of the MTT estimates for GM and WM obtained with the different TTDs, respectively.

5.5.3.3. CBV

In estimates of GM CBV, the Shapiro-Wilk test indicated normality ($p = 0.720, 0.747, 0.161, 0.844$, for skewed-Gaussian, gamma, gamma-variate, and Weibull distributions, respectively; see Appendix 2: Table 1) and Mauchly's test indicated non-sphericity ($p < 0.001, \chi^2(5) = 37.724$, Appendix 2: Table 8a). After Greenhouse-Geisser correction of degrees of freedom, ANOVA revealed no significant differences between estimates using different TTDs ($F(1.135, 9.081) = 0.054; p = 0.851$, Appendix 2: Table 8b).

In WM estimates of CBV, normality and non-sphericity were verified respectively by the Shapiro-Wilk test ($p = 0.358, 0.105, 0.321, 0.074$, for skewed-gaussian, gamma, gamma-variate, and Weibull distributions, respectively; see Appendix 2: Table 1) and Mauchly's test ($p = 0.002, \chi^2(5) = 19.181$, Appendix 2: Table 9a). ANOVA followed by Greenhouse-Geisser correction revealed no significant differences between estimates using different TTDs ($F(1.292, 10.337) = 0.259; p = 0.681$, Appendix 2: Table 9b).

Figure 5.6(c) and 5.6(f) give boxplots of the CBV estimates for GM and WM obtained with the different TTDs, respectively.

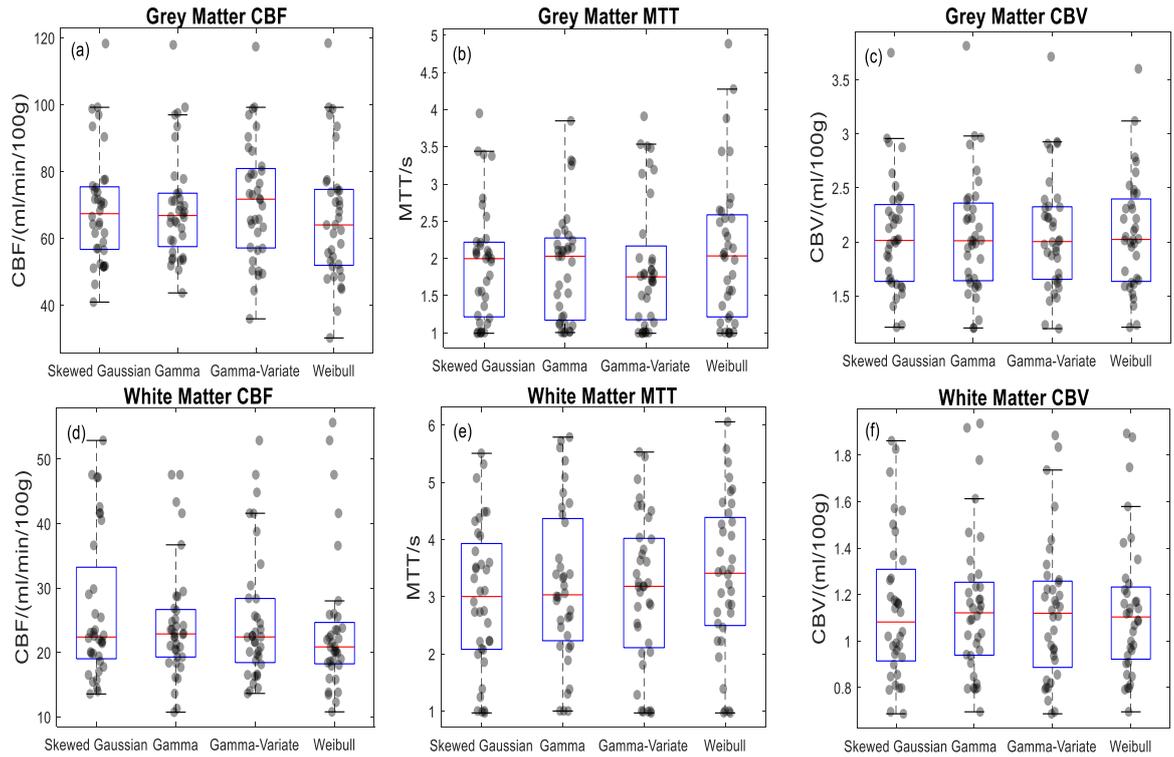


Figure 5.6: Boxplot of CBF, MTT and CBV estimates for GM (a-c) and WM (d-f) for four models. The raw data are indicated by black dots, median values by red lines, interquartile range by blue boxes, and 99% confidence bounds by the black whiskers. No significant difference was obtained between the estimates quantified from four TTDs for both GM and WM. Abbreviations: CBF, cerebral blood flow; CBV, cerebral blood volume; MTT, mean transit time.

5.5.4. Success rate

Table 5.4 presents the mean and SDs of the percentage success rate of fitting GM and WM signals with different TTDs. For GM signal fitting, the Shapiro-Wilk test indicated that the success rates significantly deviated from a normal distribution ($p < 0.001$; see Appendix 2: Table 1) and thus the non-parametric Friedman's test was performed, which indicated that there were statistically-significant differences between the success rates obtained with different TTDs ($p < 0.001$, $\chi^2(3) = 54.277$, Appendix 2: Table 10a). The post-hoc tests with Bonferroni-corrected p -values revealed that fits with the skewed Gaussian and gamma TTDs had significantly higher success rates than those with the gamma-variate and Weibull distributions ($p < 0.008$, see Appendix 2: Table 10b).

For WM signal fitting, the success rates were not normally distributed ($p < 0.001$; see Appendix 2: Table 1). Friedman's test indicated significant differences between the success rates obtained with different TTDs ($p < 0.001$, $\chi^2(3) = 34.919$, Appendix 2: Table 11a). The post-hoc tests using Bonferroni-correction revealed that fits with the gamma

distribution had a significantly higher success rate than all other distributions and the skewed-Gaussian TTD had significantly higher success rates than the Weibull distribution ($p < 0.008$, see Appendix 2: Table 11b).

Table 5.4: Mean (SD) success rate (percentage of successful fits) obtained with each TTD.

Function		Skewed Gaussian	Gamma	Gamma-variate	Weibull
Percentage Success	GM	63.2 (16.0) ^{3,4}	64.3 (28.8) ^{3,4}	21.1 (33.0)	22.0 (16.5)
	WM	61.0 (14.0) ⁴	74.4 (26.4) ^{1,3,4}	38.5 (36.3)	22.8 (13.5)

Note: ¹Statistically significant ($p < 0.008$) versus skewed Gaussian TTD; ²Statistically significant ($p < 0.008$) versus gamma TTD; ³Statistically significant ($p < 0.008$) versus gamma-variate TTD; ⁴Statistically significant ($p < 0.008$) versus Weibull TTD.

5.5.5. Computation Time

Table 5.5 gives T_{comp} 's for the four TTDs averaged over all subjects in both GM and WM. Fitting the GM signals took longer with a larger variation of T_{comp} than that of WM.

For GM signal fitting, the Shapiro-Wilk test indicated that the T_{comp} 's significantly deviated from a normal distribution ($p < 0.001$, for all four TTDs; see Appendix 2: Table 1). Friedman's test indicated that there was a significant difference between the T_{comp} 's for different TTDs ($p < 0.001$, $\chi^2(3) = 39.300$, Appendix 2: Table 12a). The post-hoc tests using Bonferroni correction revealed that fits were converging significantly faster for gamma TTDs than all other distributions ($p < 0.008$, see Appendix 2: Table 12b).

For WM signal fitting, the T_{comp} 's were not normally distributed ($p < 0.001$; see Appendix 2: Table 1). Friedman's non-parametric test indicated a significant difference between the T_{comp} 's with different TTDs ($p < 0.001$, $\chi^2(3) = 50.367$, Appendix 2: Table 13a). The post-hoc tests revealed that fits with the gamma and skewed Gaussian distributions took shorter time to converge than other distributions ($p < 0.008$, see Appendix: Table 13b).

Table 5.5: Mean (SD) of computation time (T_{comp}) obtained with each TTD.

Function	Skewed Gaussian	Gamma	Gamma-variate	Weibull	
$T_{comp, S}$	GM	533.3 (956.3) ^{2,3,4}	433 (557.4) ^{1,3,4}	766.3 (917.7)	613.6 (942.9)
	WM	249.9 (190.1) ^{3,4}	278.1(356.6) ^{3,4}	674.4 (227.8)	386.8 (216.1) ³

Note: ¹Statistically significant ($p < 0.008$) versus skewed Gaussian TTD; ²Statistically significant ($p < 0.008$) versus gamma TTD; ³Statistically significant ($p < 0.008$) versus gamma-variate TTD; ⁴Statistically significant ($p < 0.008$) versus Weibull TTD.

5.6. Discussion

The primary aim of this study was to compare the computational utility of four different forms of TTD for model-dependent deconvolution. The models were compared in terms of their goodness and stability of fits, consistency of perfusion parameters, and computation times. The study demonstrated that all functions gave similar fits and perfusion estimates. However, out of them, the gamma distribution offers at least two computational benefits: high stability of fit and shorter computation time.

5.6.1. Comparison between TTD models

Curve fits—with all functions—were excellent, with low RMSEs. All perfusion estimates from the four TTDs are congruent with previously reported values. Since all functions gave similar RMSEs and perfusion estimates, the most suitable TTD would be the one that provides at least one computational benefit, such as, shorter computation time or higher fit stability (Sourbron and Buckley, 2012).

It should be noted that while evaluating the suitability of models, both the goodness of fit and the model complexity should be considered. Models with a higher number of free parameters may give better goodness of fit, but they increase the computational complexity and can ultimately overfit the signals. The present study does not consider model complexity as there is implicit control for the number of degrees of freedom—each parametric signal model has three free parameters: CBF and two TTD parameters. Therefore, goodness of fit (measured by RMSE) alone was deemed sufficient to

determine which model was most suitable. When comparing models with different degrees of freedom, the Akaike information criterion (AIC) (Akaike, 1974) or Bayesian information criterion (BIC) (Schwarz, 1978) can be used. AIC considers the sum of squared errors with the number of free parameters in the models. Similar to the AIC, the BIC provides a trade-off between goodness of fit and model complexity.

Empirically, while BIC is biased towards simple models, AIC tends to be biased towards more complex ones (Penny et al., 2004). It is therefore often considered sensible to perform both BIC and AIC; if BIC indicates a three-parameter model as suitable while AIC chooses a five-parameter model, suitability of the models with three, four, and five parameters should be further examined for the application in hand. In future studies, a Bayesian framework—a more flexible model building approach (Mouridsen et al., 2006b)—can be used to model the DSC-MRI signals, as it facilitates the quantification of BIC using model evidence.

The present research indicates that the gamma TTD produces more stable fits; that is, a higher percentage of fits converged to the global minimum. This is likely because the gamma function results in a smoother (i.e. ‘less bumpy’) error surface with fewer local minima than other TTDs. Another reason could be the ability of the gamma TTD to adapt—by changing its α^{-1} —to other models of tissue microvasculature, such as Tofts-Kety (Tofts et al., 1999), extended Tofts-Kety (Buckley, 2002, Tofts, 1997), two compartment exchange (Hoffmann et al., 1995), and adiabatic tissue homogeneity models (St Lawrence and Lee, 1998, Schabel, 2012). Hence, it can effectively model a range of microvasculature that would have otherwise required multiple unique TTDs. Additionally, the use of a gamma TTD accelerates curve-fitting, as there is a higher chance of convergence irrespective of the initial guesses. So, compared to other TTDs, for a gamma TTD, the curve fitting step needs to be iterated fewer times with different initial guesses. Due to these computational benefits, the present research concludes gamma TTD model as more suitable for any perfusion quantification applications; specifically, it can be useful in clinical cases that require rapid analysis and decision-making, such as acute stroke.

5.6.2. Clinical applications

In addition to the usual perfusion parameters, the model-based approach taken in this work lends itself to the calculation of parameters related to the width and shape of the TTDs. Vessels created by tumour angiogenesis are chaotically structured, dilated, and irregularly shaped, so the GBCA particles require a wide range of transit times to traverse them. Consequently, the TTDs for the tumour regions will be wider than those of normal regions. The width and shape of TTD can be used to distinguish between normal and tumour vessels, as well as to grade the tumours.

The model-dependent deconvolution used in the present work can characterise the residue function using only two free parameters, without estimating it at every time point. Consequently, the present analysis is more robust against experimental noise than the voxel-by-voxel model-independent deconvolution (Ostergaard et al., 1999), and offers an improved approach for perfusion quantification at high spatial resolution. The presented ROI-based perfusion measurement can be extended to estimate pixel-wise perfusion parameters and create brain maps of CBF, CBV, and MTT in clinical settings. All four models can be used for this purpose; however, for rapid measurement, gamma TTD will be more suitable than the others.

