The Role of Arterial Spin Labelling (ASL) in Classification of Primary Adult Gliomas

Amirah Faisal S. Alsaedi

Institute of Neurology (ION)

University College London (UCL)

Doctor of Philosophy (PhD)

Declaration

I, Amirah Faisal S. Alsaedi, confirm that I carried out the work presented in this thesis for the degree of Doctor of Philosophy in MRI perfusion from the Institute of Neurology, University College London, London, UK. I confirm that the information that has been derived from other sources, as indicated in the thesis.

Acknowledgements

The support and help from many inspiring people around have enabled me to complete this PhD. I would like to express my sincere gratitude to the following people:

My supervisors, Professor Sotirios Bisdas and Professor Xavier Golay for their guidance, encouragement and thoughtful comments, as well for being there when needed. Their continuous support was fundamental to the fulfilment of this PhD thesis.

Dr David Thomas, for his generous assistance, and for always being available and his willingness to share his knowledge.

Dr Jasmina Panovska-Griffiths, who revised and helpfully provided thoughtful comments upon for the statistical methods used and the results reported in this thesis.

I also thank all the neuro-radiologists, neuropathologists, neurosurgeons, photographers and scientists who helped this study by identifying patients who were suitable for inclusion, providing the histopathological information, contacting and scanning the patients.

The participating patients for their time and trust. All the colleagues at the Brain Repair & Rehabilitation department for the interesting, enjoyable and motivating discussions.

Special thanks to my parents and siblings for their prayer, encouragements and their unconditional love. I dedicate this work to my father, who recently passed away; he used to encourage me to be a student forever and continue learning.

Thanks to the Kingdom of Saudi Arabia for supporting me financially through my academic studies and research as a lecturer at Taibah University.

Glory to the God who inspired every written word in this thesis facilitated my learning process and enabled me to understand myself better.

Publications

a. Scientific papers

1. Alsaedi A, Doniselli F, Jäger HR, et al. The value of arterial spin labelling in adults glioma grading: systematic review and meta-analysis. *Oncotarget*. 2019;10(16):1589-1601. doi:10.18632/oncotarget.26674.

2. Alsaedi A, Thomas D, Bisdas S, Golay X. Overview and Critical Appraisal of Arterial Spin Labelling Technique in Brain Perfusion Imaging. *Contrast Media Mol Imaging*. 2018;2018:1-15. doi:10.1155/2018/5360375.

b. Conference papers

1. Alsaedi A, Bisdas S, Golay X. Can Arterial Spin Labeling Metrics Serve as Imaging-Biomarkers for Isocitrate-Dehydrogenase (IDH) -Mutations in Gliomas?. E-poster at the ESMRMB 2017 Congress from October 19 – October 21, 2017, in Barcelona, Spain

2. Amirah Alsaedi, Fabio M. Doniselli, H.R. Jäger, Xavier Golay, Sotirios Bisdas. Systematic Review and Meta-analysis: Arterial Spin Labelling (ASL) Efficiency in Adults Glioma Grading. printed poster at BNOS-2018. from 4th-6thJuly 2018, in Winchester, UK

3. Amirah Alsaedi, Fabio M. Doniselli, H.R. Jäger, Xavier Golay, Sotirios Bisdas. The value of Arterial Spin Labelling in Adults Glioma Grading: Systematic Review and Meta-analysis. printed poster at ESNR-2018. from 19th – 23rd September 2018, in Rotterdam, The Netherlands

4. A. Alsaedi, D. Thomas, E. De Vita, J. Panovska-Griffths, S. Bisdas, X. Golay. Repeatability of perfusion measurements in gliomas using pulsed and pseudo-continuous arterial spin labelling MRI. printed poster at ESMRMB-2019. from 3ed – 5th October 2019, in Rotterdam, The Netherlands

5. A. Alsaedi, J. Panovska-Griffths, X. Golay, S. Bisdas. The Diagnostic Performance of Multiparametric MRI Radiomics for Classification of Untreated Adult Gliomas. ASNR 2020 (Virtual meeting)

6. A. Alsaedi, J. Panovska-Griffths, X. Golay, S. Bisdas. The Diagnostic Performance of Multiparametric MRI Radiomics for Classification of Untreated Adult Gliomas. ISMRM 2020 (Virtual meeting)

Abstract

Currently, the histological biopsy is the gold standard for classifying gliomas using the most recent histomolecular features. However, this process is both invasive and challenging, mainly when the lesion is in eloquent brain regions. Considering the complex interaction between the presence of the isocitrate dehydrogenase (IDH)-mutation, the upregulation of the hypoxia-induced factor (HIF), the neo-angiogenesis and the increased cellularity, perfusion MRI may be used indirectly for gliomas staging and further to predict the presence of key mutations, such as IDH. Recently, several studies have reported the subsidiary role of perfusion MRI in the prediction of gliomas histomolecular class. The three most common perfusion MRI methods are dynamic susceptibility contrast (DSC), dynamic contrast enhancement (DCE) and arterial spin labelling (ASL). Both DSC and DCE use exogenous contrast agent (CA) while ASL uses magnetically labelled blood water as an inherently diffusible tracer. ASL has begun to feature more prominently in clinical settings, as this method eliminates the need for CA and facilitates quantification of absolute cerebral blood flow (CBF). As a non-invasive, CA-free test, it can also be performed repeatedly where necessary. This makes it ideal for vulnerable patients, e.g. post-treatment oncological patients, who have reduced tolerance for high rate contrast injections and those suffering from renal insufficiency.

This thesis performed a systematic review and critical appraisal of the existing ASL techniques for brain perfusion estimation, followed by a further systematic review and meta-analysis of the published studies, which have quantitatively assessed the diagnostic performance of ASL for grading preoperative adult gliomas. The repeatability of absolute tumour blood flow (aTBF) and relative TBF (rTBF) ASL-derived measurements were estimated to investigate the reliability of these ASL biomarkers in the clinical routine. Finally, utilising the radiomics pipeline analysis, the added diagnostic performance of ASL compared with CA-based MRI perfusion techniques, including DSC and DCE, and diffusion-weighted imaging (DWI) was investigated for glioma class prediction according to the WHO-2016 classification.

Impact statemen

Contrasting to histological biopsy, surrogate non-invasive gliomas classification and grading are in demand for the patient management. Studies have demonstrated the diagnostic and predictive value of the quantitative biomarkers from perfusion MRI for patients with gliomas. ASL, unlike contrast-enhanced perfusion MRI techniques, measures perfusion completely non-invasive, using magnetically labelled blood water as a diffusible tracer, which eliminates the need for CA and hence it can be performed repeatedly, without concern related to the CA retention. The supplementary role of ASL to conventional imaging has been reported for classifying and grading gliomas. Nevertheless, ASL has still not been fully adopted in routine glioma MRI scanning due to non-standardised acquisition techniques and the presumably high within-patient variability. Both of which could bias the quantitative measurements and impact the interpretation of the results compared with the established DSC perfusion technique.

Nevertheless, with the continuous development, ASL has reached a stage that encourages its adoption to the routine clinical application, which does not contradict the persist in the development of the method. This PhD thesis explains the contrast-enhanced perfusion methods briefly and then emphasises on the ASL. It highlights the ASL methodological background, explaining the source and mitigation strategies of possible measurements bias. The thesis demonstrates the clinical effectiveness of the ASL for gliomas grading among the published studies by conducting a quantitative meta-analysis on the current evidence. Using this as a solid foundation, it investigates the sensitivity/specificity of the rTBF as a surrogate biomarker for the tumour pathophysiology and hence, gliomas grading. The key finding for this diagnostic task was the lowest within-patient variability recorded for the rTBF from both PASL and PCASL, suggesting its suitability for the clinical practice. Employing the radiomics pipeline approach, this thesis provides evidence of the synergistic diagnostic value of the combined ASL and ADC, which represents gadolinium-free gliomas staging approach, to replace the DSC technique as working horse for the gliomas staging in the clinical routine. The comprehensive work on the available DCE perfusion datasets showed that more detailed physiological models for tumour perfusion imaging,

such as L&L, yielded better staging accuracy compared to simplified approaches, like the mTK. This calls the attention of the researcher and the scientists on the role of perfusion, traditionally defined in terms of blood flow, derived from ASL and assisted by radiomics analysis to achieve high diagnostic accuracy in gliomas staging. Given a robust and standardised ASL acquisition, with a reproducible radiomics processing and analysing pipeline, the evidence in the submitted thesis can be the steppingstone for future large-scale clinical studies.

Table of Contents

Declaration ***2***

Acknowledgements ***3***

Publications ***4***

a. Scientific papers **4**

b. Conference papers **4**

Abstract ***6***

Impact statement ***7***

Table of Contents ***9***

Abbreviations ***13***

1 Introduction ***22***

1.1 Glioma classification **22**

1.2 Basic principles of magnetic resonance imaging (MRI) **26**

1.2.1 Underlying principles of MRI 27

1.2.2 Image contrast 34

1.2.3 Contrast agents in MRI 38

1.3 Advanced MRI **42**

1.3.1 Diffusion-weighted imaging (DWI) 42

1.3.2 Magnetic resonance spectroscopy (MRS) 44

1.3.3 Perfusion MRI 47

1.4 MRI for glioma primary diagnosis **48**

1.5 Image analysis **50**

1.5.1 Texture analysis and radiomics 52

2 Methodology ***58***

2.1 Tracer kinetics and tissue concentration-time curve basic concepts **58**

2.2 MRI perfusion **58**

2.2.1 Non-diffusible (exogenous) tracer 61

Arterial input function (AIF) 62
a. Dynamic susceptibility contrast-enhanced MRI (DSC) 64
b. Dynamic contrast-enhanced MRI (DCE) 67
2.2.2 Diffusible endogenous tracer 69
a. ASL main principle 69
b. Arterial spin labelling approaches 71
i. Continuous arterial spin labelling (CASL) 72
ii. Pulsed arterial spin labelling (PASL) 74
iii. Pseudo-continuous arterial spin labelling 79
c. Technical considerations for CBF quantification 81
D. CBF quantification models 83
i. T1 model 84
ii. General kinetic model 85
d. Improving the SNR 90
i. Hardware considerations 90
ii. Acquisition parameters and special techniques 91
e. Further parameters beyond the CBF 93
3 The value of Arterial Spin Labelling in Adults Glioma Grading: Systematic Review and Meta-analysis 95
Abstract 95
3.1 Introduction 96
3.2 Material and methods 96
3.2.1 Literature search and selection 96
3.2.2 Statistical Methods 100
3.3 Results 101
3.3.1 Eligible ASL studies 101
3.3.2 QUADAS-2 assessment 101
3.3.3 Differentiation between HGGs and LGGs 102
3.3.4 Differentiation of grade II from grade III gliomas 103