Brain maps created with the TTD parameters can inform clinicians about the capillary flow profiles and allow them to characterise tissue viability in ischaemia (Mouridsen et al., 2006b). Flow heterogeneity and oxygen extraction fraction can be determined using MTT and capillary transit time heterogeneity, the SD of the TTD (Ostergaard et al., 1999, Mouridsen et al., 2014). Due to its high fit-stability and low time-complexity, the gamma TTD can offer a fast and effective method for quantifying these parameters. Changes in flow heterogeneity can identify the effect of different pathologies on the delivery of nutrients and oxygen through the capillary. For example, elevated capillary transit time heterogeneity reduces tissue oxygen availability in acute ischaemic stroke, tumours and Alzheimer's disease (Mouridsen et al., 2014).

5.6.3. Limitations

The major limitation of this study is that perfusion estimates were not compared directly with estimates obtained by the reference-standard of positron emission tomography (PET) perfusion imaging. However, the primary purpose of the study was to compare different forms of h , rather than assessing the absolute accuracy of the approach. Nonetheless it is highly probable that some of the variation seen in estimates is due to natural variation between subjects and this might mask systematic differences in results obtained with different forms of h . A larger sample size can better capture the physiological variation and further validate the inferences made.

A second limitation is that the AIF detection stage used k -means clustering, mainly for simplicity. As discussed in Chapter 3 (Section 3.6.5), k -means clustering has several limitations, such as high sensitivity to outliers, limited adaptability to data that need clusters of different density, and low accuracy when non-spherical (e.g. elliptical) clusters are needed (Raykov et al., 2016). As noise filtering and thresholding were already performed prior to this stage, clustering only simplifies and refines the otherwise-intensive manual AIF search. Nevertheless, future works can explore different clustering methods to investigate whether any better AIFs can be obtained.

The present study searched for a ‘global AIF’ under the premise that it gives a reasonable representation of the arterial input to every ROI. Global AIFs can get delayed in reaching the ROI when they are spatially distant from the selected arterial voxel site. Additionally, there can be a significant effect of dispersion for the feeding vessel, which spreads the arterial input bolus. This dispersion effect is more pronounced when blood needs to pass through stenoses with marked turbulence or irregular collateral paths located upstream of the global AIF site (Ostergaard et al., 1999). However, the present patient cohort had no reported vessel disease, so the effect of dispersion can be regarded as trivial. Additionally, it is shown that the flow estimates are independent of vascular delay for model-dependent approaches (Ostergaard et al., 1999). Therefore, the present work avoided including delay and dispersion in the AIF. Saying that, future extensions of this work can look into finding local AIF in the tissue neighbourhood or including a function that considers the effect of delay and dispersion of AIF during its passage from arterial site to the tissue of interest.

The implications of global AIFs and some potential improvements in their detection technique are discussed in Chapter 6.

The present study quantified perfusion parameters in selected GM or WM ROIs only. When applied on real subjects to create brain perfusion maps, several thousands of voxels need to be analysed. Curve-fitting for all these signals will ultimately increase the computation time. Different configuration choices, such as initial guesses, constraints imposed on the free parameters, cost functions, algorithms, and tolerances, may negatively affect the reproducibility of parameters across centres. One way to overcome this drawback is to use an alternative dictionary-matching approach proposed in a recent DCE-MRI study (Ghudasara et al., 2020), which decreased the computation time by 140-fold, while keeping the perfusion estimates similar. Rapid generation of brain maps are further discussed in Chapter 6.

Lastly, the study did not consider pathological tissues. For normal GM and WM signals, all TTDs are expected to give similar estimates. Only when signals from pathological tissues are considered, would the TTDs be more likely to deviate significantly and thereby produce statistically-different perfusion estimates. However, the aim of the present study was not to estimate perfusion in pathology, but rather to investigate the relative utility of TTDs for automatic analysis. Moreover, the small sample size of this study may have led to it being underpowered. The fact that no statistically-significant differences were observed does not necessarily mean that no differences exist between the perfusion estimates. A much larger sample may have revealed statistical differences. Therefore, the conclusions of this exploratory study regarding perfusion estimates should be generalised cautiously. In future extensions of this study, more subjects will be included in the analysis and pathological signals will be fitted to address these limitations.

5.6.4. Future scopes

As mentioned above, one of the possible sources of error in the perfusion measurement arises from the application of a global AIF, which does not consider the effects of delay and dispersion. Future extensions of this research can convolve the detected AIF with a function that includes two additional free parameters to represent the delay (i.e. shift) and dispersion (i.e. broadening) of the global AIF (Sourbron and Buckley, 2013). This would

establish a five-parameter model-dependent deconvolution approach, which could be used as a general framework for the perfusion analysis of ischemic stroke patients (Ostergaard et al., 1999) (further discussed in Chapter 6).

The effect of noise on perfusion parameters can be investigated for each TTD. Using Monte-Carlo simulations, noise can be added to raw GM and WM data to produce signals with a range of SNRs. The perfusion estimates from the raw GM and WM signals (true parameter value) and those from the simulated noisy signals can be calculated with the process described in Section 5.4.1. The difference between the true parameter value and the estimated parameter value averaged across different noise realisations will give a measure of accuracy (i.e. bias); whereas the standard deviation of the estimates will indicate the precision (i.e. dispersion) of the perfusion parameters (Cameron et al., 2017).

When modelling the TTDs, the leakage of GBCA was not considered. This study can be extended to assess and monitor perfusion or permeability in subjects with mild ischaemic stroke (Heye et al., 2016), multiple sclerosis, optic neuritis (Larsson et al., 2017), or early signs of Alzheimer's disease (Wardlaw et al., 2017, Haar et al., 2016), where subtle leakage from BBB is present. If a small fraction of GBCA is bound irreversibly in the tissue (Larsson et al., 2017), i.e. there is no back diffusion from tissue to blood, the area under the TTD can be normalised to $1-E$, instead of unity; where E is the extraction fraction. However, this process is valid only when the leakage is small. In case of high permeability, this procedure will be inadequate (Larsson et al., 2017) and the volume of extravascular extra-cellular space should be added as another free parameter in the least-squares algorithm (Schabel, 2012).

5.7. Conclusion

This study concludes that the gamma distribution is superior to other plausible TTD functions, including the proposed Weibull distribution. Although all four functions gave perfusion estimates similar to published studies, the gamma TTD offers significantly faster convergence with higher stability of fit than other TTDs. Therefore, it can significantly decrease the computation time as a lower number of initial guesses would need to be explored to find the best fit. All the models can be applied in future research including, but not limited to, identification of tumour grade, assessment of tissue viability

in ischaemia, and estimation of flow heterogeneity and oxygen extraction capacity. Saying that, the gamma distribution can specifically facilitate rapid measurement, which can be useful for perfusion analysis of acute stroke patient.

Chapter 6: Conclusions and Future Directions

The investigations presented in this thesis assist three important intermediate steps of the dynamic susceptibility-contrast magnetic resonance imaging (DSC-MRI) analysis: arterial input function (AIF) detection, tissue segmentation, and model-dependent deconvolution. With specific emphasis on automation and data-mining, methods are proposed to avoid intensive manual labour and establish a simplified and accelerated perfusion analysis platform. A *more-systematic* approach towards tissue-voxel-elimination and subsequent AIF detection is now possible with the application of criteria thresholds (Chapter 3); other sites can use this approach as a general framework to set their own data-specific thresholds. With the knowledge of the *individual* arterial voxel (AV)- and tissue voxel-discriminatory powers of different features (Chapters 3 and 4, respectively), Radiologists or automatic algorithms can efficiently examine the arterial candidacy or tissue class of a voxel. The application of these features in the proposed feature-based segmentation (Chapter 4) can facilitate rapid, but effective, region discrimination and lesion identification. Once the tissue class for a voxel is identified and the arterial input to it is determined, its dynamic signal is analysed to estimate perfusion parameters. The identified *computational benefits* of the gamma function (Chapter 5) as transit time distribution (TTD) model will promote its use in clinical perfusion measurements, especially for pathologies requiring rapid processing and decision-making (e.g., acute stroke). More detailed discussion of these outcomes, as well as the limitations and future directions of this work are presented in the following sections.

6.1. Research outcomes

In Chapters 3 and 4, several semi-quantitative parameters were extracted from the DSC-MRI data and their effectiveness in discriminating arterial and tissue voxels was evaluated. The studies demonstrated that some of these semi-quantitative parameters can effectively classify the voxels into artery, grey matter (GM), white matter (WM), and lesion, without transferring the data to a separate segmentation platform. It has also been

shown that with an optimal threshold unique to each effective criterion, tissue voxels can be discarded with high sensitivity and specificity. Although these thresholds were not generic for all datasets and imaging protocols, the approach can serve as a general framework for the determination of tissue-elimination threshold for other sites. Further, using the relative criteria-effectiveness knowledge, future studies can make more informed choices about which features to prioritise while searching global and local AIF, or distinguishing specific tissue regions.

Data-mining has been implemented here to accelerate the automatic AIF detection and segmentation. Clustering reduced the manual labour otherwise necessary for labelling voxels as arterial or of a specific tissue class. It was demonstrated (in Chapter 4) that conventional raw-data-based clustering can be outperformed by the proposed feature-based approaches that cluster a dimensionality-reduced data space. Although for some regions its performance was similar to other approaches, the feature-based segmentation should be implemented for future perfusion analysis due to its lower computation time. Besides increasing the automation and decreasing the computation time, the present study proposes a simplified analysis pipeline. Here, segmentation was performed using the raw DSC-MRI data or its dimensionality-reduced version. No prior knowledge of feature distributions was required and the dynamics were not transferred to other processing platforms, for example, Statistical Parametric Mapping (SPM) (Ashburner and Friston, 2000), Functional Magnetic Resonance Imaging of the Brain (FMRIB) software library (FSL) (Jenkinson et al., 2012), or some semi-automated integration of these, like the Diffusion/Perfusion Project (DPP) Suite (Revenaz et al., 2016). Additionally, for AV detection, no Radiologists were required to delineate any ‘preferred’ region-of-interest manually, rather the algorithm searched for AVs from all brain voxels of all slices. Consequently, the proposed approaches comprise a simple, objective, rapid, and effective perfusion quantification platform that can be used in the clinical setting to mitigate the time-delay during data processing and treatment planning.

Despite their reported high AV- and tissue voxel-discriminatory power, physiological interpretation using the semi-quantitative parameters is complex. These parameters have a convoluted relationship to perfusion parameters and are sensitive to non-tissue-related characteristics, such as injection rate, cardiac output, vascular structure, and arterial input to the tissue (Calamante, 2013, Perthen et al., 2002, Calamante, 2012). So, perfusion estimates were made using quantitative model-dependent deconvolution approaches.

Model-independent approaches were avoided as they are not only unable to ensure the monotonicity and non-negativity constraints imposed on the residue functions, but they are also sensitive to the delay and dispersion of the bolus.

For perfusion quantification, in Chapter 5, different models of transit time distribution (TTD) were compared. All models gave estimates congruent with those reported in published studies (Bjornerud and Emblem, 2010, Larsson et al., 2008, Hakyemez et al., 2005, Larsson et al., 2017, Ibaraki et al., 2007, Ostergaard et al., 1998). However, the gamma distribution was found to be the most suitable TTD for perfusion estimation, as using it provides at least two additional computational benefits over the other TTD models: lower computation time and higher stability of fit.

Throughout the thesis, simplicity, objectivity, and rapidity of the analysis is prioritised over absolute accuracy and computational intricacy. For example, *k*-means and *k*-medoids clustering were used instead of more complex and time-consuming unsupervised or supervised algorithms, such as hierarchical clustering, support vector machines, expectation-maximisation, or self-organising maps. This trade-off between accuracy and complexity permits rapid lesion-diagnosis, decision-making, and progression-assessment for future perfusion analyses. The methods can specifically assist in ‘time is brain’ situations, such as acute stroke where rapid diagnosis and decision-making is of crucial importance.

6.2. Limitations

The methods chapters of this thesis (Chapters 3, 4, and 5) contain their specific limitations. In this section, some general limitations of the overall work will be discussed.

6.2.1. Global AIF

The present study used an idealised arterial input (typically referred to as ‘global’ AIF) for all tissue voxels. This global AIF can be delayed (i.e., shifted) and dispersed (i.e., broadened) as it reaches the voxel of interest; thus, it may not accurately represent the

arterial input to a given voxel. Consequently, using this idealised AIF may have introduced systematic error to the perfusion estimation study. Bolus delay can cause both over-estimation and under-estimation of cerebral blood flow (Wu et al., 2003a). This can negatively affect the detection of lesions and even misinform about their actual hemispheric location (Calamante, 2013). Additionally, if the ‘actual’ arterial supply reaching the tissue is broader than the global AIF (due to dispersion), the perfusion analysis wrongly interprets this dispersion as occurring in the tissue and overestimates the contribution of the tissue towards the bolus spread. Consequently, the mean transit time is over-estimated, underestimating the cerebral blood flow.

In this work, the effects of bolus delay and dispersion are minimal, if not trivial, for the reasons given as follows. First, a model-dependent deconvolution approach is used here, which gives flow estimates that are vascular-delay independent (Ostergaard et al., 1999). Moreover, dispersion is minimal as the patient cohort did not have any reported arterial stenosis. Due to unavailability of local AIF detection tools on scanner workstations, the vast majority of DSC-MRI analysis tools rely on the global AIF, as in this work (Calamante, 2012, Calamante, 2013, Bjornerud and Emblem, 2010, Mouridsen et al., 2006a). Therefore, with a controlled global AIF, the overall inferences drawn from the comparison of TTDs are still valid.

Saying that, there is scope for establishing a rigorous global AIF detection method by including more systematic quality control. For example, multiple Radiologists could be asked to verify the suitability of the finally-selected AVs by assessing their spatial locations. The difference between manually- and automatically-obtained AIFs can be evaluated via root-mean-square error (Yin et al., 2015, Yin et al., 2014), and their agreement can be evaluated by Bland-Altman analysis (Wong et al., 1998). Additionally, the relative agreement can be evaluated using a non-metric version of multidimensional scaling: as discussed by Mouridsen et al. (2006) and Yin et al. (2014).

6.2.2. Partial volume effect

The arterial signal can be distorted due to the partial volume effect (PVE). This arises due to the lower spatial resolution of DSC-MRI, which can lead to voxels containing both artery and tissues. Their complex DSC-MRI signals are superimposed, which leads to

distortion or shape-errors in the AIF. This PVE can create narrow and sharp concentration time curves (CTCs) that survive the thresholding stage and are erroneously considered 'correct' AIFs. The AIF-detection algorithm here did not explicitly consider PVE. Hence, there exists a good chance that many peak-shaped CTCs passed through the thresholding stage.

Many published studies removed PV-affected CTCs by using a venous-outflow function (Bjornerud and Emblem, 2010, Larsson et al., 2017) or by fitting gamma-variate to the CTC first pass (Bleeker et al., 2011, Yin et al., 2014, Yin et al., 2015). However, these methods require further manual involvement and add more complexity, while reducing objectivity and reproducibility. The present work circumvented these issues by using clustering, which grouped CTCs with similar set of features. The PV-affected CTCs with no first-passage peaks and those with two first-passage peaks in close succession were never selected, as the features extracted from them were completely different to those of candidate AIFs. The remaining peak-shaped PV-affected CTCs were discarded by visual inspection of a relatively small number of arterial candidates. Although this process is somewhat subjective, it is no more so than the available PV-correction approaches mentioned above. Moreover, the present process is independent of 'T₁-shine-through' and does not require measurement of the venous output function, the steady state of the CTC, or the complex gamma-variate fitting to obtain the area under the CTC first pass. Yet, future studies could assist PV-correction by finding novel, composite parameters that have high PV-discriminatory power, as discussed in Section 3.6.6.

6.2.3. Size of Patient Cohort

The small patient cohort may affect the statistical power of the analysis. However, for an exploratory study like the present work, this does not undermine the inferences made from the comparative analyses here. Furthermore, several findings of this work are already congruent with published studies, such as: the optimal threshold of the area under the curve for tissue-voxel elimination (Mouridsen et al., 2006a); the suitability of area under the curve and peak concentration as AV-detection criteria (Yin et al., 2014, Peruzzo et al., 2011, Bjornerud and Emblem, 2010); the spatial locations of the finally identified AVs (Zaro-Weber et al., 2012, Wismuller et al., 2006); the lower time-complexity of the

proposed feature-based clustering (Wang et al., 2006b); the suitability of the gamma distribution (Schabel, 2012); and the resultant perfusion parameters (see Table 5.3). Nevertheless, in future studies, a larger, more diverse patient cohort should be used to validate the conclusions of this study.

The comparative analysis of this study should be applied to different pathologies, to investigate whether models of TTD need to be varied with pathology. For example, for subjects with arterial abnormalities, inclusion of delay and dispersion is necessary; a brain map of delay and dispersion can supplement the perfusion information. The perfusion estimates for different healthy age-groups, races, or genders can be recorded to better identify the abnormalities observed in the patients. For example, perfusion in age- or gender-, if possible, race-, or activity-matched healthy subjects can be analysed to isolate the effect of pathology from those of age, gender, race, or physical activity.

6.3. Future directions

The work succeeded at simplifying and accelerating the perfusion analysis and making it more automatic and objective. Future extensions of this work can aim at establishing an end-to-end automatic perfusion platform, investigating the diagnostic utility of the extracted semi-quantitative parameters, validating the inferences using available *in silico* datasets, and assisting the local AIF detection. These opportunities are discussed in the following sections.

6.3.1. Relating semi-quantitative parameters to perfusion estimates

As mentioned before, the diagnostic utility of semi-quantitative parameters is not clear due to their complex relationship with perfusion estimates. The present study can be extended to address this knowledge gap by investigating the relationship between the extracted semi-quantitative parameters and the quantified perfusion or TTD parameters. This can indicate the utility of semi-quantitative parameters as biomarkers of different pathologies. Additionally, investigations can ascertain the clinical utility of TTD

parameters in distinguishing normal and tumour vessels, informing the tumour grades, assessing tissue viability in ischaemia, and estimating flow heterogeneity.

6.3.2. Developing composite parameters

Multiple semi-quantitative parameters can be combined using different parametric forms to develop novel, composite parameters. With the help of receiver operating characteristic curves, experiments can find the most suitable parametric form for efficient identification of tissue, arteries, lesions, and PV- or noise-affected voxels. The association between these computationally-simple parameters and perfusion estimates of a voxel may provide supplementary information about its oxygen extraction fraction, capillary tissue inhomogeneity, or tumour grade (Mouridsen et al., 2006a, Mouridsen et al., 2014)

Composite feature spaces or parameters can also be created by combining features extracted from different registered perfusion MRI modalities into a single platform. Instead of dynamic intensity values, a voxel can be characterised by a feature vector containing these cross-modality features; thus, more information can be gathered into a dimensionality-reduced, composite feature space. This will increase the orthogonality between the voxels, allowing the clustering algorithm to better discriminate the regions into different types of lesions (i.e., penumbra, infarct, haemorrhage, tumour), normal-appearing GM and WM, arteries, veins, ventricles etc. Clinicians can simultaneously analyse different signal attributes for a voxel of interest to obtain better insights about its pathological status; researchers can investigate the utility of these composite parameters in initial diagnosis, characterisation, and treatment monitoring.

6.3.3. Local AIF detection

As discussed in Section 6.2.2 and also in Chapter 5, the present work does not consider the effects of delay and dispersion on AIF. Future studies can consider this by convolving AIF with a function that describes the transport of contrast agent from any global AV site to the voxel of interest (Sourbron and Buckley, 2013). However, this requires knowledge

of the vascular architecture and can often be impossible to predict for a pathological region (Calamante, 2013). Instead, detection of *local* AIF (a unique arterial input for each voxel or its neighbourhood) may improve the analysis.

In many published local AIF detection works, certain criteria, such as effective bolus arrival time, identify the tissue neighbourhoods that are supplied by the same arterial input; within this small neighbourhood, an established global AIF detection methodology (like the work of Mouridsen et al. (2006), discussed in 3.2.1) is applied (Tabbara et al., 2020, Willats et al., 2011). Many aspects of the present work can assist these local AIF detection studies. For example, the individual criteria-effectiveness knowledge would inform which criteria to prioritise while generating a tissue neighbourhood; the proposed feature-based clustering would accelerate grouping similar CTCs; data-specific optimal thresholds would discard tissue voxels with high sensitivity and specificity. Additionally, AVs and different tissue regions in the neighbourhoods—of sizes similar to the 16×16 -dimensional digital brain phantom (Section 4.4.2)—can be classified rapidly by applying the kmed-feature approach (Section 4.4.7.1).

6.3.4. Validation using digital reference object

In this thesis, real-subject DSC-MRI data were analysed to create a ground truth for both AIF detection and the segmentation work. Application of simulated dataset could have been used to validate the workflow. Many published studies have simulated DSC-MRI data (Mouridsen et al., 2006a, Yin et al., 2014, Yin et al., 2015); however, their choice of exponential residue functions violates one of the basic tenets of DSC-MRI: the intactness of the blood-brain barrier. Moreover, they only modelled the CTC first pass, neglecting any recirculation. By using real-subject data, this work overcomes these limitations and gave an indication of the applicability of the assistive approaches in the clinical setting, on noisy data. Although this work decreases the human component in the analysis, it still depends on manual interventions, especially for these ground truth generation stages (Sections 3.4.4.5, 4.4.1, and 4.4.2).

In future extensions of this work, the total workflow (i.e., AIF identification, brain segmentation, and perfusion quantification) could be validated using population-based DSC-MRI digital reference objects (DROs) that can capture the heterogeneous signal

characteristics found in *in vivo* glioblastoma datasets (Semmineh et al., 2017, Semmineh et al., 2014). There are collection of DROs available for download from The Cancer Imaging Archive (www.cancerimagingarchive.net) under the collection name “Borrow-DRO”. These DROs contain cohorts of virtual patients with AIFs, normal tissue and tumour voxels similar to clinical DSC-MRI datasets; however, they differ in pulse sequence parameters and preload dosing schemes. Using these diverse DROs, the tissue-voxel-elimination thresholds can be optimised; the feature-based segmentation approaches can be validated; and the most effective TTD across different imaging protocols and concentration dosages can be ascertained. Once validated across diverse DROs, the optimised workflow can then be applied to clinical glioblastoma patients.

6.3.5. Brain map creation

The present work of ROI-based perfusion estimation (Chapter 5) can be extended to obtain pixel-wise perfusion estimates by analysing all brain signals. Then, brain maps of relevant perfusion parameters can be created. Although curve-fitting converges within the order of several minutes for ROI-based estimation, when applied pixel-by-pixel, time-complexity will evidently increase. Application of the proposed feature-based segmentation and gamma TTD will accelerate pixel-by-pixel perfusion estimation. Further acceleration can be achieved through parallelisation of the analysis over multiple professional workstations. As an alternative to time-consuming curve-fitting approaches, future investigations could also explore dictionary-matching (Ghudasara et al., 2020), where trial CTCs are produced using a subject-specific dictionary that consists of permutations of perfusion parameters. Trial CTCs are then matched with the measured data to obtain the best perfusion estimate.

This dictionary-matching process may significantly accelerate the production of brain maps, as demonstrated in a recent dynamic contrast-enhanced MRI study that characterised uterine fibroid (Ghudasara et al., 2020). Once generated, these brain perfusion maps can be examined by Radiologists and clinicians to identify and characterise the lesion region, decide an appropriate treatment plan, or assess the response to an antiangiogenic drug.

6.3.6. End-to-end automated approach

As mentioned before, the present work has significantly decreased the human involvement in the DSC-MRI analysis. However, there is further scope for decreasing the manual involvements in the steps like: nominal AIF listing (Chapter 3), digital brain phantom generation, and the classification of cluster centroid (Chapter 4), as well as region of interest selection (Chapter 5). End-to-end automation is an important future target.

The end-to-end automated perfusion analysis platform can use the proposed criteria thresholds to discard the tissue voxels. AV detection can be accelerated by clustering a dimensionality-reduced feature space that contains only the effective criteria (i.e. AUC, PeakConc, M_{peak} , and MWO). Then, the proposed kmed-feature segmentation approach could automatically segment the GM, WM, and lesion regions, for which perfusion parameters can be quantified afterwards using the gamma TTD. The only manual component left in this pipeline would be the listing of nominal AIF (Section 3.4.4.5); future investigations should target its automation first. However, this development can only be materialised once the presented methods have gone through rigorous quality control, and satisfactory agreement between manual and automatic methods has been ensured, both in DROs and patients.