3.3.5 Differentiation of grade II from grade IV gliomas 103
3.3.6 Differentiation of grade III from grade IV gliomas 103
3.3.7 HROC curve analysis of gliomas grading 106
3.3.8 HROC curve analysis for differentiation between HGGs and LGGs 108
3.4 Discussion 111
3.5 Conclusion 116
4 Repeatability of Perfusion Measurements in Gliomas Using Pulsed and Pseudo-continuous Arterial Spin Labelling (ASL) MRI 118
4.1 Introduction 119
4.2 Methods 120
4.2.1 Patients 120
4.2.2 ASL acquisitions 120
4.2.3 CBF maps calculation 122
4.2.4 Image analysis 125
4.2.5 Statistical analysis 126
4.3 Results 128
4.4 Discussion 141
4.5 Conclusion 143
5 The Diagnostic Performance of Different Advanced-MRI Based on Radiomics for Noninvasive Histomolecular Staging of Untreated Adult Gliomas 145
5.1 Introduction 146
5.2 Material and Methods 147
5.2.1 Study design 147
5.2.2 Patients 147
5.2.3 Hand-crafted radiomics 148
5.2.4 Image acquisition and processing 150
5.2.5 Image analysis and feature extraction 155
5.2.6 Feature reduction and statistical analysis 156
5.3 Results 158

5.3.1 Histogram features 158
5.3.2 Diagnostic performance 158
ii. The significant features separately 158
iii. The combined significant features for different MRI methods 163
5.4 Discussion 168
5.5 Conclusion 174
6 Conclusion and further work 185
6.1 Summary 185
6.2 Recommendation for further work 186
6.2.1 Image acquisition and processing 187
6.2.2 Further quantitative estimates from ASL 187
6.2.3 Sample size and validation 189
6.3 Role of ASL in other brain tumours and surveillance 189
6.3.1 ASL in other brain tumours 189
6.3.2 ASL in brain tumours surveillance 190
6.4 ASL in paediatric brain tumours 191
7 References 193
Appendix A: Figures 214
Tables 225
Appendix B 233

Abbreviations

λ	the coefficient for water blood/water partition
(-M0a)	inverted blood water magnetisation
Δt	initial transit time
Δt+τ	ongoing delivery of the tagged blood to the imaging slice with a length equal to the bolus width
^{13}C	carbon
2-Hg	2-hydroxyglutarate
^{31}P	phosphorus
3D	three dimensions
95tile	95th-percentile
AAT	arterial arrival time
aBV	arterial blood volume
ADC	apparent diffusion coefficient
AIF/CA(t)	arterial input function/concentration in arterial blood
ASL	arterial spin labelling
astro-mut-HG	grade III IDH-mutant astrocytoma
astro-mut-LG	grade II IDH-mutant astrocytoma
aTBF	absolute tumour blood flow
ATRX protein	alpha-thalassemia/mental retardation syndrome X-linked
B0	static magnetic field
BAT	bolus arrival time
BBB	blood-brain barrier
BE	backward elimination
Beff	effective field
BS	background suppression

CA	contrast agent
CASL	continuous arterial spin labelling
caud	caudate nucleus
CBF	cerebral blood flow
CDF	cumulative distribution function
CHESS	chemical shift selective
Cho	choline compounds
CI	confidence intervals
CNAH	contralateral normal-appearing hemisphere
CNS	central nervous system
Cr	Creatine compounds
CR	Coefficient of Repeatability
CSF	cerebrospinal fluid
CSI	multi-voxel chemical shift imaging
CT(t)	tissue concentration
CV(t)	venous concentration
DCE-MRI	dynamic contrast-enhanced MRI
DICOM	Digital Imaging and Communications in Medicine
DSC-MRI	dynamic susceptibility contrast-enhanced MRI
DWI	diffusion-weighted imaging
E%	extraction fraction
EES	extravascular-extracellular space
EGFR	epidermal growth factor receptor
EMEA	European Medicines Agency
EPISTAR	Echo-planar imaging and signal-targeting with alternating radiofrequency

f	regional blood flow
f0	Larmor frequency
FA	flip angle
FAIR	flow-sensitive inversion pulse
Fb	plasma flow
FFT	Fourier transformation
FID	free induction decay
FLAIR	fluid-attenuated inversion recovery
FLASH	fast low angle shot
FMRIB	Functional MRI of the Brain
FN	false-negative
FP	false-positive
FSL	(FMRIB)-Software Library
GBM	glioblastoma
GBM-wt	glioblastoma IDH-wild type
Gd	gadolinium
GFR	glomerular filtration rate
GM	grey matter
GRASE	Gradient- and spin-echo
GRE	gradient-echo
H2O	water molecule
HGGs	high-grade gliomas
HGGs	high-grade gliomas
HIF	hypoxia-induced factor
HROC	summary receiver operating characteristic
Hz	Hertz

ICC	intra-class correlation coefficient
IDH	isocitrate dehydrogenase
IDH-mut	IDH mutation
IDH-wt	IDH-wildtype
Int16	16-bit integers
iqr	inter-quartile range
IR	inversion recovery
ISMRM	International Society for Magnetic Resonance in Medicine
IT	inversion time/ delay time in PASL
IVIM	intravoxel incoherent motion
Kep, min-1	reflux constant
Ktrans, min-1	permeability constant
kur	kurtosis
L&L	St Lawrence and Lee model
LGGs	low-grade gliomas
LGGs	low-grade gliomas
M0a	fully relaxed blood water magnetisation
Ma(t)	magnetisations for the arterial blood
max	maximum
mI	Myo Inositol
min	minimum
MP	maximum peak
MP	maximum peak
MPRAGE	magnetisation-prepared rapid gradient-echo
MRA	MR-angiography
MRI	Magnetic resonance imaging

MRS	MR-spectroscopy
MT	magnetisation transfer effect
Mt(t)	magnetisations for the tissue
mTK	modified Tofts-Kermode model
MTT	mean transit time
Mv(t)	magnetisations for the venous blood
Mxy	transverse magnetisation
Mz	longitudinal magnetisation
NAA	N-acetyl aspartate
NADPH	Nicotinamide adenine dinucleotide phosphate
NIfTI	Neuroimaging Informatics Technology Initiative
NMV	net magnetisation vector
NOS	not otherwise specified
NPV	negative predictive value
NSF	nephrogenic systemic fibrosis
oligo-mut-HG	grade III IDH-mutant oligodendroglioma
oligo-mut-LG	grade II IDH-mutant oligodendroglioma
PASL	pulsed arterial spin labelling
pc	phase contrast
PC	pairwise-correlation
PCASL	pseudo-continuous arterial spin labelling
PD	Proton density
PD-W	PD weighted imaging
PHD	prolyl hydroxylases
PICO	Population/Intervention/Comparator/Outcomes
PICORE	Proximal inversion with a control for off-resonance effect

PK	pharmacokinetic model
PLD	post labelling delay time
ppm	parts per million
PPV	positive predictive value
PS, min-1	separate estimation of permeability
PVE	partial volume effect
Q1	first quartile
Q2TIPS	QUIPSS II with thin-slice TI1 periodic saturation
Q3	third quartile
QIBA	quantitative imaging biomarkers alliance
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies-2
QUASAR	quantitative STAR labelling of arterial region
QUIPSS II	quantitative image of perfusion using a single subtraction
R(t)	residue function
r1/r2	the relaxivity of the used contrast agent
R10	longitudinal relaxation without the CA
R20	transverse relaxation without the CA
rCBV	Relative normalised cerebral blood volume
RF	radiofrequency
ROC	receiver operating curve analysis
ROI	region of interest
ROS	reactive oxygen species
rTBF	relative tumour blood flow
rTBFmax	maximum rTBF
SAR	specific absorption rate
SD	standard deviation

SE	spin-echo
SI	signal intensity
skew	skewness
SMD	standardised mean difference
SNR	signal-to-noise ratio
SRD	smallest real difference
SVS	single-voxel spectroscopy
T0	onset time
T1-c	T1-W post-contrast
T1-W	T1-weighted image
T1a	The T1 relaxation time of the arterial blood
T1app	apparent longitudinal relaxation time
T2-W	T2-weighted image
TBF	tumour blood flow
TC, min	mean capillary transit time
TE	echo time
TI1	initial inversion time
Tmax	maximum time
TMS	tetramethylsilane
TN	true-negative
TP	true-positive
TR	are repetition time
Tt1	The T1 relaxation time of the tissue
TTP	time to peak
var	and variance
Ve%	the fractional volume of the EES

VEGF	vascular endothelial growth factor
Vp%	fractional plasma volume
w-average-F1	weighted average F1 score
WCoV	Within-Coefficient-of-Variation
WHO	world health organisation
WM	white matter
WS-SD	within-subject standard deviation
zscore	median z-score
α	labelling efficiency
α-Kg	α-Ketoglutarate
γ	gyromagnetic ratio
τ	bolus width

Chapter 1

1 Introduction 22

1.1 Glioma classification 22

1.2 Basic principles of magnetic resonance imaging (MRI) 26

1.2.1 Underlying principles of MRI 27

1.2.2 Image contrast 34

1.2.3 Contrast agents in MRI 38

1.3 Advanced MRI 42

1.3.1 Diffusion-weighted imaging (DWI) 42

1.3.2 Magnetic resonance spectroscopy (MRS) 44

1.3.3 Perfusion MRI 47

1.4 MRI for glioma primary diagnosis 48

1.5 Image analysis 50

1.5.1 Texture analysis and radiomics 52

This thesis evaluates the added value of ASL perfusion in gliomas staging as a standalone method and in combination with other perfusion techniques and DWI. The first chapter provides a brief introduction into the fundamentals of MRI and the principles of advanced MRI in neuro-oncology. The second chapter focuses primarily on the methodological background of ASL and secondarily of DSC and DCE. The third chapter aims to summarise quantitatively the previous studies that have explored the clinical application of the ASL in grading adult gliomas. The fourth chapter investigates the repeatability of ASL estimates (TBF and rTBF) for gliomas and the impact of the internal reference tissue size and type (white or grey brain matter) for the estimation of rTBF, which is a common metric for ASL-based gliomas staging. The fifth chapter studies the diagnostic performance of the radiomics approach of different MRI perfusion methods (DSC, DCE and ASL), as well as the ubiquitously available diffusion weighted imaging (DWI) to classify untreated adult gliomas. In this chapter, the diagnostic performance of DCE is particularly investigated using both the modified Tofts-Kermode (mTK) and the St Lawrence and Lee (L&L) models. The final chapter concludes and summarises the findings and provides recommendations for future work.

1 Introduction

Magnetic resonance imaging (MRI) plays a fundamental role in the characterisation of gliomas as it allows the generation of various image contrast types, including conventional structural and advanced functional contrasts, which enable non-invasive assessment in details of the underlying lesion's morphology and physiology. This chapter thus provides a comprehensive review of glioma classification, in this manner, including an overview of the basic principles of MRI, and advanced MRI techniques and the relevant image analysis.