In conclusion, this work promotes more automation of DSC-MRI analysis steps, giving priority to simple and rapid approaches. The application of the assistive techniques presented here is not limited to glioblastoma patients only; with minor modifications, diverse patient cohorts can be analysed. Some additional steps such as motion- and PV-correction, and local AIF detection can potentially increase the sophistication and accuracy. Saying that, the presented workflow can be applied readily to any comparative analysis on clinical datasets, with several computational benefits over current methods, such as simpler but more systematically-optimised automation approaches, lower operator bias, faster perfusion quantification, and a shorter delay in diagnosis. Future extensions of this exploratory work can materialise a fully-automated perfusion analysis for clinics; where rapid, accurate, and efficient analysis will permit clinicians to initiate

treatment earlier—and rapidly monitor it thereafter—with minimal delay in processing, diagnosis, and treatment planning.

Appendices

Appendix 1

Table 1: p -value for pairwise comparison between criteria. Note: statistical significance ($p < 0.001$) is highlighted in bold.

Pairwise compared criteria		p value
AUC	FWHM	1.269e-07
AUC	PeakConc	0.99
AUC	FM	1.269e-07
AUC	TTP	1.269e-07
AUC	BAT	1.269e-07
AUC	MTD	1.269e-07
AUC	MWI	0.002
AUC	M_{peak}	0.999
AUC	MWO	0.999
FWHM	PeakConc	1.269e-07
FWHM	FM	0.999
FWHM	TTP	0.999
FWHM	BAT	0.933
FWHM	MTD	0.660
FWHM	MWI	1.269e-07
FWHM	M_{peak}	1.269e-07
FWHM	MWO	1.269e-07
PeakConc	FM	1.269e-07
PeakConc	TTP	1.269e-07
PeakConc	BAT	1.269e-07
PeakConc	MTD	1.269e-07
PeakConc	MWI	0.001
PeakConc	M_{peak}	1.000
PeakConc	MWO	0.999
FM	TTP	0.999
FM	BAT	0.867
FM	MTD	0.773
FM	MWI	1.269e-07
FM	M_{peak}	1.269e-07
FM	MWO	1.269e-07
TTP	BAT	0.982
TTP	MTD	0.487
TTP	MWI	1.269e-07
TTP	M_{peak}	1.269e-07
TTP	MWO	1.269e-07
BAT	MTD	0.053
BAT	MWI	1.269e-07

BAT	M_{peak}	1.269e-07
BAT	MWO	1.269e-07
MTD	MWI	1.269e-07
MTD	M_{peak}	1.269e-07
MTD	MWO	1.269e-07
MWI	M_{peak}	0.001
MWI	MWO	0.002
M_{peak}	MWO	0.999

Abbreviations: AUC, area under the curve; FWHM, full width at half maximum; PeakConc, peak concentration; FM, first moment; TTP, time to peak; BAT, bolus arrival time; MTD, mean time duration; MWI, average wash-in rate; M_{peak} , a combination of multiple criteria (equation 3.8), MWO, average wash-out rate.

Appendix 2

Table 1: Tests of Normality for each Perfusion Parameter, Success Rate, Computation time and root-mean-square error (RMSE).

	Tests of Normality					
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
CBF_GM_sg	.178	9	.200*	.949	9	.674
CBF_GM_gamma	.218	9	.200*	.930	9	.483
CBF_GM_Weibull	.181	9	.200*	.952	9	.713
CBF_GM_GV	.200	9	.200*	.945	9	.631
CBF_WM_sg	.176	9	.200*	.908	9	.300
CBF_WM_gamma	.194	9	.200*	.910	9	.314
CBF_WM_Weibull	.280	9	.040	.831	9	.046
CBF_WM_GV	.192	9	.200*	.852	9	.078
MTT_GM_sg	.177	9	.200*	.918	9	.374
MTT_GM_gamma	.181	9	.200*	.912	9	.327
MTT_GM_Weibull	.191	9	.200*	.924	9	.429
MTT_GM_GV	.180	9	.200*	.898	9	.239
MTT_WM_sg	.173	9	.200*	.979	9	.960
MTT_WM_gamma	.121	9	.200*	.979	9	.957
MTT_WM_Weibull	.148	9	.200*	.950	9	.687
MTT_WM_GV	.128	9	.200*	.985	9	.984
CBV_GM_sg	.174	9	.200*	.953	9	.720
CBV_GM_gamma	.170	9	.200*	.955	9	.747
CBV_GM_weibull	.177	9	.200*	.965	9	.844
CBV_GM_GV	.188	9	.200*	.881	9	.161
CBV_WM_sg	.174	9	.200*	.916	9	.358
CBV_WM_gamma	.185	9	.200*	.863	9	.105
CBV_WM_weibull	.201	9	.200*	.850	9	.074
CBV_WM_GV	.173	9	.200*	.911	9	.321

GM_SG_SUC	.255	35	.000	.806	35	.000
GM_GAMMA_SUC	.220	35	.000	.855	35	.000
GM_WEIBULL_SUC	.159	35	.026	.915	35	.010
GM_GV_SUC	.418	35	.000	.591	35	.000
WM_SG_SUC	.270	35	.000	.714	35	.000
WM_GAMMA_SUC	.214	35	.000	.847	35	.000
WM_WEIBULL_SUC	.206	35	.001	.868	35	.001
WM_GV_SUC	.314	35	.000	.708	35	.000
GM_SG_TIME	.290	35	.000	.468	35	.000
GM_GAMMA_TIME	.276	35	.000	.613	35	.000
GM_WEIBULL_TIME	.294	35	.000	.517	35	.000
GM_GV_TIME	.318	35	.000	.453	35	.000
WM_SG_TIME	.321	35	.000	.643	35	.000
WM_GAMMA_TIME	.305	35	.000	.586	35	.000
WM_WEIBULL_TIME	.122	35	.200*	.921	35	.015
WM_GV_TIME	.318	35	.000	.453	35	.000
RMSE_GM_SG	.107	35	.200*	.962	35	.257
RMSE_GM_GAMMA	.117	35	.200*	.955	35	.158
RMSE_GM_WEIBULL	.114	35	.200*	.966	35	.335
RMSE_GM_GV	.102	35	.200*	.967	35	.374
RMSE_WM_SG	.309	35	.000	.714	35	.000
RMSE_WM_SG	.229	35	.000	.624	35	.000
RMSE_WM_WEIBULL	.276	35	.000	.615	35	.000
RMSE_WM_GV	.212	35	.000	.620	35	.000

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Abbreviations: CBF_GM_sg, CBF value for GM with skewed-Gaussian TTD; CBF_GM_gamma, CBF value for GM with gamma TTD; CBF_GM_Weibull, CBF value for GM with Weibull TTD; CBF_GM_GV, CBF value for GM with gamma-variate TTD; MTT_GM_sg, MTT value for GM with skewed-Gaussian TTD; MTT_GM_gamma, MTT value for GM with gamma TTD; MTT_GM_Weibull, MTT value for GM with Weibull TTD; MTT_GM_GV, MTT value for GM with gamma-variate TTD; CBV_GM_sg, CBV value for GM with skewed-Gaussian TTD; CBV_GM_gamma, CBV value for GM with gamma TTD; CBV_GM_Weibull, CBV value for GM with Weibull TTD; CBV_GM_GV, CBV value for GM with gamma-variate TTD; GM_SG_SUC, Success rate for GM fitting with skewed-Gaussian TTD; GM_GAMMA_SUC, Success rate for GM fitting with gamma TTD; GM_WEIBULL_SUC, Success rate for GM fitting with Weibull TTD; GM_GV_SUC, Success rate for GM fitting with gamma-variate TTD; GM_SG_TIME, Computation time for GM fitting with skewed Gaussian TTD; GM_GAMMA_TIME, Computation time for GM fitting with gamma TTD; GM_WEIBULL_TIME, Computation time for GM fitting with Weibull TTD; GM_GV_TIME, Computation time for GM fitting with gamma-variate TTD; RMSE_GM_SG, RMSE for GM fitting with skewed Gaussian TTD; RMSE_GM_GAMMA, RMSE for GM fitting with gamma TTD; RMSE_GM_WEIBULL, RMSE for GM fitting with Weibull TTD; RMSE_GM_GV, RMSE for GM fitting with gamma-variate TTD.

Note: Corresponding parameters for white matter (WM) are expressed using 'WM' instead of 'GM'.

Table 2a: Mauchly's test of Sphericity for root-mean-square error (RMSE) of grey matter (GM) fitting

Mauchly's Test of Sphericity^a

Measure: RMSE

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Distribution	.004	186.606	5	.000	.361	.363	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: Distribution

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Table 2b: Tests of Within-Subjects effects to assess significant effect of any transit time distribution (TTD) on the root-mean-square error (RMSE) of grey matter (GM) signal fitting

Tests of Within-Subjects Effects

Measure: RMSE

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Distribution	Sphericity Assumed	3.035E-5	3	1.012E-5	2.595	.056
	Greenhouse-Geisser	3.035E-5	1.082	2.806E-5	2.595	.113
	Huynh-Feldt	3.035E-5	1.089	2.787E-5	2.595	.113
	Lower-bound	3.035E-5	1.000	3.035E-5	2.595	.116
Error(Distribution)	Sphericity Assumed	.000	105	3.899E-6		
	Greenhouse-Geisser	.000	37.856	1.081E-5		
	Huynh-Feldt	.000	38.116	1.074E-5		
	Lower-bound	.000	35.000	1.170E-5		

Table 3a: Non-parametric Friedman's test to assess significant effect of any transit time distribution (TTD) on the root-mean-square error (RMSE) of white matter (WM) signal fitting

Test Statistics^a

N	36
Chi-Square	9.733
df	3
Asymp. Sig.	.021

a. Friedman Test

Table 4a: Mauchly's test of Sphericity for grey matter (GM) cerebral blood flow (CBF)

Mauchly's Test of Sphericity^a

Measure: CBF

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Distribution	.071	17.759	5	.004	.438	.490	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: Distribution

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Table 4b: Tests of Within-Subjects effects to assess significant effect of any distribution on grey matter (GM) cerebral blood flow (CBF) value

Tests of Within-Subjects Effects

Measure: CBF

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Distribution	Sphericity Assumed	120.062	3	40.021	2.409	.092
	Greenhouse-Geisser	120.062	1.314	91.387	2.409	.147
	Huynh-Feldt	120.062	1.469	81.715	2.409	.141
	Lower-bound	120.062	1.000	120.062	2.409	.159
Error(Distribution)	Sphericity Assumed	398.733	24	16.614		
	Greenhouse-Geisser	398.733	10.510	37.938		
	Huynh-Feldt	398.733	11.754	33.923		
	Lower-bound	398.733	8.000	49.842		

Table 5a: Mauchly's test of Sphericity for white matter (WM) cerebral blood flow (CBF)

Mauchly's Test of Sphericity^a

Measure: CBF

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Distribution	.032	23.045	5	.000	.399	.429	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: Distribution

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Table 5b: Tests of Within-Subjects effects to assess significant effect of any transit time distribution (TTD) on white matter (WM) cerebral blood flow (CBF) value

Tests of Within-Subjects Effects

Measure: CBF

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Distribution	Sphericity Assumed	40.334	3	13.445	1.617	.212
	Greenhouse-Geisser	40.334	1.196	33.721	1.617	.239
	Huynh-Feldt	40.334	1.288	31.309	1.617	.239
	Lower-bound	40.334	1.000	40.334	1.617	.239
Error(Distribution)	Sphericity Assumed	199.609	24	8.317		
	Greenhouse-Geisser	199.609	9.569	20.860		
	Huynh-Feldt	199.609	10.306	19.368		
	Lower-bound	199.609	8.000	24.951		

Table 6a: Mauchly's test of Sphericity for grey matter (GM) mean transit time (MTT)

Mauchly's Test of Sphericity^a

Measure: MTT

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Distribution	.043	21.109	5	.001	.441	.495	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept

Within Subjects Design: Distribution

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Table 6b: Tests of Within-Subjects effects to assess significant effect of any transit time distribution (TTD) on grey matter (GM) mean transit time (MTT) value

Tests of Within-Subjects Effects

Measure: MTT

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Distribution	Sphericity Assumed	.222	3	.074	1.601	.215
	Greenhouse-Geisser	.222	1.324	.168	1.601	.241
	Huynh-Feldt	.222	1.486	.149	1.601	.239
	Lower-bound	.222	1.000	.222	1.601	.241
Error(Distribution)	Sphericity Assumed	1.109	24	.046		
	Greenhouse-Geisser	1.109	10.595	.105		
	Huynh-Feldt	1.109	11.888	.093		
	Lower-bound	1.109	8.000	.139		

Table 7a: Mauchly's test of Sphericity for white matter (WM) mean transit time (MTT)

Mauchly's Test of Sphericity^a

Measure: MTT

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Distribution	.112	14.723	5	.013	.571	.712	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: Distribution

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Table 7b: Tests of Within-Subjects effects to assess significant effect of any transit time distribution (TTD) on white matter (WM) mean transit time (MTT) value

Tests of Within-Subjects Effects

Measure: MTT

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Distribution	Sphericity Assumed	.643	3	.214	3.864	.022
	Greenhouse-Geisser	.643	1.714	.375	3.864	.052
	Huynh-Feldt	.643	2.136	.301	3.864	.039
	Lower-bound	.643	1.000	.643	3.864	.085
Error(Distribution)	Sphericity Assumed	1.331	24	.055		
	Greenhouse-Geisser	1.331	13.711	.097		
	Huynh-Feldt	1.331	17.085	.078		
	Lower-bound	1.331	8.000	.166		

Table 8a: Mauchly's test of Sphericity for grey matter (GM) cerebral blood volume (CBV)

Mauchly's Test of Sphericity^a

Measure: CBV

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Distribution	.004	37.724	5	.000	.378	.399	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: Distribution

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Table 8b: Tests of Within-Subjects effects to assess significant effect of any transit time distribution (TTD) on grey matter (GM) cerebral blood volume (CBV) value

Tests of Within-Subjects Effects

Measure: CBV

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Distribution	Sphericity Assumed	.001	3	.000	.054	.983
	Greenhouse-Geisser	.001	1.135	.001	.054	.851
	Huynh-Feldt	.001	1.197	.001	.054	.862
	Lower-bound	.001	1.000	.001	.054	.823
Error(Distribution)	Sphericity Assumed	.168	24	.007		
	Greenhouse-Geisser	.168	9.081	.019		
	Huynh-Feldt	.168	9.575	.018		
	Lower-bound	.168	8.000	.021		

Table 9a: Mauchly's test of Sphericity for white matter (WM) cerebral blood volume (CBV)

Mauchly's Test of Sphericity^a

Measure: CBV

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Distribution	.058	19.181	5	.002	.431	.479	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept

Within Subjects Design: Distribution

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Table 9b: Tests of Within-Subjects effects to assess significant effect of any transit time distribution (TTD) on white matter (WM) cerebral blood volume (CBV) value

Tests of Within-Subjects Effects

Measure: CBV

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Distribution	Sphericity Assumed	.002	3	.001	.259	.854
	Greenhouse-Geisser	.002	1.292	.002	.259	.681
	Huynh-Feldt	.002	1.436	.002	.259	.704
	Lower-bound	.002	1.000	.002	.259	.624
Error(Distribution)	Sphericity Assumed	.077	24	.003		
	Greenhouse-Geisser	.077	10.337	.007		
	Huynh-Feldt	.077	11.484	.007		
	Lower-bound	.077	8.000	.010		

Table 10a: Non-parametric Friedman's test to assess significant effect of any transit time distribution (TTD) on the success rate of grey matter (GM) signal fitting

Test Statistics^a

N	36
Chi-Square	54.277
df	3
Asymp. Sig.	.000

a. Friedman Test

Table 10b: Pairwise significance Comparisons between transit time distributions (TTDs) for Success Rate in grey matter (GM) signal fitting.

Test Statistics^a

	GM_GAMMA_ SUC - GM_SG_SUC	GM_WEIBUL L_SUC - GM_SG_SUC	GM_GV_SUC - GM_SG_SUC	GM_WEIBUL L_SUC - GM_GAMMA_ SUC	GM_GV_SUC - GM_GAMMA_ SUC	GM_GV_SUC - GM_WEIBUL L_SUC
Z	-.204 ^b	-5.233 ^c	-4.660 ^c	-5.111 ^c	-4.243 ^c	-1.073 ^c
Asymp. Sig. (2-tailed)	.838	.000	.000	.000	.000	.283

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

Table 11a: Non-parametric Friedman's test to assess significant effect of any transit time distribution (TTD) on the success rate of white matter (WM) signal fitting

Test Statistics^a

N	36
Chi-Square	34.919
df	3
Asymp. Sig.	.000

a. Friedman Test

Table 11b: Pairwise significance Comparisons between transit time distributions (TTDs) for Success Rate in white matter (WM) signal fitting.

Test Statistics^a

	WM_GAMMA_ SUC - WM_SG_SUC	WM_WEIBUL L_SUC - WM_SG_SUC	WM_GV_SUC - WM_SG_SUC	WM_WEIBUL L_SUC - WM_GAMMA_ SUC	WM_GV_SUC - WM_GAMMA_ SUC	WM_GV_SUC - WM_WEIBUL L_SUC
Z	-3.129 ^b	-5.185 ^c	-2.530 ^c	-5.075 ^c	-3.514 ^c	-2.467 ^b
Asymp. Sig. (2-tailed)	.002	.000	.011	.000	.000	.014

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

Table 12a: Non-parametric Friedman’s test to assess significant effect of any transit time distribution (TTD) on the computation time of grey matter (GM) signal fitting

Test Statistics^a

N	36
Chi-Square	39.300
df	3
Asymp. Sig.	.000

a. Friedman Test

Table 12b: Pairwise significance Comparisons between transit time distribution (TTDs) for Computation Time in grey matter (GM) signal fitting.

Test Statistics^a

	GM_GAMMA_ TIME - GM_SG_TIME	GM_WEIBUL L_TIME - GM_SG_TIME	GM_GV_TIME - GM_SG_TIME	GM_WEIBUL L_TIME - GM_GAMMA_ TIME	GM_GV_TIME - GM_GAMMA_ TIME	GM_GV_TIME - GM_WEIBUL L_TIME
Z	-3.158 ^b	-3.959 ^c	-3.299 ^c	-3.221 ^c	-2.891 ^c	-2.639 ^c
Asymp. Sig. (2-tailed)	.002	.000	.001	.001	.004	.008

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

Table 13a: Non-parametric Friedman’s test to assess significant effect of any transit time distribution (TTD) on the computation time of white matter (WM) signal fitting

Test Statistics^a

N	36
Chi-Square	50.367
df	3
Asymp. Sig.	.000

a. Friedman Test

Table 13b: Pairwise significance Comparisons between transit time distribution (TTDs) for Computation Time in white matter (WM) signal fitting.

Test Statistics^a

	WM_GAMMA_ TIME - WM_SG_TIM E	WM_WEIBUL L_TIME - WM_SG_TIM E	WM_GV_TIME - WM_SG_TIM E	WM_WEIBUL L_TIME - WM_GAMMA_ TIME	WM_GV_TIME - WM_GAMMA_ TIME	WM_GV_TIME - WM_WEIBUL L_TIME
Z	-.990 ^b	-3.645 ^c	-4.383 ^c	-3.236 ^c	-4.006 ^c	-3.488 ^c
Asymp. Sig. (2-tailed)	.322	.000	.000	.001	.000	.000

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

References

- ABERNETHY, R., BRENNEMAN, J., MEDLIN, C. & REINMAN, G. 1983. Weibull Analysis Handbook. 243.
- AKAIKE, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716-723.
- AKBARI, H., MACYSZYN, L., DA, X., WOLF, R. L., BILELLO, M., VERMA, R., O'ROURKE, D. M. & DAVATZIKOS, C. 2014. Pattern analysis of dynamic susceptibility contrast-enhanced MR imaging demonstrates peritumoral tissue heterogeneity. *Radiology*, 273, 502-10.
- ANDERSEN, I. K., SZYMKOWIAK, A., RASMUSSEN, C. E., HANSON, L. G., MARSTRAND, J. R., LARSSON, H. B. & HANSEN, L. K. 2002. Perfusion quantification using Gaussian process deconvolution. *Magn Reson Med*, 48, 351-61.
- ARTHUR, D. & VASSILVITSKII, S. 2007. k-means++: the advantages of careful seeding. *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. New Orleans, Louisiana: Society for Industrial and Applied Mathematics.
- ARTZI, M., LIBERMAN, G., NADAV, G., VITINSHTAIN, F., BLUMENTHAL, D. T., BOKSTEIN, F., AIZENSTEIN, O. & BEN BASHAT, D. 2015. Human cerebral blood volume measurements using dynamic contrast enhancement in comparison to dynamic susceptibility contrast MRI. *Neuroradiology*, 57, 671-8.
- ASHBURNER, J. & FRISTON, K. J. 2000. Voxel-based morphometry--the methods. *Neuroimage*, 11, 805-21.
- AXEL, L. 1980. Cerebral blood flow determination by rapid-sequence computed tomography: theoretical analysis. *Radiology*, 137, 679-86.
- BJORNERUD, A. & EMBLEM, K. E. 2010. A fully automated method for quantitative cerebral hemodynamic analysis using DSC-MRI. *J Cereb Blood Flow Metab*, 30, 1066-78.
- BJORNERUD, A., JOHANSSON, L. O., BRILEY-SAEBO, K. & AHLSTROM, H. K. 2002. Assessment of T1 and T2* effects in vivo and ex vivo using iron oxide nanoparticles in steady state--dependence on blood volume and water exchange. *Magn Reson Med*, 47, 461-71.
- BLAND, M. 1995. *An Introduction to Medical Statistics*, Oxford University Press.
- BLEEKER, E. J., VAN OSCH, M. J., CONNELLY, A., VAN BUCHEM, M. A., WEBB, A. G. & CALAMANTE, F. 2011. New criterion to aid manual and automatic selection of the arterial input function in dynamic susceptibility contrast MRI. *Magn Reson Med*, 65, 448-56.
- BOROGOVAC, A. & ASLLANI, I. 2012. Arterial Spin Labeling (ASL) fMRI: Advantages, Theoretical Constrains and Experimental Challenges in Neurosciences. *International Journal of Biomedical Imaging*, 2012, 818456.
- BOXERMAN, J. L., SCHMAINDA, K. M. & WEISSKOFF, R. M. 2006. Relative cerebral blood volume maps corrected for contrast agent extravasation

- significantly correlate with glioma tumor grade, whereas uncorrected maps do not. *AJNR Am J Neuroradiol*, 27, 859-67.
- BUCKLEY, D. L. 2002. Uncertainty in the analysis of tracer kinetics using dynamic contrast-enhanced T1-weighted MRI. *Magn Reson Med*, 47, 601-6.
- CALAMANTE, F. 2012. Perfusion Magnetic Resonance Imaging Quantification in the Brain. In: BADOER, E. (ed.) *Visualization Techniques: From Immunohistochemistry to Magnetic Resonance Imaging*. Totowa, NJ: Humana Press.
- CALAMANTE, F. 2013. Arterial input function in perfusion MRI: a comprehensive review. *Prog Nucl Magn Reson Spectrosc*, 74, 1-32.
- CALAMANTE, F., CHRISTENSEN, S., DESMOND, P. M., OSTERGAARD, L., DAVIS, S. M. & CONNELLY, A. 2010. The physiological significance of the time-to-maximum (Tmax) parameter in perfusion MRI. *Stroke*, 41, 1169-74.
- CALAMANTE, F., CONNELLY, A. & VAN OSCH, M. J. 2009. Nonlinear DeltaR*2 effects in perfusion quantification using bolus-tracking MRI. *Magn Reson Med*, 61, 486-92.
- CALAMANTE, F., GADIAN, D. G. & CONNELLY, A. 2000. Delay and dispersion effects in dynamic susceptibility contrast MRI: simulations using singular value decomposition. *Magn Reson Med*, 44, 466-73.
- CALAMANTE, F., GADIAN, D. G. & CONNELLY, A. 2003. Quantification of bolus-tracking MRI: Improved characterization of the tissue residue function using Tikhonov regularization. *Magn Reson Med*, 50, 1237-47.
- CAMERON, D., BOUHRARA, M., REITER, D. A., FISHBEIN, K. W., CHOI, S., BERGERON, C. M., FERRUCCI, L. & SPENCER, R. G. 2017. The effect of noise and lipid signals on determination of Gaussian and non-Gaussian diffusion parameters in skeletal muscle. *NMR Biomed*, 30.
- CHA, S., JOHNSON, G., WADGHIRI, Y. Z., JIN, O., BABB, J., ZAGZAG, D. & TURNBULL, D. H. 2003. Dynamic, contrast-enhanced perfusion MRI in mouse gliomas: correlation with histopathology. *Magn Reson Med*, 49, 848-55.
- CHA, S., LUPO, J. M., CHEN, M. H., LAMBORN, K. R., MCDERMOTT, M. W., BERGER, M. S., NELSON, S. J. & DILLON, W. P. 2007. Differentiation of glioblastoma multiforme and single brain metastasis by peak height and percentage of signal intensity recovery derived from dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging. *AJNR Am J Neuroradiol*, 28, 1078-84.
- CHA, S., PIERCE, S., KNOPP, E. A., JOHNSON, G., YANG, C., TON, A., LITT, A. W. & ZAGZAG, D. 2001. Dynamic contrast-enhanced T2*-weighted MR imaging of tumefactive demyelinating lesions. *AJNR Am J Neuroradiol*, 22, 1109-16.
- CHOU, Y. C., TENG, M. M., GUO, W. Y., HSIEH, J. C. & WU, Y. T. 2007. Classification of hemodynamics from dynamic-susceptibility-contrast magnetic resonance (DSC-MR) brain images using noiseless independent factor analysis. *Med Image Anal*, 11, 242-53.
- CHRISTENSEN, S., MOURIDSEN, K., WU, O., HJORT, N., KARSTOFT, H., THOMALLA, G., RÖTHER, J., FIEHLER, J., KUCINSKI, T. & ØSTERGAARD, L. 2009. Comparison of 10 Perfusion MRI Parameters in 97 Sub-6-Hour Stroke Patients Using Voxel-Based Receiver Operating Characteristics Analysis. *Stroke*, 40, 2055-2061.
- COLEMAN, G. B. & ANDREWS, H. C. 1979. Image segmentation by clustering. *Proceedings of the IEEE*, 67, 773-785.