1.1 Glioma classification

Gliomas are the most common neuroepithelial neoplasms of primary brain tumours, which generally arise from progenitor glial cells, including astrocytes and oligodendrocytes. The World Health Organization's (WHO) classification of brain tumours is the most widely adopted classification system, and this is updated approximately every seven years. The WHO-2007[1] edition classified

gliomas according to histological criteria, grading the tumours according to their biological behaviour from grade I - IV. This grading system reflected the degree of tumour malignancy, with grades I and II indicating slow-growing lesions, grouped together as low-grade gliomas (LGGs), and grades III and IV representing highly malignant lesions, grouped together as high-grade gliomas (HGGs). HGGs represent the largest group of adult gliomas, tending to occur in older patients, while LGGs more frequently occur in younger patients. This histopathological grading is tightly linked to the prognosis and has hugely influenced the selection of treatment options. The median overall survival time for patients with LGGs exceeds in most of the entities the five years. In comparison, patients with HGGs reach median overall survival rates of 2 to 3 years for grade III gliomas and less than one year for grade IV. The updated WHO-2016 classification[2] represented a paradigm shift towards a more multi-layered approach, with genetic and molecular information (genotyping) incorporated into the histological grading (phenotyping) to form an integrated system. This approach resulted in tumour groups that are biologically homogenous; this, in turn, facilitated more precise categorisation than the previous classification system and hence, improved overall diagnostic accuracy, more optimised patient management, and accurate prognoses with reliable treatment response predictions.

All adult diffuse gliomas are currently grouped together based on shared genotypes and phenotypes, resulting in the following categories: a) WHO grade II and III astrocytomas (IDH-mutant, ATRX-mutant, 1p/19q-intact), b) WHO grade II and III oligodendrogliomas (IDH-mutant, ATRX-wildtype, 1p/19q-codeleted), and c) WHO grade IV glioblastoma primary/secondary (IDH-wildtype/mutant, respectively). In rare situations, however, gliomas may be histologically inconsistent with the predefined observed molecular and genetic classification criteria, e.g. oligoastrocytomas, or may not be adequately described as distinct entities in the past and therefore are denoted as "not otherwise specified" (NOS). Thus, for the time being, the WHO-2016 [2] classification is not solely based on the genotype.

A key mutation in the gliomas tumorigenesis with crucial importance for the integrated histomolecular classification is the isocitrate dehydrogenase (IDH) mutation. This mutation occurs early and can potentially initiate the glioma pathogenesis[3]. It occurs in 70-80% of grade II and III astrocytomas, oligodendrogliomas, and secondary glioblastomas[4]. In contrast, about 90% of all de novo/primary glioblastomas are IDH-wt[2]. IDH-mutation is considered to be a strong prognostic marker independent of other known prognostic factors such as age[5,6]. Importantly, the presence of IDH-mutation and 1p/19q co-deletion in oligodendrogliomas is reported to presage better responses to chemotherapy and longer overall survival times compared to IDH-mutation and 1p/19q retained gliomas [7–9].

Mutant IDH proteins develop neo-morphic enzymatic activities that are involved in the cancer cell metabolism and along with tumour suppressor genes controls the progression of the tumour and its response to chemoradiation, which in turn impacts the prognosis[10,11]. Such changes to cellular metabolism lead to uncontrolled growth, which requires high energy levels, and this increased demand for energy results in elevated oxidative stress (high reactive oxygen species (ROS)). The mechanisms controlling the resultant oxidative stress are influenced by the IDH-mutation status, which in turn affect the hypoxia-induced factor (HIF), and regulation of the angiogenesis and cellularity (see Figure 1.1). As the metabolic alteration impacts the tumour microenvironment, it can result in sporadic hypoxia, acidity, and nutrition starvation. This elevates the ROS, where Nicotinamide adenine dinucleotide phosphate (NADPH) helps reduce it; NADPH is generated via the catalysis of isocitrate during conversion, using isocitrate dehydrogenase (IDH)-enzymes. This process produces α-Ketoglutarate (α-Kg) and NADPH. However, in the case of mutant IDH proteins, the α-Kg converts to 2-hydroxyglutarate (2-Hg) on the consumption of NADPH, and thus any reduction of NADPH leads to an increase in ROS.

Further, as 2-Hg is chemically similar to α-Kg, it can bind to α-Kg-dependent enzymes and disturb their functions. For example, 2-Hg binding to prolyl hydroxylases (PHD) inhibits the degradation of HIF and stabilises its formation, subsequently inducing vascular endothelial growth factor (VEGF) signalling and

angiogenesis. Also, 2-Hg can bind to histone demethylase, resulting in the promotion of methylation and consequently inactivating both oncogenesis and tumour suppressor genes. The presence of IDH-mutations thus promotes tumour progression only slowly, perhaps explaining the better survival rates, while high ROS levels may sensitise tumours to irradiation and chemotherapy[12]. However, IDH-mutation alone is not enough for tumourigenesis[13]; besides, the presence of IDH-mutation with other molecular markers can be used to differentiate astrocytoma from oligodendroglioma with prognostic and predictive outcomes. IDH-mutant astrocytoma involves more genetic alterations, including ATRX (alpha-thalassemia/mental retardation syndrome X-linked) that associated with tumour progression from diffuse to anaplastic[14]. Anaplastic astrocytoma with mutated IDH and ATRX loss has a favourable prognosis comparing to anaplastic astrocytoma that only has IDH mutation[3]. On the other hand, IDH-mutant oligodendroglioma coexists with 1p/19q codeletion, which is indicated as a hallmark feature, and present approximately in 60%-90% of the oligodendrogliomas patients[14]. The existence of 1p/19q codeletion is associated with longer survival time and enabled identifying patients who benefit from chemotherapy[3].

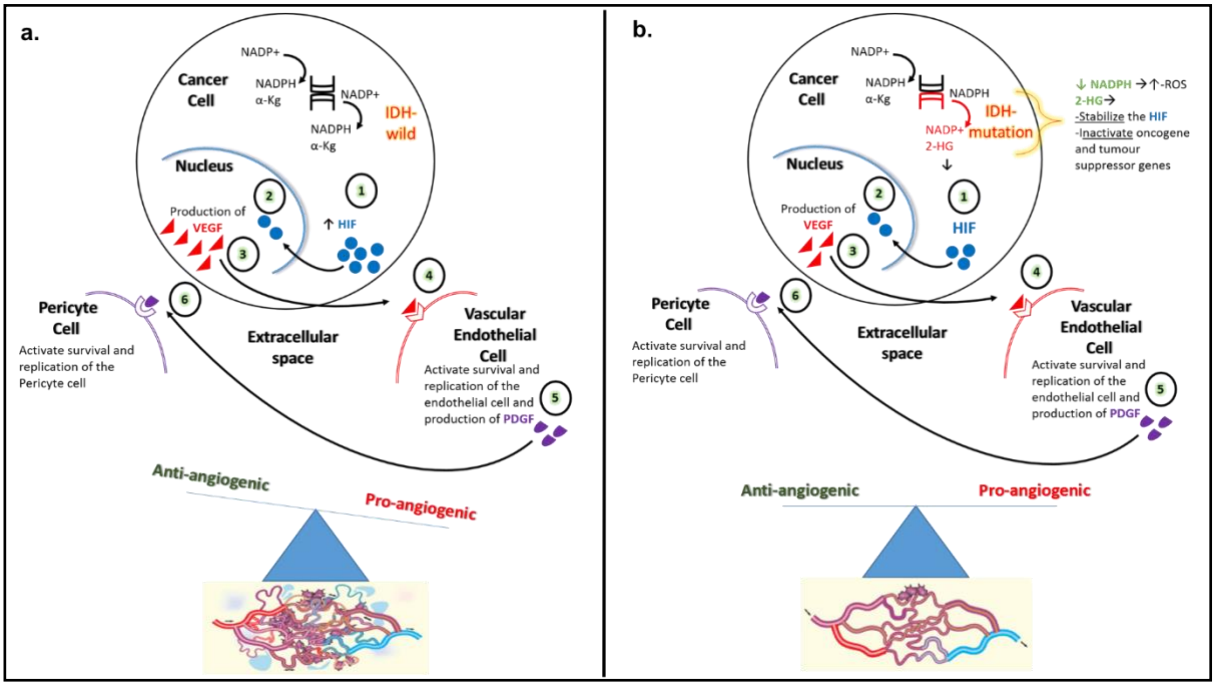


Figure 1.1 Explain how the IDH status involved in cancer cell metabolism; Presence of gene mutations results in cell metabolic changes, which is lead to high demand for energy and high oxidative stress (high reactive oxygen species (ROS)). Nicotinamide adenine dinucleotide phosphate (NADPH) helps to reduce ROS. NADPH generated via various reaction inside the cell, conversion of isocitrate is one of them. Catalysis of isocitrate, during conversion, uses isocitrate dehydrogenase (IDH) enzymes. (a.) in case of absence of IDH mutation (IDH-wild), the catalysis of isocitrate using isocitrate dehydrogenase (IDH) enzymes produce α-Kg (α-Ketoglutarate) and NADPH.

(b.) In contrast, in the case of mutated IDH, α-Kg converts to 2-Hg (2-hydroxyglutarate) on the consumption of NADPH. Resulted in a reduction of NADPH and leads to an increase in ROS. As 2-Hg is chemically similar to α-Kg, it can bind to α-Kg dependent enzymes and disturb their functions. For example, 2-Hg binding to prolyl hydroxylases (PHD) inhibits the degradation of hypoxia-inducible factor (HIF) and stabilises its formation, subsequently, the inducing of vascular endothelial growth factor (VEGF) signalling and angiogenesis. 2-Hg also binds to histone demethylase, resulting in the promotion of methylation and consequently inactivating oncogenesis and tumour suppressor genes. Thus, the presence of IDH-mutations promotes tumour progression process slowly, allowing better survival rates, while high ROS levels may sensitise tumours to irradiation and chemotherapy.

1.2 Basic principles of magnetic resonance imaging (MRI)

For brain tumour diagnosis, histopathologic assessment is the gold standard. However, it suffers from several limitations, including sampling errors and dependence on invasive biopsies or the surgical removal of a tumour[15]. Imaging tests have thus been proposed as an alternative non-invasive method to stage the brain tumours. The application of appropriate imaging modalities in a clinical setting can make a great contribution toward the detection, diagnosis, and monitoring of gliomas by allowing the correlation of molecular biomarkers to the imaging feature outputs.

Due to its non-invasive nature and high sensitivity to soft tissue changes, MRI plays a leading role in all phases of tumour management, including diagnosis, therapy, and follow-up. Typical standard clinical MRI protocols for gliomas diagnosis consist of both conventional MRI and diffusion-weighted imaging (DWI)[16]. Also, advanced MRI methods such as perfusion MRI and/or MR-spectroscopy (MRS) can be used where it is believed they may offer supportive diagnostic information[16].

1.2.1 Underlying principles of MRI

Nuclear magnetic resonance imaging (NMR) study the resonance of the nuclei under magnetic field. Incorporation the spatial information of the resonance of these nuclei to produce an image is then called magnetic resonance imaging (MRI). Thus, MRI uses three main components to create images: a strong magnet to align the protons; radiofrequency (RF)-coils to excite the protons and detect their return signals; and gradient coils for spatial localisation of those signals, as shown in Figure 1.2. It is thus essential to understand the underlying basics of MRI to grasp its complexities, and this, in turn, starts with understanding of atomic structure.