- COLEMAN, T. F. & LI, Y. 1996. An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds. *SIAM Journal on Optimization*, 6, 418-445.
- COLLIER, D. C., BURNETT, S. S., AMIN, M., BILTON, S., BROOKS, C., RYAN, A., RONIGER, D., TRAN, D. & STARKSCHALL, G. 2003. Assessment of consistency in contouring of normal-tissue anatomic structures. *J Appl Clin Med Phys*, 4, 17-24.
- DESPOTOVI, I., GOOSSENS, B. & PHILIPS, W. 2015. MRI Segmentation of the Human Brain: Challenges, Methods, and Applications. *Computational and Mathematical Methods in Medicine*, 2015, 23.
- EMBLEM, K. E., NEDREGAARD, B., HALD, J. K., NOME, T., DUE-TONNESSEN, P. & BJORNERUD, A. 2009. Automatic glioma characterization from dynamic susceptibility contrast imaging: brain tumor segmentation using knowledge-based fuzzy clustering. *J Magn Reson Imaging*, 30, 1-10.
- EMBLEM, K. E., NEDREGAARD, B., NOME, T., DUE-TONNESSEN, P., HALD, J. K., SCHEIE, D., BOROTA, O. C., CVANCAROVA, M. & BJORNERUD, A. 2008. Glioma grading by using histogram analysis of blood volume heterogeneity from MR-derived cerebral blood volume maps. *Radiology*, 247, 808-17.
- ESSIG, M., ANZALONE, N., COMBS, S. E., DORFLER, A., LEE, S. K., PICOZZI, P., ROVIRA, A., WELLER, M. & LAW, M. 2012. MR imaging of neoplastic central nervous system lesions: review and recommendations for current practice. *AJNR Am J Neuroradiol*, 33, 803-17.
- ESSIG, M., SHIROISHI, M. S., NGUYEN, T. B., SAAKE, M., PROVENZALE, J. M., ENTERLINE, D., ANZALONE, N., DÖRFLER, A., ROVIRA, A., WINTERMARK, M. & LAW, M. 2013. Perfusion MRI: the five most frequently asked technical questions. *AJR Am J Roentgenol*, 200, 24-34.
- EUGENE, C. L. & ABASS, A. 2019. *PET and PET/CT : A Clinical Guide*, New York, Thieme.
- EVERITT, B. S., LANDAU, S., LEESE, M. & STAHL, D. 2011. *Cluster Analysis*, Wiley.
- FIELD, A. 2013. *Discovering Statistics using IBM SPSS Statistics*, Sage Publications Ltd.
- FUBINI, G. 1907. *Sugli integrali multipli: nota*, Tipografia della R. Accademia dei Lincei.
- GHASEMI, A. & ZAHEDIASL, S. 2012. Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism*, 10, 486-489.
- GHODASARA, S., CHEN, Y., PAHWA, S., GRISWOLD, M. A., SEIBERLICH, N., WRIGHT, K. L. & GULANI, V. 2020. Quantifying Perfusion Properties with DCE-MRI Using a Dictionary Matching Approach. *Scientific Reports*, 10, 10210.
- GOBBEL, G. T. & FIKE, J. R. 1994. A deconvolution method for evaluating indicator-dilution curves. *Physics in Medicine and Biology*, 39, 1833-1854.
- GORDON, Y., PARTOVI, S., MÜLLER-ESCHNER, M., AMARTEIFIO, E., BÄUERLE, T., WEBER, M. A., KAUCZOR, H. U. & RENGIER, F. 2014. Dynamic contrast-enhanced magnetic resonance imaging: fundamentals and application to the evaluation of the peripheral perfusion. *Cardiovasc Diagn Ther*, 4, 147-64.
- GRANDIN, C. B., BOL, A., SMITH, A. M., MICHEL, C. & COSNARD, G. 2005. Absolute CBF and CBV measurements by MRI bolus tracking before and after acetazolamide challenge: repeatability and comparison with PET in humans. *Neuroimage*, 26, 525-35.

- GUIJO-RUBIO, D., DURÁN-ROSAL, A., GUTIÉRREZ, P. A., TRONCOSO, A. & MARTÍNEZ, C. 2018. *Time series clustering based on the characterisation of segment typologies*.
- GUIJO-RUBIO, D., DURAN-ROSAL, A. M., GUTIERREZ, P. A., TRONCOSO, A. & HERVAS-MARTINEZ, C. 2020. Time-Series Clustering Based on the Characterization of Segment Typologies. *IEEE Trans Cybern*.
- GUZMAN-DE-VILLORIA, J. A., FERNANDEZ-GARCIA, P., MATEOS-PEREZ, J. M. & DESCO, M. 2012. [Studying cerebral perfusion using magnetic susceptibility techniques: technique and applications]. *Radiologia*, 54, 208-20.
- HAAR, H. J. V. D., BURGMANS, S., JANSEN, J. F. A., OSCH, M. J. P. V., BUCHEM, M. A. V., MULLER, M., HOFMAN, P. A. M., VERHEY, F. R. J. & BACKES, W. H. 2016. Blood-Brain Barrier Leakage in Patients with Early Alzheimer Disease. *Radiology*, 281, 527-535.
- HADJIPROCOPIIS, A., RASHID, W. & TOFTS, P. S. 2005. Unbiased segmentation of diffusion-weighted magnetic resonance images of the brain using iterative clustering. *Magn Reson Imaging*, 23, 877-85.
- HAKYEMEZ, B., ERDOGAN, C., ERCAN, I., ERGIN, N., UYSAL, S. & ATAHAN, S. 2005. High-grade and low-grade gliomas: differentiation by using perfusion MR imaging. *Clin Radiol*, 60, 493-502.
- HARTMANN, M., HEILAND, S., HARTING, I., TRONNIER, V. M., SOMMER, C., LUDWIG, R. & SARTOR, K. 2003. Distinguishing of primary cerebral lymphoma from high-grade glioma with perfusion-weighted magnetic resonance imaging. *Neurosci Lett*, 338, 119-22.
- HELENIUS, J., PERKIÖ, J., SOINNE, L., ØSTERGAARD, L., CARANO, R. A., SALONEN, O., SAVOLAINEN, S., KASTE, M., ARONEN, H. J. & TATLISUMAK, T. 2003. Cerebral hemodynamics in a healthy population measured by dynamic susceptibility contrast MR imaging. *Acta Radiol*, 44, 538-46.
- HEYE, A. K., THRIFFLETON, M. J., ARMITAGE, P. A., VALDÉS HERNÁNDEZ, M. D. C., MAKIN, S. D., GLATZ, A., SAKKA, E. & WARDLAW, J. M. 2016. Tracer kinetic modelling for DCE-MRI quantification of subtle blood-brain barrier permeability. *Neuroimage*, 125, 446-455.
- HOFFMANN, U., BRIX, G., KNOPP, M. V., HEB, T. & LORENZ, W. J. 1995. Pharmacokinetic Mapping of the Breast: A New Method for Dynamic MR Mammography. *Magnetic Resonance in Medicine*, 33, 506-514.
- IBARAKI, M., ITO, H., SHIMOSEGAWA, E., TOYOSHIMA, H., ISHIGAME, K., TAKAHASHI, K., KANNO, I. & MIURA, S. 2007. Cerebral vascular mean transit time in healthy humans: A comparative study with PET and dynamic susceptibility contrast-enhanced MRI. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, 27, 404-13.
- IBARAKI, M., SHIMOSEGAWA, E., TOYOSHIMA, H., TAKAHASHI, K., MIURA, S. & KANNO, I. 2005. Tracer delay correction of cerebral blood flow with dynamic susceptibility contrast-enhanced MRI. *J Cereb Blood Flow Metab*, 25, 378-90.
- JACKSON, A., KASSNER, A., ANNESLEY-WILLIAMS, D., REID, H., ZHU, X. P. & LI, K. L. 2002. Abnormalities in the recirculation phase of contrast agent bolus passage in cerebral gliomas: comparison with relative blood volume and tumor grade. *AJNR Am J Neuroradiol*, 23, 7-14.

- JAHNG, G. H., LI, K. L., OSTERGAARD, L. & CALAMANTE, F. 2014. Perfusion magnetic resonance imaging: a comprehensive update on principles and techniques. *Korean J Radiol*, 15, 554-77.
- JENKINSON, M., BECKMANN, C. F., BEHRENS, T. E., WOOLRICH, M. W. & SMITH, S. M. 2012. FSL. *Neuroimage*, 62, 782-90.
- JIANBO, S. & MALIK, J. 2000. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, 888-905.
- JOHNSON, G., WETZEL, S. G., CHA, S., BABB, J. & TOFTS, P. S. 2004. Measuring blood volume and vascular transfer constant from dynamic, T(2)*-weighted contrast-enhanced MRI. *Magn Reson Med*, 51, 961-8.
- KANE, I., CARPENTER, T., CHAPPELL, F., RIVERS, C., ARMITAGE, P., SANDERCOCK, P. & WARDLAW, J. 2007. Comparison of 10 different magnetic resonance perfusion imaging processing methods in acute ischemic stroke: effect on lesion size, proportion of patients with diffusion/perfusion mismatch, clinical scores, and radiologic outcomes. *Stroke*, 38, 3158-64.
- KANUNGO, T., MOUNT, D. M., NETANYAHU, N. S., PIATKO, C. D., SILVERMAN, R. & WU, A. Y. 2002. An Efficient k-Means Clustering Algorithm: Analysis and Implementation %J IEEE Trans. Pattern Anal. Mach. Intell. 24, 881-892.
- KAO, Y. H., TENG, M. M., ZHENG, W. Y., CHANG, F. C. & CHEN, Y. F. 2010. Removal of CSF pixels on brain MR perfusion images using first several images and Otsu's thresholding technique. *Magn Reson Med*, 64, 743-8.
- KASS, M., WITKIN, A. & TERZOPOULOS, D. 1988. Snakes: Active contour models. *International Journal of Computer Vision*, 1, 321-331.
- KAUFMAN, L. & ROUSSEEUW, P. J. 2005. *Finding Groups in Data : An Introduction to Cluster Analysis*, New York, UNITED STATES, John Wiley & Sons, Incorporated.
- KAYA, I. E., PEHLIVANLI, A. Ç., SEKIZKARDEŞ, E. G. & IBRIKCI, T. 2017. PCA based clustering for brain tumor segmentation of T1w MRI images. *Computer Methods and Programs in Biomedicine*, 140, 19-28.
- KNUTSSON, L., STAHLBERG, F. & WIRESTAM, R. 2010. Absolute quantification of perfusion using dynamic susceptibility contrast MRI: pitfalls and possibilities. *Magma*, 23, 1-21.
- KOH, T. S., BISDAS, S., KOH, D. M. & THNG, C. H. 2011. Fundamentals of tracer kinetics for dynamic contrast-enhanced MRI. *J Magn Reson Imaging*, 34, 1262-76.
- KOH, T. S., ZEMAN, V., DARKO, J., LEE, T. Y., MILOSEVIC, M. F., HAIDER, M., WARDE, P. & YEUNG, I. W. 2001. The inclusion of capillary distribution in the adiabatic tissue homogeneity model of blood flow. *Phys Med Biol*, 46, 1519-38.
- KREMER, S., GRAND, S., BERGER, F., HOFFMANN, D., PASQUIER, B., REMY, C., BENABID, A. L. & BAS, J. F. 2003. Dynamic contrast-enhanced MRI: differentiating melanoma and renal carcinoma metastases from high-grade astrocytomas and other metastases. *Neuroradiology*, 45, 44-9.
- KRZANOWSKI, W. J. & HAND, D. J. 2009. *ROC Curves for Continuous Data*, Chapman & Hall/CRC.
- KUHN, M. J., PICOZZI, P., MALDJIAN, J. A., SCHMALFUSS, I. M., MARAVILLA, K. R., BOWEN, B. C., WIPPOLD, F. J., 2ND, RUNGE, V. M., KNOPP, M. V., WOLANSKY, L. J., GUSTAFSSON, L., ESSIG, M. & ANZALONE, N. 2007. Evaluation of intraaxial enhancing brain tumors on magnetic resonance imaging: intraindividual crossover comparison of gadobenate dimeglumine and

- gadopentetate dimeglumine for visualization and assessment, and implications for surgical intervention. *J Neurosurg*, 106, 557-66.
- L BOXERMAN, J., SCHMAINDA, K. & M WEISSKOFF, R. 2006. Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not. *AJNR. American journal of neuroradiology*, 27, 859-67.
- LAI, C.-D., MURTHY, D. N. & XIE, M. 2006. Weibull Distributions and Their Applications. In: PHAM, H. (ed.) *Springer Handbook of Engineering Statistics*. London: Springer London.
- LARSSON, H. B., HANSEN, A. E., BERG, H. K., ROSTRUP, E. & HARALDSETH, O. 2008. Dynamic contrast-enhanced quantitative perfusion measurement of the brain using T1-weighted MRI at 3T. *J Magn Reson Imaging*, 27, 754-62.
- LARSSON, H. B. W., VESTERGAARD, M. B., LINDBERG, U., IVERSEN, H. K. & CRAMER, S. P. 2017. Brain capillary transit time heterogeneity in healthy volunteers measured by dynamic contrast-enhanced T(1) -weighted perfusion MRI. *Journal of magnetic resonance imaging : JMRI*, 45, 1809-1820.
- LAW, M., OH, S., BABB, J. S., WANG, E., INGLESE, M., ZAGZAG, D., KNOPP, E. A. & JOHNSON, G. 2006. Low-grade gliomas: dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging--prediction of patient clinical response. *Radiology*, 238, 658-67.
- LAW, M., YANG, S., BABB, J. S., KNOPP, E. A., GOLFINOS, J. G., ZAGZAG, D. & JOHNSON, G. 2004. Comparison of cerebral blood volume and vascular permeability from dynamic susceptibility contrast-enhanced perfusion MR imaging with glioma grade. *AJNR Am J Neuroradiol*, 25, 746-55.
- LAW, M., YANG, S., WANG, H., BABB, J. S., JOHNSON, G., CHA, S., KNOPP, E. A. & ZAGZAG, D. 2003. Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. *AJNR Am J Neuroradiol*, 24, 1989-98.
- LEU, K., BOXERMAN, J. L., CLOUGHESY, T. F., LAI, A., NGHIEMPHU, P. L., LIAU, L. M., POPE, W. B. & ELLINGSON, B. M. 2016. Improved Leakage Correction for Single-Echo Dynamic Susceptibility Contrast Perfusion MRI Estimates of Relative Cerebral Blood Volume in High-Grade Gliomas by Accounting for Bidirectional Contrast Agent Exchange. *AJNR Am J Neuroradiol*, 37, 1440-6.
- LIU, H. L., PU, Y., LIU, Y., NICKERSON, L., ANDREWS, T., FOX, P. T. & GAO, J. H. 1999. Cerebral blood flow measurement by dynamic contrast MRI using singular value decomposition with an adaptive threshold. *Magn Reson Med*, 42, 167-72.
- LLOYD, S. 1982. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28, 129-137.
- MADSEN, M. T. 1992. A simplified formulation of the gamma variate function. *Physics in Medicine & Biology*, 37, 1597.
- MANGLA, R., KOLAR, B., ZHU, T., ZHONG, J., ALMAST, J. & EKHOLM, S. 2011. Percentage Signal Recovery Derived from MR Dynamic Susceptibility Contrast Imaging Is Useful to Differentiate Common Enhancing Malignant Lesions of the Brain. *American Journal of Neuroradiology*, 32, 1004-1010.
- MCDONALD, J. H. & DELAWARE, U. O. 2009. *Handbook of Biological Statistics*, Sparky House Publishing.
- MEHNDIRATTA, A., MACINTOSH, B. J., CRANE, D. E., PAYNE, S. J. & CHAPPELL, M. A. 2013. A control point interpolation method for the non-

- parametric quantification of cerebral haemodynamics from dynamic susceptibility contrast MRI. *Neuroimage*, 64, 560-70.
- MEIER, P. & ZIERLER, K. L. 1954. On the theory of the indicator-dilution method for measurement of blood flow and volume. *J Appl Physiol*, 6, 731-44.
- MILES, K. A. 2004. Brain perfusion: computed tomography applications. *Neuroradiology*, 46 Suppl 2, s194-200.
- MOHAJER, M., BRIX, G. & ENGLMEIER, K.-H. 2010. *A novel and fast method for cluster analysis of DCE-MR image series of breast tumors*, SPIE.
- MOHAMAD, I. & USMAN, D. 2013. *Standardization and Its Effects on K-Means Clustering Algorithm*.
- MOUANNES-SROUR, J. J., SHIN, W., ANSARI, S. A., HURLEY, M. C., VAKIL, P., BENDOK, B. R., LEE, J. L., DERDEYN, C. P. & CARROLL, T. J. 2012. Correction for arterial-tissue delay and dispersion in absolute quantitative cerebral perfusion DSC MR imaging. *Magn Reson Med*, 68, 495-506.
- MOURIDSEN, K., CHRISTENSEN, S., GYLDENSTED, L. & OSTERGAARD, L. 2006a. Automatic selection of arterial input function using cluster analysis. *Magn Reson Med*, 55, 524-31.
- MOURIDSEN, K., FRISTON, K., HJORT, N., GYLDENSTED, L., OSTERGAARD, L. & KIEBEL, S. 2006b. Bayesian estimation of cerebral perfusion using a physiological model of microvasculature. *Neuroimage*, 33, 570-9.
- MOURIDSEN, K., HANSEN, M. B., OSTERGAARD, L. & JESPERSEN, S. N. 2014. Reliable estimation of capillary transit time distributions using DSC-MRI. *J Cereb Blood Flow Metab*, 34, 1511-21.
- MUNDIYANAPURATH, S., RINGLEB, P. A., DIATSCHEK, S., HANSEN, M. B., MOURIDSEN, K., ØSTERGAARD, L., WICK, W., BENDSZUS, M. & RADBRUCH, A. 2016. Capillary Transit Time Heterogeneity Is Associated with Modified Rankin Scale Score at Discharge in Patients with Bilateral High Grade Internal Carotid Artery Stenosis. *PLOS ONE*, 11, e0158148.
- MURASE, K., KIKUCHI, K., MIKI, H., SHIMIZU, T. & IKEZOE, J. 2001a. Determination of arterial input function using fuzzy clustering for quantification of cerebral blood flow with dynamic susceptibility contrast-enhanced MR imaging. *Journal of Magnetic Resonance Imaging*, 13, 797-806.
- MURASE, K., SHINOHARA, M. & YAMAZAKI, Y. 2001b. Accuracy of deconvolution analysis based on singular value decomposition for quantification of cerebral blood flow using dynamic susceptibility contrast-enhanced magnetic resonance imaging. *Physics in Medicine and Biology*, 46, 3147-3159.
- NAGESH, V., CHENEVERT, T. L., TSIEN, C. I., ROSS, B. D., LAWRENCE, T. S., JUNCK, L. & CAO, Y. 2007. Quantitative characterization of hemodynamic properties and vasculature dysfunction of high-grade gliomas. *NMR Biomed*, 20, 566-77.
- NEWTON, A. T., PRUTHI, S., STOKES, A. M., SKINNER, J. T. & QUARLES, C. C. 2016. Improving Perfusion Measurement in DSC-MR Imaging with Multiecho Information for Arterial Input Function Determination. *AJNR Am J Neuroradiol*, 37, 1237-43.
- O'CONNOR, J. P., TOFTS, P. S., MILES, K. A., PARKES, L. M., THOMPSON, G. & JACKSON, A. 2011. Dynamic contrast-enhanced imaging techniques: CT and MRI. *Br J Radiol*, 84 Spec No 2, S112-20.
- ORTIZ, A., GORRIZ, J. M., RAMIREZ, J. & SALAS-GONZALEZ, D. 2014. Improving MR brain image segmentation using self-organising maps and entropy-gradient clustering. *Information Sciences*, 262, 117-136.

- ØSTERGAARD, L. 2005. Principles of cerebral perfusion imaging by bolus tracking. *Journal of Magnetic Resonance Imaging*, 22, 710-717.
- OSTERGAARD, L., CHESLER, D. A., WEISSKOFF, R. M., SORENSEN, A. G. & ROSEN, B. R. 1999. Modeling cerebral blood flow and flow heterogeneity from magnetic resonance residue data. *J Cereb Blood Flow Metab*, 19, 690-9.
- OSTERGAARD, L., SMITH, D. F., VESTERGAARD-POULSEN, P., HANSEN, S. B., GEE, A. D., GJEDDE, A. & GYLDENSTED, C. 1998. Absolute cerebral blood flow and blood volume measured by magnetic resonance imaging bolus tracking: comparison with positron emission tomography values. *J Cereb Blood Flow Metab*, 18, 425-32.
- OSTERGAARD, L., WEISSKOFF, R. M., CHESLER, D. A., GYLDENSTED, C. & ROSEN, B. R. 1996. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis. *Magn Reson Med*, 36, 715-25.
- OTSU, N. 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9, 62-66.
- PARK, H.-S. & JUN, C.-H. 2009. A simple and fast algorithm for K-medoids clustering. *Expert Systems with Applications*, 36, 3336-3341.
- PATIL, V., JENSEN, J. H. & JOHNSON, G. 2013. Intravascular contrast agent T2* relaxivity in brain tissue. *NMR Biomed*, 26, 392-9.
- PATIL, V. & JOHNSON, G. 2011. An improved model for describing the contrast bolus in perfusion MRI. *Med Phys*, 38, 6380-3.
- PATIL, V. & JOHNSON, G. 2013. DeltaR2 (*) gadolinium-diethylenetriaminepentacetic acid relaxivity in venous blood. *Magn Reson Med*, 69, 1104-8.
- PENNY, W. D., STEPHAN, K. E., MECHELLI, A. & FRISTON, K. J. 2004. Comparing dynamic causal models. *Neuroimage*, 22, 1157-72.
- PERL, W., LASSEN, N. A. & EFFROS, R. M. 1975. Matrix proof of flow, volume and mean transit time theorems for regional and compartmental systems. *Bull Math Biol*, 37, 573-88.
- PERTHEN, J. E., CALAMANTE, F., GADIAN, D. G. & CONNELLY, A. 2002. Is quantification of bolus tracking MRI reliable without deconvolution? *Magn Reson Med*, 47, 61-7.
- PERUZZO, D., BERTOLDO, A., ZANDERIGO, F. & COBELLI, C. 2011. Automatic selection of arterial input function on dynamic contrast-enhanced MR images. *Comput Methods Programs Biomed*, 104, e148-57.
- RASYID, L. & ANDAYANI, S. 2018. Review on Clustering Algorithms Based on Data Type: Towards the Method for Data Combined of Numeric-Fuzzy Linguistics. *Journal of Physics: Conference Series*, 1097, 012082.
- RAUSCH, M., SCHEFFLER, K., RUDIN, M. & RADÜ, E. W. 2000. Analysis of input functions from different arterial branches with gamma variate functions and cluster analysis for quantitative blood volume measurements. *Magnetic Resonance Imaging*, 18, 1235-1243.
- RAYKOV, Y. P., BOUKOUVALAS, A., BAIG, F. & LITTLE, M. A. 2016. What to Do When K-Means Clustering Fails: A Simple yet Principled Alternative Algorithm. *PLOS ONE*, 11, e0162259.
- REMPP, K. A., BRIX, G., WENZ, F., BECKER, C. R., GUCKEL, F. & LORENZ, W. J. 1994. Quantification of regional cerebral blood flow and volume with dynamic susceptibility contrast-enhanced MR imaging. *Radiology*, 193, 637-41.
- REVENAZ, A., RUGGERI, M., LAGANÀ, M., BERGSLAND, N., GROPPA, E., ROVARIS, M. & FAINARDI, E. 2016. A semi-automated measuring system of