Atom is the fundamental unit of matter. All tissues are made of molecules, with atoms organised in specific ways to form these molecules, such as the hydrogen and oxygen atoms in a water molecule (H_2O). Atoms themselves consist of a central nucleus surrounded by orbital electrons. The nucleus of the atom is composed of protons and neutrons both represents the atomic mass. Both neutrons and protons have an opposite magnetic moment[17]. Thus, nuclei with even protons and neutrons have zero net spins while nuclei with odd protons or neutrons have a net spin (angular moment), making them MR active nuclei, which act like tiny magnets with north and south poles (dipoles)[18]. As MR active nuclei have an electrical charge and net spin, they automatically induce magnetic moment under an external magnetic field, according to electromagnetic induction laws.

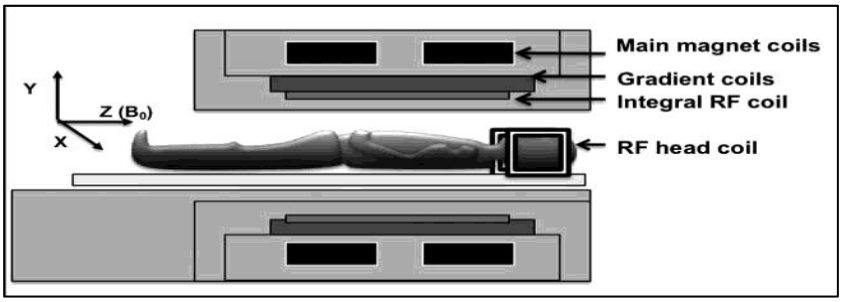


Figure 1.2 Schematic demonstrating the main components of an MRI machine. Reprinted from Understanding MRI: Basic MR physics for physicians. *Postgrad Med J.* 2013;89(1050):209-223. doi:10.1136/postgradmedj-2012-131342.

Hydrogen atoms, often called protons in MRI terminology, are the most frequently used MR active nuclei due to their abundance in living tissue (~75% of fat and water). These also have odd mass numbers, as a hydrogen nucleus consists of a solitary proton with a positive charge, as seen in Figure 1.3.a. Other MR active nuclei that are less abundant than hydrogen is also used, such as Fluorine, Carbon, and Sodium. Hydrogen under an external strong static magnetic field (B_0) experience a torque result in the precession of their magnetic moment about B_0. This precession occurs at a specific angular frequency, also called precessional frequency, denoted by ω_0 or f_0 (in cycle/sec or Hertz (Hz)) which is proportional to the strength of the applied magnetic field and the gyromagnetic ratio (γ) of the nuclei as stated in Larmor's equation:

$$w0 = B0.\gamma \qquad \text{Equation 1.1}$$

Where γ is a constant known as the gyro-magnetic ratio which represents the relationship between the angular moment and the induced magnetic moment for a specific MR active nucleus; hence each MR active nuclei has its γ, and nuclei of the same type (e.g. hydrogen) will nevertheless have changeable precessional frequencies as they experience different field strengths.

The magnetic moment can be described as a vector, its arrow represents the alignment direction, and its length represents the size of the magnetic moment. The alignment direction of the magnetic moment of the protons can be either parallel (spin-up) or anti-parallel (spin-down) to B_0. Nuclei with low energy lack the capacity to oppose the B_0 and align parallel to it. In contrast, a few nuclei have

higher energy levels and can thus align themselves anti-parallel to B_0. The excess low energy nuclei that align with B_0 form a net magnetisation vector (NMV), as shown in Figure 1.3.b and 1.4.a. Any increase in B_0 field strength increases the number of low energy nuclei, subsequently increase the excess spin-up nuclei, improving MR signals accordingly.

The NMV in space can be described using a rotating frame, which facilitates understanding of its complex motion. This frame views the magnetisation precession in slow motion observed from frame rotating at Larmor frequency. Thus, expressing the magnetisation as the component of a static vector on X', Y' and Z' axes. The magnetisation on Z-axis is the same direction as that in which the NMV aligned with B_0, known as the longitudinal magnetisation (M_z). The X-Y axes magnetisation (transverse plane) is the flipped vector of the rotating magnetisation around the Z-direction at an angle under the influence of an external force, known as transverse magnetisation (M_{xy}).

The longitudinal magnetisation (M_z) precess at Larmor frequency (f_0) about the B_0 is at an equilibrium state. In order to detect signal, the M_z need to be perturbed from its equilibrium by applying oscillating electromagnetic field (B_1) generated using the transmitted RF-coil[19]. If the precessional frequency (f_1) of the applied B_1 field in X-Y plane matches that of the NMV ($f_0 = f_1$), the resonance condition achieved; hence M_z flips toward the X-Y plane with a particular angle to the B_0, known as the flip angle. The magnetisation precesses about B_1 at f_1, which forming the M_{xy}. The M_z nutates toward X-Y plane until B_1 field removed. Thus, the magnitude of the flip angle is determined by the amplitude and duration of the applied RF-pulse. There are different types of RF-pulses, including excitation and inversion pulses. Excitation RF-pulses tip the NMV away from the B_0 towards the transverse plane which generates a voltage in the receiver coil, while inversion RF-pulses flip the NMV from the B_0 direction to the negative B_0 direction, with no or only a negligible transverse component.

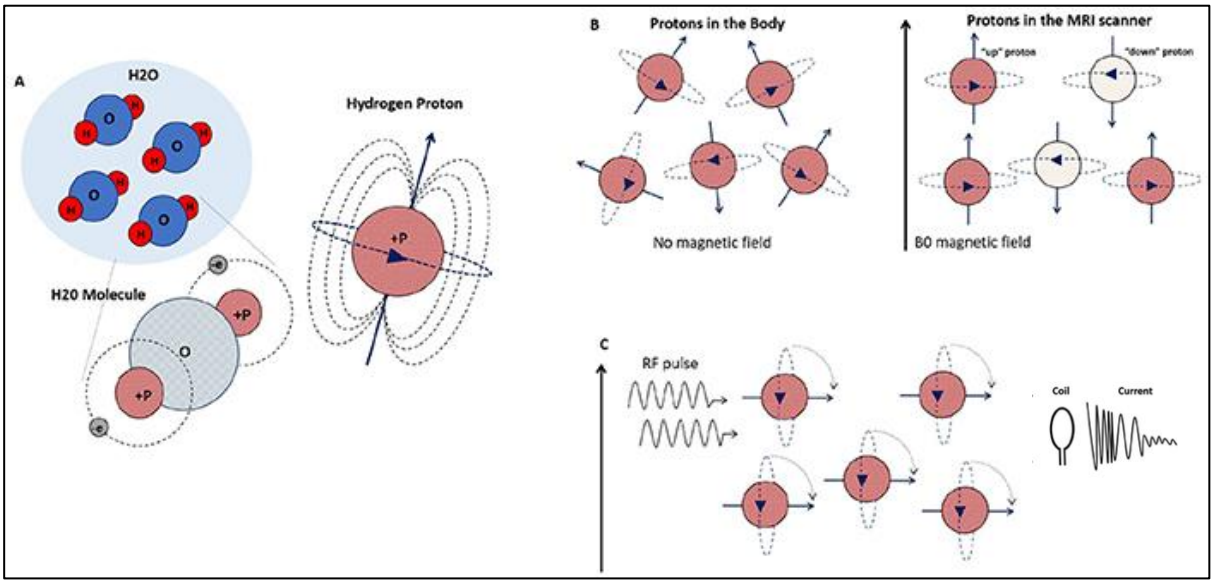


Figure 1.3 A) Water, composed of one oxygen atom (blue) and two hydrogen atoms (red). Hydrogen atom (+P, in red) is one of the active MRI nuclei, with net spin due to the solitary proton which acts as a tiny magnet. B) These protons usually spin randomly and cancel each other out; however, in the presence of a strong static magnetic field (B_0) the low- and high-energy protons align with the B_0 in either a parallel (red) or an anti-parallel (white) manner, respectively. The parallel and anti-parallel protons cancel each other out, but there remain excess parallel protons, which form a net magnetisation vector (NMV). C) The transmitted RF-pulse delivers electromagnetic energy to those aligned protons with the same precessional frequency as the protons. The protons flip away from B_0 as they absorb the energy, and as the spinning protons cut across the coil in the transverse plane, a current is induced, known as free induction decay (FID)[48]. Reprinted from https://kids.frontiersin.org/article/10.3389/frym.2019.00023

As the spinning protons flipped into the transverse plane after the B_1 field removed, a current is induced in the receiver coil that constitutes the expected MR signals; the induced voltage decreases exponentially with time, this is called free induction decay (FID), as illustrated in Figure 1.3.c. All of the protons flipped to the transverse plane have the same position in the precessional path initially, which means that they are in the same phase (in-phase) and provide the highest possible signal, as illustrated in Figure 1.4.b. When the RF-pulse ceases, the protons give up the absorbed energy, and the spinning protons' position in the precessional path at a given time point depends on their precessional speed, causing them to adopt different phases (dephase) over time, and signal decay. As a vector quantity, once the RF-pulse is removed, the M_z magnetisation increases while the M_{xy} magnetisation decays in a process called relaxation, shown in Figures 1.4.c and 1.4.d. The protons loss the energy in two main mechanisms synchronic but independent; the first mechanism is losing energy to the surrounding molecules (spin-lattice relaxation) where the M_z magnetisation

exponentially increase and back to the original state after T1 constant time, and called T1 recovery; the second mechanism is losing energy to the neighbour nuclei (spin-spin relaxation) and hence become out of the phase where the M_{xy} magnetisation exponentially decrease after T2 constant time and called T2 decay. T2 time constant is always shorter than the T1 time constant[18]. Different inherent tissue properties in a magnetic field result in various T1 and T2 rates and hence generate different amounts of signal intensity over time, creating contrast[20].