- brain diffusion and perfusion magnetic resonance imaging abnormalities in patients with multiple sclerosis based on the integration of coregistration and tissue segmentation procedures. *BMC medical imaging*, 16, 4-4.
- ROSEN, B. R., BELLIVEAU, J. W., VEVEA, J. M. & BRADY, T. J. 1990. Perfusion imaging with NMR contrast agents. *Magn Reson Med*, 14, 249-65.
- ROWLEY, H. A., SCIALFA, G., GAO, P. Y., MALDJIAN, J. A., HASSELL, D., KUHN, M. J., WIPPOLD, F. J., 2ND, GALLUCCI, M., BOWEN, B. C., SCHMALFUSS, I. M., RUSCALLEDA, J., BASTIANELLO, S. & COLOSIMO, C. 2008. Contrast-enhanced MR imaging of brain lesions: a large-scale intraindividual crossover comparison of gadobenate dimeglumine versus gadodiamide. *AJNR Am J Neuroradiol*, 29, 1684-91.
- SCHABEL, M. C. 2012. A unified impulse response model for DCE-MRI. *Magn Reson Med*, 68, 1632-46.
- SCHWARZ, G. 1978. Estimating the Dimension of a Model. *The Annals of Statistics*, 6, 461-464, 4.
- SEMMINEH, N. B., STOKES, A. M., BELL, L. C., BOXERMAN, J. L. & QUARLES, C. C. 2017. A Population-Based Digital Reference Object (DRO) for Optimizing Dynamic Susceptibility Contrast (DSC)-MRI Methods for Clinical Trials. *Tomography*, 3, 41-49.
- SEMMINEH, N. B., XU, J., BOXERMAN, J. L., DELANEY, G. W., CLEARY, P. W., GORE, J. C. & QUARLES, C. C. 2014. An Efficient Computational Approach to Characterize DSC-MRI Signals Arising from Three-Dimensional Heterogeneous Tissue Structures. *PLOS ONE*, 9, e84764.
- SINGH, A., HARIS, M., RATHORE, D., PURWAR, A., SARMA, M., BAYU, G., HUSAIN, N., RATHORE, R. K. & GUPTA, R. K. 2007. Quantification of physiological and hemodynamic indices using T(1) dynamic contrast-enhanced MRI in intracranial mass lesions. *J Magn Reson Imaging*, 26, 871-80.
- SINGH, V., RANA, R. K. & SINGHAL, R. 2013. Analysis of repeated measurement data in the clinical trials. *J Ayurveda Integr Med*, 4, 77-81.
- SKINNER, J. T., MOOTS, P. L., AYERS, G. D. & QUARLES, C. C. 2016a. On the Use of DSC-MRI for Measuring Vascular Permeability. *AJNR. American journal of neuroradiology*, 37, 80-87.
- SKINNER, J. T., MOOTS, P. L., AYERS, G. D. & QUARLES, C. C. 2016b. On the Use of DSC-MRI for Measuring Vascular Permeability. *AJNR Am J Neuroradiol*, 37, 80-7.
- SOURBRON, S., INGRISCH, M., SIEFERT, A., REISER, M. & HERRMANN, K. 2009. Quantification of cerebral blood flow, cerebral blood volume, and blood-brain-barrier leakage with DCE-MRI. *Magn Reson Med*, 62, 205-17.
- SOURBRON, S., LUYPAERT, R., VAN SCHUERBEEK, P., DUJARDIN, M., STADNIK, T. & OSTEAX, M. 2004. Deconvolution of dynamic contrast-enhanced MRI data by linear inversion: choice of the regularization parameter. *Magn Reson Med*, 52, 209-13.
- SOURBRON, S. P. & BUCKLEY, D. L. 2012. Tracer kinetic modelling in MRI: estimating perfusion and capillary permeability. *Phys Med Biol*, 57, R1-33.
- SOURBRON, S. P. & BUCKLEY, D. L. 2013. Classic models for dynamic contrast-enhanced MRI. *NMR Biomed*, 26, 1004-27.
- ST LAWRENCE, K. S. & LEE, T. Y. 1998. An adiabatic approximation to the tissue homogeneity model for water exchange in the brain: I. Theoretical derivation. *J Cereb Blood Flow Metab*, 18, 1365-77.

- SUETENS, P. 2009. *Fundamentals of Medical Imaging*, Cambridge, Cambridge University Press.
- SUGAHARA, T., KOROJI, Y., KOCHI, M., IKUSHIMA, I., HIRAI, T., OKUDA, T., SHIGEMATSU, Y., LIANG, L., GE, Y., USHIO, Y. & TAKAHASHI, M. 1998. Correlation of MR imaging-determined cerebral blood volume maps with histologic and angiographic determination of vascularity of gliomas. *AJR Am J Roentgenol*, 171, 1479-86.
- TABBARA, R., CONNELLY, A. & CALAMANTE, F. 2020. Multi-stage automated local arterial input function selection in perfusion MRI. *Magma*, 33, 357-365.
- TAXT, T., LUNDERVOLD, A., FUGLAAS, B., LIEN, H. & ABELER, V. 1992. Multispectral analysis of uterine corpus tumors in magnetic resonance imaging. *Magn Reson Med*, 23, 55-76.
- THERASSE, P., ARBUCK, S. G., EISENHAEUER, E. A., WANDERS, J., KAPLAN, R. S., RUBINSTEIN, L., VERWEIJ, J., VAN GLABBEKE, M., VAN OOSTEROM, A. T., CHRISTIAN, M. C. & GWYTHYER, S. G. 2000. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst*, 92, 205-16.
- TOFTS, P. S. 1997. Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *J Magn Reson Imaging*, 7, 91-101.
- TOFTS, P. S., BRIK, G., BUCKLEY, D. L., EVELHOCH, J. L., HENDERSON, E., KNOPP, M. V., LARSSON, H. B., LEE, T. Y., MAYR, N. A., PARKER, G. J., PORT, R. E., TAYLOR, J. & WEISSKOFF, R. M. 1999. Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusible tracer: standardized quantities and symbols. *J Magn Reson Imaging*, 10, 223-32.
- TOFTS, P. S. & KERMODE, A. G. 1991. Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts. *Magn Reson Med*, 17, 357-67.
- TZIKA, A. A., ASTRAKAS, L. G., ZARIFI, M. K., PETRIDOU, N., YOUNG-POUSSAINT, T., GOUMNEROVA, L., ZURAKOWSKI, D., ANTHONY, D. C. & BLACK, P. M. 2003. Multiparametric MR assessment of pediatric brain tumors. *Neuroradiology*, 45, 1-10.
- VAN OSCH, M. J., VONKEN, E. J., VIERGEVER, M. A., VAN DER GROND, J. & BAKKER, C. J. 2003. Measuring the arterial input function with gradient echo sequences. *Magn Reson Med*, 49, 1067-76.
- VANSTEENKISTE, E. 2007. *Quantitative Analysis of Ultrasound Images of the Preterm Brain*.
- VELMURUGAN 2010. Computational Complexity between K-Means and K-Medoids Clustering Algorithms for Normal and Uniform Distributions of Data Points. *Journal of Computer Science*, 6, 363-368.
- VONKEN, E.-J. P. A., VAN OSCH, M. J. P., BAKKER, C. J. G. & VIERGEVER, M. A. 2000. Simultaneous quantitative cerebral perfusion and Gd-DTPA extravasation measurement with dual-echo dynamic susceptibility contrast MRI. 43, 820-827.
- WANG, X., SMITH-MILES, K. & HYNDMAN, R. 2006a. *Characteristic-Based Clustering for Time Series Data*.
- WANG, X., SMITH, K. & HYNDMAN, R. 2006b. Characteristic-Based Clustering for Time Series Data. *Data Mining and Knowledge Discovery*, 13, 335-364.

- WARDLAW, J. M., MAKIN, S. J., VALDÉS HERNÁNDEZ, M. C., ARMITAGE, P. A., HEYE, A. K., CHAPPELL, F. M., MUÑOZ-MANIEGA, S., SAKKA, E., SHULER, K., DENNIS, M. S. & THRIFFLETON, M. J. 2017. Blood-brain barrier failure as a core mechanism in cerebral small vessel disease and dementia: evidence from a cohort study. *Alzheimer's & Dementia*, 13, 634-643.
- WELKER, K., BOXERMAN, J., KALNIN, A., KAUFMANN, T., SHIROISHI, M. & WINTERMARK, M. 2015. ASFN Recommendations for Clinical Performance of MR Dynamic Susceptibility Contrast Perfusion Imaging of the Brain. *American Journal of Neuroradiology*, 36, E41-E51.
- WILLATS, L., CHRISTENSEN, S., MA, H. K., DONNAN, G. A., CONNELLY, A. & CALAMANTE, F. 2011. Validating a local Arterial Input Function method for improved perfusion quantification in stroke. *J Cereb Blood Flow Metab*, 31, 2189-98.
- WILLIAMS, D. S., DETRE, J. A., LEIGH, J. S. & KORETSKY, A. P. 1992. Magnetic resonance imaging of perfusion using spin inversion of arterial water. *Proc Natl Acad Sci U S A*, 89, 212-6.
- WIRESTAM, R. 2012. Using contrast agents to obtain maps of regional perfusion and capillary wall permeability. *Imaging in Medicine*, 4, 423-442.
- WISMULLER, A., MEYER-BAESE, A., LANGE, O., REISER, M. F. & LEINSINGER, G. 2006. Cluster analysis of dynamic cerebral contrast-enhanced perfusion MRI time-series. *IEEE Trans Med Imaging*, 25, 62-73.
- WONG, E. T., JACKSON, E. F., HESS, K. R., SCHOMER, D. F., HAZLE, J. D., KYRITSIS, A. P., JAECKLE, K. A., YUNG, W. K., LEVIN, V. A. & LEEDS, N. E. 1998. Correlation between dynamic MRI and outcome in patients with malignant gliomas. *Neurology*, 50, 777-81.
- WU, O., ØSTERGAARD, L., KOROSHETZ, W. J., SCHWAMM, L. H., O'DONNELL, J., SCHAEFER, P. W., ROSEN, B. R., WEISSKOFF, R. M. & SORENSEN, A. G. 2003a. Effects of tracer arrival time on flow estimates in MR perfusion-weighted imaging. *Magn Reson Med*, 50, 856-64.
- WU, O., ØSTERGAARD, L., M WEISSKOFF, R., BENNER, T., ROSEN, B. & SORENSEN, A. 2003b. Tracer arrival timing-insensitive technique for estimating flow in MR perfusion-weighted imaging using singular value decomposition with a block-circulant deconvolution matrix. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 50, 164-74.
- XUE, J.-H., PIZURICA, A., PHILIPS, W., KERRE, E., VAN DE WALLE, R. & LEMAHIEU, I. 2003. An integrated method of adaptive enhancement for unsupervised segmentation of MRI brain images. *Pattern Recognition Letters*, 24, 2549-2560.
- YABLONSKIY, D. A. & HAACKE, E. M. 1994. Theory of NMR signal behavior in magnetically inhomogeneous tissues: the static dephasing regime. *Magn Reson Med*, 32, 749-63.
- YIN, J., SUN, H., YANG, J. & GUO, Q. 2015. Automated detection of the arterial input function using normalized cut clustering to determine cerebral perfusion by dynamic susceptibility contrast-magnetic resonance imaging. *J Magn Reson Imaging*, 41, 1071-8.
- YIN, J., YANG, J. & GUO, Q. 2014. Evaluating the feasibility of an agglomerative hierarchy clustering algorithm for the automatic detection of the arterial input function using DSC-MRI. *PLoS One*, 9, e100308.

- ZANDERIGO, F., BERTOLDO, A., PILLONETTO, G. & COBELLI AST, C. 2009. Nonlinear stochastic regularization to characterize tissue residue function in bolus-tracking MRI: assessment and comparison with SVD, block-circulant SVD, and Tikhonov. *IEEE Trans Biomed Eng*, 56, 1287-97.
- ZARO-WEBER, O., MOELLER-HARTMANN, W., HEISS, W.-D. & SOBESKY, J. 2012. Influence of the Arterial Input Function on Absolute and Relative Perfusion-Weighted Imaging Penumbra Flow Detection. *Stroke*, 43, 378-385.
- ZHU, X. P., LI, K. L. & JACKSON, A. 2005. Dynamic Contrast-Enhanced MRI in Cerebral Tumours. In: JACKSON, A., BUCKLEY, D. L. & PARKER, G. J. M. (eds.) *Dynamic Contrast-Enhanced Magnetic Resonance Imaging in Oncology*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- ZOU, K. H., WARFIELD, S. K., BHARATHA, A., TEMPANY, C. M. C., KAUS, M. R., HAKER, S. J., WELLS, W. M., JOLESZ, F. A. & KIKINIS, R. 2004. Statistical validation of image segmentation quality based on a spatial overlap index: scientific reports. *Academic Radiology*, 11, 178-189.