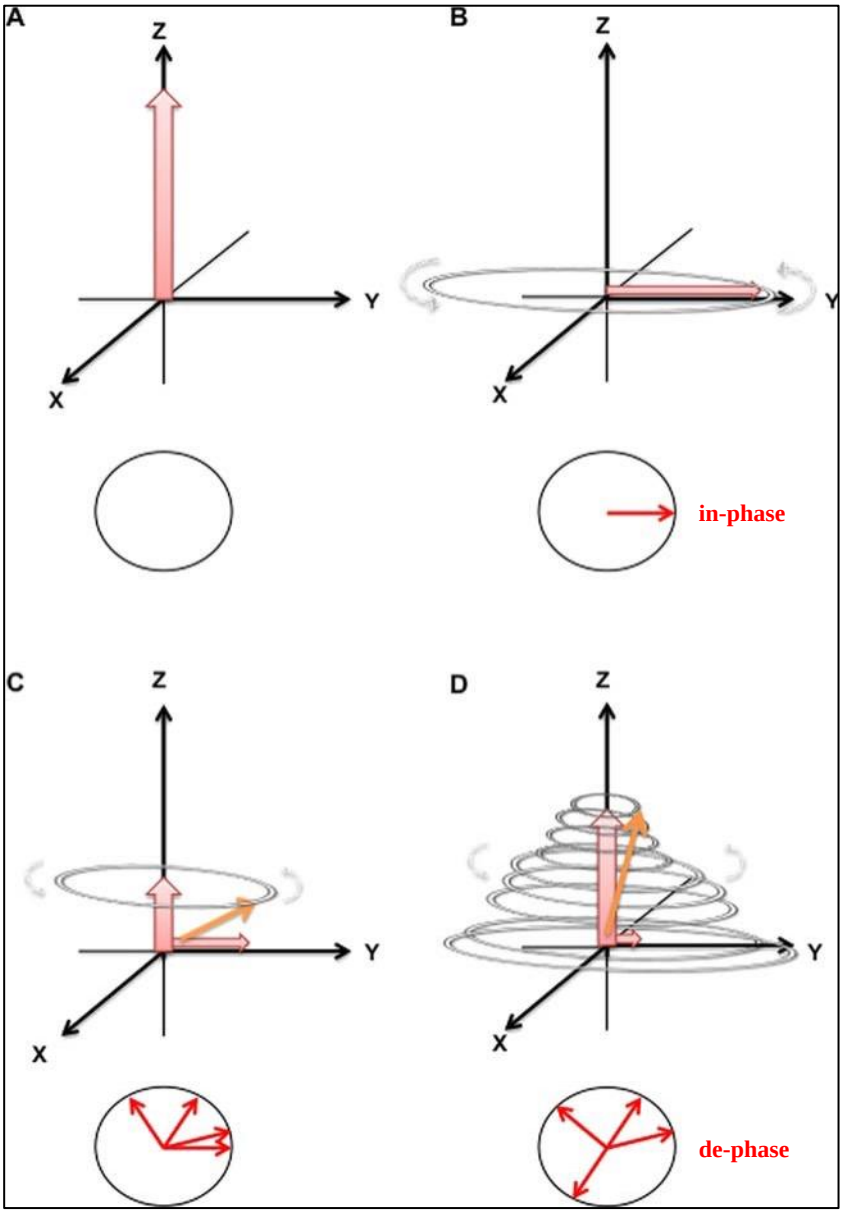


Figure 1.4 demonstrates the relaxation process. A) The net magnetisation vector (NMV) aligns with B_0, longitudinal magnetisation M_z. B) As the RF-pulse switches on, the transverse magnetisation (M_{xy}) propagates while the M_z is reduced. C) Immediately following the discontinuation of the RF-pulse, the protons precess in-phase and provide the highest possible signal. C and D) The M_{xy} magnetisation decays over time (T2 decay) while the M_z regrows (T1 recovery). Reprinted from Understanding MRI: Basic MR physics for physicians. Postgrad Med J. 2013;89(1050):209-223. doi:10.1136/postgradmedj-2012-131342.

The relaxation process is enhanced when the precessional frequencies of the protons and the surrounded moving molecules match. The molecular motion is impacted by both molecular size itself and the physical state of the material in terms of it being fluid or rigid. For example, smaller molecules such as water display faster movement than larger ones such as lipids. Molecules in fluid state

materials tend to display faster molecular motion over a wide range than the those in rigid materials, which have both slower molecular motion and less range. Thus, rigid state materials offer a better frequency match with the protons and are thus able to absorb energy from them more readily, enhancing the relaxation and reducing the time taken. Although T2 values are generally much shorter than T1 values, they are proportional to one another such that tissues that have longer T1 periods will generally have longer T2 times as well. Which means large molecules that are closely packed allow both T1-recovery and T2-decay to be relatively rapid, with short T1 and T2 times. In contrast, small molecules that are spaced further apart have slower T1-recovery and T2-decay and thus display longer T1 and T2 times. Tissues generally contain a combination of free water and combined water to a variety of small and large molecules which impact the relaxation process, and changes in the water composition of tissue, therefore, alter its relaxation process, which enables depicting the normal anatomy and detection of abnormalities by observing these changes.

On an MR image, the signal intensity (brightness/darkness) is directly related to the magnitude of the MR signals received by the receiver coil after the RF-pulse. This magnitude depends on the size of the M_z magnetisation prior to the RF-pulse and the size of the M_{xy} magnetisation after the RF-pulse. The larger the longitudinal vector component at the time the RF-pulse is employed, the greater the transverse magnetisation. Large M_{xy} magnetisation induces a high signal intensity that appears as a bright white area on the MRI image; the lower the M_{xy} magnetisation, the lower the resulting signal intensity, so that very low levels appear as dark signals, shading to black. Areas with intermediate signal intensity are expressed in shades of grey proportional in brightness to the relevant intensity. Each voxel of tissue is a functional unit of protons that generates an MR signal of a particular strength that is displayed on the MR image matrix as a single shade of grey. The spatial location of the representative MR signal for each tissue voxel is then encoded using an MRI gradient, coil, which is an additional magnetic coil implemented inside the bore of the MR machine on the three orthogonal axes, X, Y and Z. This gradient coil interacts with the main magnetic field and alters its strength in a linear manner along each axis, which can be used for slice selection

as well as being applied in advanced MRI methods such as diffusion-weighted imaging (DWI)[20], as will be seen later on.

1.2.2 Image contrast

A combination of RF-pulses, gradients, and predefined timing parameters outline the main components of a pulse sequence. The primary timing parameters in a pulse sequence are repetition time (TR) and echo time (TE), both measured in ms. TR measures the elapsed time between the excitation RF-pulses, which control the overall regrowth of the M_z magnetisation, and thus T1-recovery. At the same time, TE determines the time of reading of the received signal, is defined as the elapsed time duration from the RF-pulse to the time the signal is read. Thus, TE controls how much of the M_{xy} magnetisation is allowed to decay, which represents the T2-decay.

Variation in the signals between areas of an MR image allows the observation of contrasts between normal anatomical features and enables abnormality detection. MRI procedures allow a wide range of contrasts production, such as T1 weighted (T1-W) and T2 weighted (T2-W) images, which depend on the selected timing parameters. Obtaining a given MR image contrast depends on predetermined extrinsic parameters set by the operator console that can control the relative contribution of each of the intrinsic properties of tissues, such as T1-recovery and T2-decay, in a process referred to as image weight contrast. The process of image weighting requires setting the image contrast toward one property (e.g. T1) and away from the other (e.g. T2)[21].

A T1 weighted image (T1-W) is where the image contrast is predominantly determined by the T1 properties of the tissues being imaged. TR controls T1 recovery as it determines how far the M_Z magnetisation is permitted to recover prior to the next RF pulse occurring. In order to achieve T1-W, a short TR is required prior to full recovery of M_z magnetisation. Thus, the tissues that recover its M_z magnetisation faster develop larger M_z magnetisation before the next RF-pulse, resulting in larger M_{xy} magnetisation. This creates image contrast based on the differences in T1 values of the imaged tissues. The M_z magnetisation of tissues with small T1 values displays faster T1-recovery as well as larger M_z

magnetisation before the next RF pulse, resulting in larger M_{xy} magnetisation levels that thus appear bright, such as white matter (WM), while the reverse is true for tissues with longer T1 values such as cerebrospinal fluid (CSF), which appear dark[20], Figure 1.5.

In contrast, a T2 weighted image (T2-W) takes its image contrast predominantly from the T2 properties of the tissues being imaged. TE controls T2 decay as it determines the time allotted for the M_{xy} magnetisation to dephase or decay after the RF pulse prior to reading the signals. Longer TE provides enough time for the M_{xy} magnetisation to dephase, leading to lower M_{xy} magnetisation. To achieve a T2-W contrast, a long TE is required to produce image contrast based on differences in the decay rates (T2 values) of the scanned tissues, although this must still be short enough to preserve the signals. On the contrary to the T1-W image, the M_{xy} magnetisation of the tissues with small T2 values have quicker T2-decay end up with small M_{xy} magnetisation and appear dark (e.g. WM); The reverse occurs for the tissues with longer T2 values, which is appear white (e.g. CSF)[20], Figure 1.5.

Proton density (PD) is another tissue property based on the number of protons per unit volume of tissue. Producing an image contrast based on differences in the PD between tissues is referred to as PD weighted imaging (PD-W). Achieve a PD-W contrast image require reducing the effect of both T1 and T2. According to the image weighting definition, fulfil a T1-W requires the use of a short TR to maximise T1 and a short TE to minimise T2, while a T2-W requires a long TE to maximise T2 and a long TR to minimise T1. Accordingly, a long TR reduces T1 contrast and a short TE reduces T2 contrast, and PD contrast can thus be dominated image contrast. Tissues with higher proton densities produce higher levels of M_{xy} magnetisation and appear bright, while those with lower proton densities are dark on PD-W images[20], Figure 1.5.

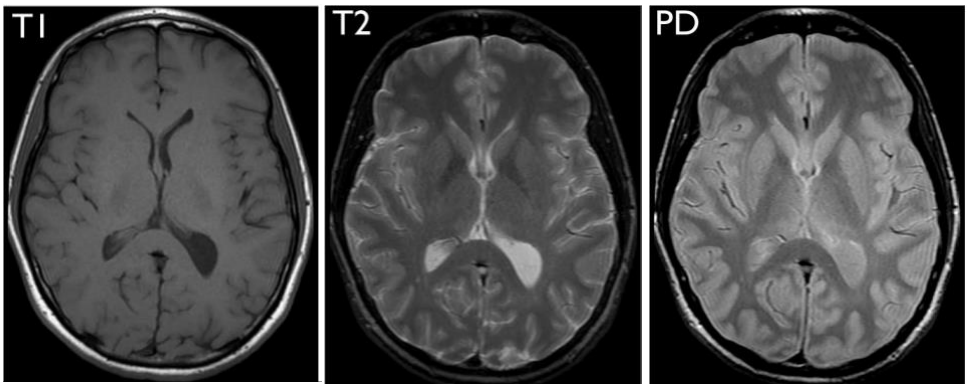


Figure 1.5 MRI images with different weighted contrast, T1-W, T2-W and PD. Reprinted from *https://radiology.ucsf.edu/blog/neuroradiology/exploring-the-brain-how-are-brain-images-made-with-mri*

Proton dephasing of M_{xy} magnetisation is not only influenced by the T2 property of the tissues but also by inhomogeneities in the magnetic field. Inhomogeneities are those areas within the main magnetic field with slightly lower or higher magnetic strength than their general surroundings. Such inhomogeneities produce faster dephasing, designated as $T2^*$, than that arising from the T2 characteristic of the tissue (true relaxation) and may tend to mask it. This occurs because the precessional frequency of protons is proportional to the strength of the magnetic field they have experienced, and inhomogeneity of the magnetic field results in relatively larger acceleration and deceleration of the protons, leading to a rapid dephasing, which produces the FID. Loss of the signals in a short time produce a small signal; besides, most of the tissues do not have enough time to attain their T1 and T2 relaxation states. To maximise the signal and allow measurement of the relaxation times, it is thus necessary to reduce the effect of the $T2^*$ to allow enough time for the overall tissues to reach their T1 and T2 relaxation states. Reducing the $T2^*$ effect can be done by rephasing or refocusing the spins so that they gradually return to the initial phase after the excitation pulse. After the excitation RF pulse, protons start to dephase in a fan formation. The slowest dephased protons represent the trailing edge, while the fastest dephased protons are located at the leading edge. Refocusing the magnetisation reverses the location of these fan edges, causing the faster dephased protons to become the trailing edge and the slower dephased protons to become the leading edge. In this case, the faster dephased protons will gradually catch up with the slower dephased protons so that two edges become

superimposed (in-phase) where the signal reaches its peak. This gradual growth of the M_z magnetisation mirrors the M_{xy} magnetisation gradual decay, generating the maximum possible signal. The refocusing of the magnetisation can be performed using either a refocusing RF-pulse in a spin-echo (SE) sequence or by using a refocusing gradient in the gradient-echo (GRE) sequence[20].

Several factors can result in magnetic field inhomogeneity and hence $T2^*$. One is inherent inhomogeneity in the main magnetic field, while another is magnetic susceptibility in different tissue components such as ferritin and hemosiderin; or materials such as gadolinium (Gd) that are susceptible to magnetism and become magnetised in an applied magnetic field. Magnetic susceptibility is defined as the ability of the applied magnetic field to affect active nuclei and magnetise them. The magnetic moment of susceptible materials affects the external magnetic field, whether in a positive or negative way, resulting in local increases or decreases of the magnetic field, respectively. Tissue susceptibility to the magnetic field cannot be fully cancelled in the GRE sequence; unlike that seen in the SE sequence, the signal decay is dictated by both the T2 and T2* in GRE. Any T2-W image obtained by a GRE sequence is thus T2*-W rather than T2-W[20].

Heavy T1-W contrast image can be produced using an inversion recovery (IR) sequence, one of the SE sequences, which rephases the magnetisation using a refocusing RF-pulse. The IR sequence begins with a 180° inverting RF pulse, which flips the magnetisation to the negative B0, allowing the magnetisation vector to begin recovery from a full inversion and thus enabling greater T1 time differences between different type of tissues and maximising the signal producing the heavy T1-W contrast image. Several different image contrasts can be acquired using this sequence (e.g. T2-W and PD-W), although the TR is always long, the flip angle determines the degree of saturation and hence T1 recovery, while the TE, as always, controls the amount of T2 decay. Also, pathology weighted image can be generated using the IR sequence, by nulling signals from certain tissue so the adjacent pathology to that tissue or lesions within it can be easily recognised. Suppressing signals of specific tissue can be achieved by employing the excitation RF-pulse while that tissue has not recovered (no M_z magnetisation component), designated as the nulling point for that tissue or

inversion recovery time (TI). Fluid attenuation-IR (FLAIR) is the most commonly used sequence in brain imaging. It produces a T2-W image contrast while suppressing the CSF signal, which illuminates the pathology[20], Figure 1.6.

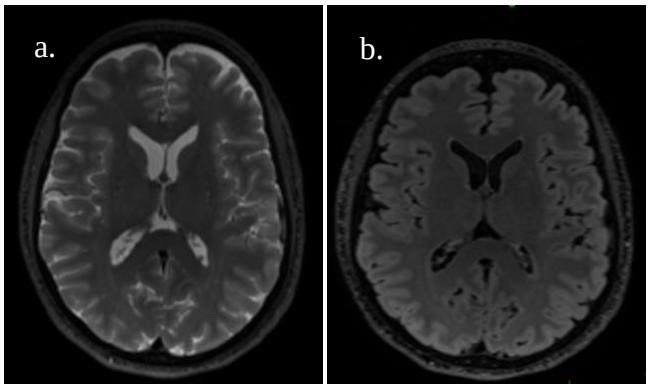


Figure 1.6 a) T2-W and b) FLAIR MRI images

1.2.3 Contrast agents in MRI

Although the intrinsic T1 and T2 relaxation processes in tissues cannot be changed, they can nevertheless be influenced. Introducing magnetically susceptible materials into tissue results in fluctuations of the local magnetic field that enhance the relaxation process of the magnetisation, which in turn reduces the relaxation times (T1 and T2) of nearby protons. This process enhances the contrast between tissues, and such materials are thus called contrast agents (CAs). CAs can be either a) formed endogenously, such as the de-oxygenated blood used in functional MRI (fMRI) and the magnetically labelled blood used in arterial spin labelling (ASL), or b) injected exogenously such as gadolinium (Gd). The accomplished contrast enhancement is depending on the susceptibility degree of the used CA and its concentration (in mmol/kg). The degree of the susceptibility of tissue (the magnetic behaviour of that tissue) is related to the movements of the electrons in its atoms, which induce magnetic moment and thus influence local magnetic fields. Inside the atom, in addition to the nucleus's spinning movement around its axis, electrons exhibit two types of movement: orbiting the nucleus and spinning about the axis. Electrons are organised in opposite polarities in the energy shells of an atom. In case of equal numbers of electrons, the opposing polarities cancel each other out to create a zero magnetic moment, known as a diamagnetic state. In contrast, atoms with partially filled

shells will have unpaired electrons that generate a net magnetic moment; this introduces positive susceptibility, whether paramagnetic, ferromagnetic, or superparamagnetic.

Paramagnetic exogenous CAs, often gadolinium-based CAs, demonstrates low positive susceptibility, are the most used in neuroimaging. Gd does not provide signals directly; instead, its magnetic properties impact the relaxation times of the surrounding tissue molecules. The effect of the CA on tissue relaxation times per molar concentration is called the relaxivity, and this is specific for each agent. As the concentration of the CA in mmol/kg increases, the generated effect also increases. Such paramagnetic agents reduce the T1 of nearby tissue molecules, thus increasing the signals on T1-W imaging so that they appear brighter (positive effect). They also reduce T2 and T2*, and on T2-W and T2*-W images, such signal reductions appear dark (negative effect). Since the T2* and T2 time of the biological tissues are much shorter than their T1 values, the impacts of the same concentration of Gd produce a more significant reduction of T1 compared to T2* and T2. Accordingly, a higher concentration is needed to produce significant reductions of T2* and T2[20].

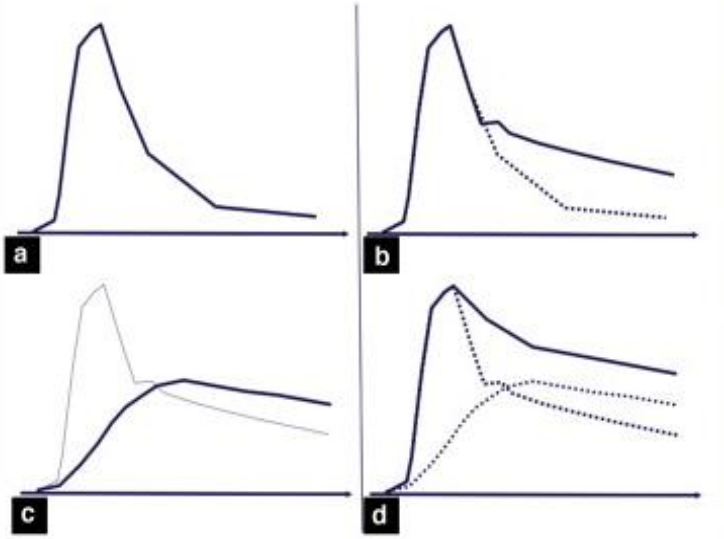


Figure 1.7 Changes in tissue enhancement following injected contrast agent with time. a: due to the bolus arterial first pass, this curve would not be seen in-vivo, because the contrast agent is not eliminated immediately but recirculate in brain vasculatures with intact blood-brain barriers (BBB), which is demonstrated on b) the curve represents the total tracer concentration in the blood (first pass plus the recirculation in the vasculature) and it represents the typical arterial input function curve (AIF) in-vivo. c: Uptake curve in tissue with a breakdown of the BBB. Note that the second part of the curve decays much slower (solid line) because the contrast agent molecules leak and accumulate in the interstitium and do not decay until the bolus is eliminated from the interstitium. d: Tissue enhancement (solid line) due to the presence of contrast agent in the vasculature plus that in the interstitium. Reprinted from Perfusion and vascular permeability: Basic concepts and measurement in DCE-CT and DCE-MRI. Diagn Interv Imaging. 2013;94(12):1187-1204. doi:10.1016/j.diii.2013.10.010.

The tissue concentration of the injected CA changes with time[22], as shown in Figure 1.7. The injected bolus requires ~20sec of transportation before the concentration in the blood circulation increases to form a peak (Fig. 1.7a). This peak represents the first pass of the bolus in the blood circulation with the highest CA concentration. Blood vessels in the central nervous system (CNS) have a blood-brain barrier (BBB) with selective permeability; thus, in normal circumstances, the CA cannot cross through the intact BBB. The next portion of the curve (Fig. 1.7b), therefore, is due to bolus recirculation in vasculatures as the BBB prohibits the contrast agent molecules from crossing into the brain tissue interstitium. In contrast, in brain tissue with dysfunctional BBB (e.g. gliomas), the curve has a completely different shape (Fig. 1.7c), due to bolus recirculation in the vasculature as well as bolus leakage and accumulation in brain tissue interstitium, which slowly appear in brain tissue and accumulates over minutes.

About 90 minutes after CA injection, the concentration is steadily reduced by excretion through the kidneys. In general, the first pass of the CA can be measured using fast imaging, which enables the assessment of the vessels qualitatively, or quantitatively using MR-angiography (MRA) or MR-perfusion, respectively. The second part of the curve represents the contrast extravasation / recirculation and can be assessed qualitatively or quantitively, using conventional T1-W image post-contrast or dynamic contrast-enhanced (DCE) methods, respectively.

Most of the CA is excreted by the renal system and hence affects this system. Normal kidney function excretes Gd rapidly without affecting the patient, but the half-life of Gd in the body increases for patients with impaired kidney function. There is also an association between the usage of Gd CA in MRI procedures in patients with renal insufficiency and nephrogenic systemic fibrosis (NSF) complications[23]. Gd is, therefore, contraindicated for patients with renal insufficiency. The glomerular filtration rate (GFR) is used to estimate the effectiveness of the renal function, based on laboratory blood tests. The kidney function can be categorised using GFR as a) normal kidney function, GFR > 60 ml/min/1.73m2; b) a moderate reduction in kidney function, GFR 30-59 ml/min/1.73m2; or c) severe reduction in kidney function, GFR < 30 ml/min/1.73m2.

Gadolinium is a heavy metal ion with free electrons that tends to accumulate in tissues. Fortunately, there are substances able to bind to these toxic ions and reduce its toxicity, known as chelates. The binding of gadolinium ions to chelates forms low molecular weight water-soluble agents that can be safely excreted by the kidneys. Various chelate structures form different agents of varying safety. Carrier ligands (chelates) may be linear or macrocyclic molecules, and these can also be subdivided into ionic or non-ionic types, creating four groupings. Stability of the selected gadolinium chelate is a crucial safety consideration, as this affects the way in which the chelate enables excretion of the toxic gadolinium. Unstable molecules, where the gadolinium easily separates from the chelate (de-chelation), can thus result in gadolinium deposition in the body. The European Medicines Agency (EMEA) has classified Gd CA according to the risk of de-

chelation into high- medium- and low-risk groups. Linear non-ionic agents are the riskiest, as they are most likely to separate from the Gd3+ ions, while linear ionic agents offer a medium risk of releasing free Gd3+ into the body. Macrocyclic agents are the safest, that is, the least likely to de-chelate from Gd. The EMEA Committee recommends (https://www.ema.europa.eu/en/medicines/human/referrals/gadolinium-containing-contrast-agents) that linear Gd-agents should be suspended except gadoxetic acid and gadobenic acid should only be used for liver scan in situations where they are necessarily required. The use of low-risk macrocyclic agents can be continuing to be used, with caution, considering the clinical risk to benefit ratios.

1.3 Advanced MRI

The MRI image contrasts discussed in the previous sections are conventional images that provide morphological information about tissues with excellent contrast and resolution. Advanced MRI permits more precise assessment and facilitates the understanding of normal physiological and pathophysiology processes. The most commonly used advanced MRI techniques for glioma diagnosis include diffusion-weighted imaging (DWI), MR spectroscopy (MRS), and MRI-perfusion, with dynamic susceptibility contrast (DSC), is the most frequently utilised format, along with dynamic contrast-enhanced (DCE) and arterial spin-labelling (ASL).

1.3.1 Diffusion-weighted imaging (DWI)

Diffusion-weighted imaging (DWI)[24,25] produces image contrast based on the mobility of particles inside tissues. Diffusion is the random movement of the molecules within and between various tissues in the body; these moving particles thus change their locations over time. In addition, the molecular diffusion inside living tissues encounters different degrees of restriction depending on the various structures of those tissues or the presence of any pathology. MRI can be used to track the diffused protons in tissues by employing two gradients following the excitation RF pulse that have the same magnitude but in opposing directions, thus dephasing and rephasing the spinning protons. Unlike a typical gradient,

these are stronger in magnitude (G) and last for a longer duration (δ), being separated by a time interval (Δ). The strength and timing of the diffusion gradient is thus represented by a factor known as the b-value, measured in s/mm², which is given by

$$b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3) \qquad \text{Equation 1.2}$$

Stationary protons return to the same phase as they dephase and rephase, and thus produce a strong signal. Moving protons, however, end up out of phase at various degrees in a manner directly proportional to the molecule displacement across a given area of tissue per unit of time. This is because the moving protons do not experience the equivalent oppositional gradient (the second gradient) due to their displacement and hence produce lower or no signals. Any movement type may contribute to this diffusion of signals, including blood flow at the capillary level (perfusion). True diffusion, which reflects the slow motion of the protons in the intra- and intercellular tissue, is thus best detected by using high b-values, as the effect of blood perfusion becomes negligible at high b-values[26]. A form of diffusion imaging called intravoxel incoherent motion (IVIM)[26,27] has also emerged that can provide information regarding perfusion and diffusion based on measuring the signals at multiple low and high b-values.

Generally, the protons can move randomly in three dimensions (3D) X, Y and Z. Therefore, MRI needs to sensitise the diffusion gradients across the three cartesian planes and average the acquired signal over the voxel. On DWI the free diffusible protons show high signal loss and appear dark (e.g. CSF), while the restricted protons will show as bright or various shade of grey, depending on the amount of the local diffusion restriction.

DWI has a long TE to accommodate the diffusion gradients, generating a T2 predominant MR image. The long TR and a TE with a b-value=0 s/mm² provides a T2-W image while increasing the b-value changes the image contrast from T2 to DWI. Subsequently, pathological tissues with long T2 decay, which are usually bright on T2-W images, remain bright on the DWI, causing artefacts that are referred to as T2 shine through[24,25]. Bright areas on the DWI are thus often

confusing, as it is difficult to determine whether they are areas of a restricted diffusion or just T2 shine through. An apparent diffusion coefficient (ADC) map facilitates differentiation of T2 shine through artefacts, and true diffusion restriction; as the ADC map estimates the moving protons' displacement across the tissue area per unit of time independently of the baseline T2 decay of the DWI, hence reflecting the effects of the diffusion itself. On the ADC map, the freely moving protons show as bright areas (e.g. CSF), while the restricted protons remain dark. Thus, if an area demonstrates bright signals on DWI but remains dark on the ADC map, this confirms it as a true diffusion restriction. However, if an area shows bright signals on both the DWI and ADC map or bright on DWI while no change on ADC, this suggests that it is simply a T2 shine through artefact, illustrated in Figure 1.8.

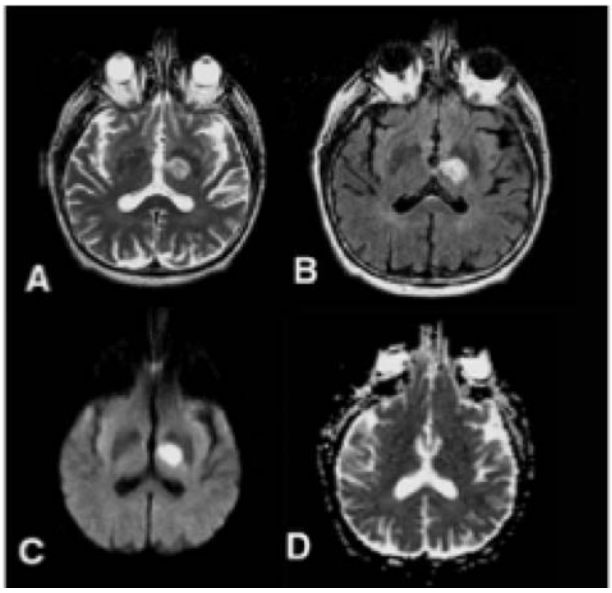


Figure 1.8 T2 shine-through effect. A) hyperintense lesion in the left hemisphere on T2-W image; B) also appears bright on the FLAIR image; C) appears bright on DWI while no changes appear on D) ADC map. reprinted from Diffusion-weighted imaging: Basic concepts and application in cerebral stroke and head trauma. Eur Radiol. 2003;13(10):2283-2297. doi:10.1007/s00330-003-1843-6.

1.3.2 Magnetic resonance spectroscopy (MRS)

Magnetic resonance spectroscopy (MRS)[28] is a technique that allows analysis of the chemical composition of tissues, in particular the presence and concentration of various metabolites. As previously mentioned, MRI signals can be acquired from the MR active nuclei, which have an angular moment. Each of these MR active nuclei has its gyromagnetic ratio, resulting in the production of different precessional frequencies under the application of the external magnetic field. This frequency variation enables the imaging of various MR active nuclei based on tuning the RF coil to transfer electromagnetic energy at the appropriate

frequency. MR active nuclei such as hydrogen, which are found in molecules with different chemical environments, also express varying precessional frequencies due to the impacts of the electron shields of the neighbouring nuclei within a given molecule. This diversity in precessional frequencies among hydrogen atoms in different molecules is referred to as a chemical shift, and it is on the order of parts per million (ppm) equal to 10^{-6}. The ppm scale is used instead of a Hertz scale, as it is not affected by the strength of the used magnetic field. The chemical shift effect forms the basis of MRS, with metabolite quantified in terms of its specific chemical shift (($\delta(ppm)$):

$$\delta(ppm) = \frac{fsamp - fref}{fref} \qquad \text{Equation 1.3}$$

Where $fsamp$ symbolise precessional frequency of the metabolite under investigation; $fref$ symbolise the precessional frequency of the reference compound, which is tetramethylsilane (TMS). The reference point at zero ppm is represented by the resonance frequency of hydrogen atoms in TMS ($Si(CH_3)_4$) because the hydrogen atoms symmetrically distributed and experience the same environment. Using TMS as a reference point, thus, the resonance signal from the other metabolites hence shifted downfield to TMS.

MRI measures the signals from the abundantly available hydrogen/proton atoms in a living body in different ways, reflecting the various magnetic properties of these tissues (T1, T2, diffusion). MRI signals acquired in 2D or 3D in the time domain are converted to the spatial domain using a fast Fourier transformation (FFT), due to the use of gradients that encode the position in different frequencies. MRS, however, provides quantitative biochemical information about the scanned tissue acquired in the time-domain, without spatial information, and transforms it into the frequency domain. This forms a spectrum in which the x-axis represents the chemical shift of various metabolites in ppm, and the y-axis demonstrates signal intensity and the area under each peak is proportional to the concentration of that metabolite. MRS signals or spectra can be generated from hydrogen (^{1}H-MRS) as well as from the other MR active nuclei. For example, phosphorus (^{31}P) or carbon (^{13}C), which are referred to as ^{31}P-MRS and ^{13}C-MRS,

respectively. ^{1}H-MRS spectrum is the most widely used in clinical imaging in the brain, and it allows the quantification of protons in different chemical environments including N-acetyl aspartate (NAA), creatine (Cr), Myo Inositol (mI), and choline (Cho)[28]. These metabolites (see table 1.1) are involved in cellular metabolism, and thus MRS enables assessment of tissue metabolism, with changes in amplitude providing clinically useful information to aid diagnosis. The concentration of these molecules is generally low relative to the preponderance of water and fat molecules, however, making it important to eliminate fat and water signals to allow observation of the metabolites of interest. Fat can be avoided by putting the MRS voxels within the brain tissue at a point away from fat and bone marrow. At the same time, water suppression is achieved by either chemical shift selective (CHESS) or IR (Inversion Recovery) techniques[28,29]. The 1D spectrum of the MRS is acquired using single-voxel spectroscopy (SVS) or multi-voxel chemical shift imaging (CSI), and table 1.2 summarises the main differences between them. As in MRI, TE impacts the observed molecular information in the MRS spectrum: MRS with a short TE observes metabolites with short and long T2s, while MRS with a long TE monitors only those metabolites with long T2s[28,29].

Table 1.1 Table.1.1 Summary of the detectable metabolites from 1H-MRS spectra. Reproduced from https://ispub.com/IJRA/17/1/17730

ppm	Metabolite	Properties
0.9-1.4	Lipids	Products of brain destruction
1.3	Lactate	Product of anerobic glycolysis
2	NAA	Neuronal marker
2.2-2.4	Glutamine, glutamate	Neurotransmitter
3	Creatinine/GABA	Energy metabolism
3.2	Choline	Cell membrane marker
3.5	Myo-inositol	Glial cell marker, osmolyte

Table 1.2 Single voxel and multi-voxel method comparison. Reproduced from *http://mriquestions.com/single-v-multi-voxel.html*

	Single Voxel (SVS)	Multi-voxel (CSI)
Operator set-up	Fast and easy	A little harder and slower
Shimming	Limited volume of interest allows very good shim to be obtained	Difficult to shim well over entire region
Spectral quality and peak separation	Excellent with high signal-to-noise, quantifiable	Lower signal-to-noise, problems with quantification
Spectral contamination	From adjacent tissues due to partial volume and chemical shift displacement effects	Bleeding of spectra from adjacent voxels due to chemical shift aliasing
Imaging time	Fast (3-5 min per voxel)	Slower, depends on resolution: 5-8 min for 2D, 7-15 min for 3D
Suitability based on size/ characteristics of lesion	Best for medium-sized homogeneous lesions in large organs	Best for lesions in small organs or for inhomogeneous lesions in larger organs

1.3.3 Perfusion MRI

Perfusion MRI is a functional imaging method that produces advanced image contrast; the resulting images allow for quantitative measurements that reflect relevant functional characteristics such as cerebral blood flow (CBF) and cerebral blood volume (CBV). These measurements mirror the regional haemodynamics of the micro-vasculature, which in turn represent cerebral metabolic demand[30] and tumour pathophysiology[31]. MRI methods for perfusion imaging can be classified into non-diffusible and diffusible perfusion, which use exogenous (e.g. paramagnetic gadolinium (Gd)) and endogenous (e.g. blood water) tracers, respectively. Non-diffusible perfusion MRI methods include dynamic susceptibility contrast (DSC), which exploits the magnetic field susceptibility caused by the first-pass of a high concentration of an extrinsic CA using dynamic T2*-W; and dynamic contrast enhancement (DCE), which uses extrinsic CA followed by dynamic T1-W. In contrast, the main diffusible perfusion MRI method is arterial spin labelling (ASL), which uses magnetically labelled arterial blood water as an inherently diffusible tracer. ASL thus provides a non-invasive imaging method that generates absolute CBF quantification. Chapter 2 provides more details regarding these MRI perfusion methods, as they are the main techniques used in this thesis.

1.4 MRI for glioma primary diagnosis

The biological changes inside gliomas, due to the involved molecular profile, induce changes mirrored on images as radiological features. Link these features to the biological changes (e.g. cellularity, neoangiogenesis) and molecular profile (e.g. IDH-mutation, 1p/19q codeletion) provide important information that can be used to characterise gliomas as a surrogate non-invasive biomarker for the gliomas genotypes. Conventional MRI images help to determine tumour morphology and location in brain tissue based on the use of various imaging contrasts, such as T1- and T2-weighted images (T1-W and T2-W, respectively), while fluid-attenuated inversion recovery (FLAIR) is used to provide information regarding oedema. Post-contrast T1-W (T1-c) allows the clinician to identify any regions of disruption to the blood-brain barrier (BBB)[3233]. Moreover, the advanced MRI images give complementary information regarding the tumour pathophysiology leading to better gliomas characterisation. For example, tumour cellularity can be evaluated using DWI while tumour proliferation and angiogenesis can be evaluated using perfusion MRI and/or MRS.

Gliomas grading is a cornerstone in treatment planning. Enhancement considered as HGGs predictor, nevertheless, enhanced T1-c can be misinforming as some LGGs exhibit contrast enhancement, while lack of enhancement can also be noticed in some HGGs[34]. Necrosis is a hallmark of glioblastoma, which is due to the increase of cell proliferation, mitotic activity with insufficient blood supply. The enhancement pattern in glioblastomas appears as irregular enhancing thick rind surround hypointense (non-enhanced) necrotic tissues. DWI allows determination of tumour grade based on cellular density, where HGGs have higher cell density than LGGs; accordingly, restricted diffusion with lower apparent diffusion coefficient (ADC)[35]. Angiogenesis development also reflects the tumour type[36], making it a key element of tumour pathophysiology[31]. Tumour growth exceeding 1 to 2 cm in diameter depends on angiogenesis to supply blood and nutrition to the tumour tissue, and this plays a critical role in tumour survival, development, and malignant transformation[37,38]. Measurement of the cerebral perfusion at the capillary level yields more accurate diagnosis [39–41]. Relative normalised cerebral blood volume (rCBV) from dynamic susceptibility

contrast (DSC) MRI is the most validated perfusion MRI parameter, thus the most widely clinically utilised. HGGs generally display higher rCBVs than LGGs[42]. However, oligodendrogliomas produce a confounding effect in that, and even when they are low grade, they express perfusion values that exceed the assigned threshold (≤1.75 mL/100 g to characterise LGGs from HGGs)[42,43]. Alternative MRI perfusion methods such as dynamic contrast enhancement (DCE) or arterial spin labelling (ASL) also appear beneficial[44,45]; unfortunately, these are not generally integrated into clinical practice due to more elaborate, non-standardised post-processing and validation. Metabolites changes using MRS gives valuable insight into the tumour metabolism and aid the diagnosis process. The MRS principal benefit is in deciphering the metabolic profile of gliomas and hence helping in their staging and in differentiating them from non-neoplastic lesion[46]. The main altered metabolites in gliomas are N-acetyl aspartate (NAA), total choline compounds (Cho), and in less extent, the total creatine compounds (Cr). Increased Cho levels accompanying reduced NAA levels suggest tumour malignancy, as these are attributed to increased cellular membrane turnover and breakdown of neuronal integrity, respectively. Presence of lipid/lactate peak correlates with a necrotic portion of the tumour and may associate with grade IV[47]. Another potential advantage of metabolic imaging like MRS lies in the better characterisation of oligodendrogliomas grade II, which have been commonly misclassified as HGGs due to increased rCBV levels[48]; Oligodendroglioma WHO grad II were differentiated form anaplastic oligodendroglioma WHO grade III using Cho/Cr ratio with 100% sensitivity and 83.3% specificity with a cut-off value of 2.33[49].

In the current WHO classification (2016), IDH testing is the first step in molecular characterisation. IDH mutation occurs early in gliomas with high frequency in WHO grade II and III astrocytoma, oligodendroglioma and secondary glioblastoma. Gliomas that harbour IDH mutation reported with longer survival time than those without such a mutation[3]. Presence of IDH mutation results in changes in metabolism, cellularity and angiogenesis, which in turn demonstrated as imaging features. On conventional MRI, frontal location, well-defined margin, with less frequent enhancement reported as radiological differentiation feature for

IDH-mutant gliomas[50–53]. In contrast, dense enhancement pattern is predictor sign for the gliomas with IDH-wt[52,54]. This different enhancement patterns on T1-c is due to the influence of the IDH-status on the VEGF expression. Enhanced brain lesion indicated dysfunction BBB which is impacted by the VEGF expression[55]. Higher VEGF level occurs in gliomas cells than that in normal brain cells, which promotes angiogenesis and is associated with higher WHO grade[55]. Besides, the lower expression of VEGF in gliomas is associated with mutated IDH[56]. High rCBV is also reported as a sign for IDH-wt gliomas[52,57]. High ADC has recently been highlighted as characteristic of IDH-mutated glioma[52]. Recent facilitation of detection of 2-hydroxyglutarate (2HG) by MRS has been reported with excellent diagnostic performance for the prediction of IDH mutant gliomas[58].

IDH mutation can be found in astrocytoma and oligodendroglioma, with concomitant codeletion of 1p/19q is a molecular signature of oligodendroglioma and can be used to differentiate astrocytomas from oligodendrogliomas. Gliomas' patients with 1p/19q codeleted have been reported with better response to chemotherapy which improves the clinical response and predicted long-term survival. Therefore, discriminate codeleted from intact 1p/19q has diagnostic, prognostic and predictive indispensable value. On conventional MRI, frontal location with an indistinct border is more likely to be IDH mutated with codeleted 1p/19q[54]. In contrast, the presence of T2-FLAIR mismatch has been reported as a highly specific imaging marker for IDH-mutation and intact 1p/19q [59]. T2-FLAIR mismatch sign is where the lesion appears hypointense on the FLAIR image except a hyperintense peripheral rim, as compared to the homogenous hyperintensity observed on T2-W images.

1.5 Image analysis

MRI images can be assessed qualitatively by visual inspection and/or in a semi- or absolute quantitatively way. Most clinical MRI acquisitions are assessed qualitatively, yet, as long as such assessment depends on human visual perception, there will be an inevitable degree of variability based on training, experience, and the individual judgement of radiologists. Quantitative MRI assessment has thus begun to be used more recently, based on the estimation of numerical maps of meaningful variables within a specific unit that can thus be

compared within and between patients. Quantitative MRI assessment uses the MRI scanner as a measuring device to measure various properties for each tissue voxel, making it important to ensure that the output of the quantitative MRI reflects the patient's biology rather than scanner fluctuations. This means that any quantitative measurements must be accurate (close to a true estimate), reproducible (repeatable measurements over a short period), and biologically relevant.

Image analysis involves three stages: pre-processing, processing and post-processing, all of which are prior to the diagnosis and clinical decision. The pre-processing step begins with a quick visual quality assessment of the raw data to facilitate preparation for further analysis by reducing sources of variation, such as adding correction for patient motion, removing non-brain tissue images, and removing any images with artefacts, thus, achieving the best possible results. After that, image processing involves quantitative maps generation in a specific unit by fitting the corresponding mathematical model(s) to the data. Finally, the image post-processing phase enhances the generated maps by using techniques such as co-registration to overlay them with high-resolution anatomical images. Segmentation of the region of interest (ROI), the cluster of voxels defining the location and extent of the lesions on the image, is also performed to increase the sensitivity and specificity of the measurement to disease addressed. ROI segmentation may be performed manually or automatically, being implemented on one slice to create a 2D image or in the form of slice by slice delineation of the tumour volume to create a 3D volume which enables more accurate assessment of the extent of the lesion. Commonly, the images are saved as 16-bit integers (Int16) ranging from -32768 to +32767; thus, following application of the image scaling, i.e. the values in a specific unit, the pixel values may retain negative values which will affect the measurements. As the post-processed images based on physical models, the negative values are generally not physically relevant but represent artefacts. Therefore, applying the tumour mask to estimate quantitative measurements values less than or equal to zero are excluded.

1.5.1 Texture analysis and radiomics

Maximum, minimum and averages over all voxels of the ROI are usually calculated and used for disease classification. The signal intensities are generally spatially variable across the ROI or VOI (volume of interest) creating the so-called image texture features. A single value, such as a maximum or average, is not sensitive to represent the precise status of a heterogeneous environment seen in tumours, and thus MRI images using the spatial variation of texture intensities permits better tissue classification. The textures of MRI images are derived from the signal intensities of the pixels/voxels throughout the ROI/VOI. These pixels/voxels are placed together in many various ways, which could represent a specific pattern. Statistical approaches can thus be used for characterisation of image textures that quantitively describe the arrangements of these intensities. Statistical texture analysis methods involve three orders: first-order, second-order, and higher-order statistics, all of which represent the statistical properties of the pixels; however, the second-order and higher-order methods, unlike the first-order method, take into account the spatial relationships between the pixels.

The first-order statistical method, commonly called histogram analysis[60], quantitatively measures the probability distribution of the signal intensities as a univariate variable in the ROI/VOI (Figure 1.9). The resulting histogram is a graphical representation of the frequency distribution of the signal intensities in the ROI/VOI, with its x-axis representing the signal intensities across a range of values (bin width), while the y-axis represents the pixel counts or percentages (bin length). The bin width of the histogram can be adjusted to accommodate different ranges of signal intensities, and the counts or percentages of pixels within a certain range can thus be estimated. The percentage (relative frequency) of pixels is calculated as the ratio of the pixel count within a range of signal intensities to the total pixel number in the ROI/VOI. A relative frequency histogram also sets the area under the histogram curve equal to one, thus removing any variation due to different tumour sizes[61].

Summary values from first-order statistics[62] describe the data in many aspects. These include: a) the central tendency of the data distribution, using the mean and median; b) the variability of the data, using standard deviation (SD) and

variance (var); c) the ways in which the values of the data are spread out, using minimum, maximum, mode, percentile/centile, inter-quartile range (iqr); and d) the shape of the data distribution, using kurtosis and skewness. The mode represents the most frequently repeated value within the data set. In contrast, percentile/centile scoring of the values in the data set (e.g. 95th percentile (95-tile)), identifies where specific values fall as compared to the other values within a data set. The inter-quartile range (iqr) measures the distance between the first quartile (Q1), or 25-percentile, and the third quartile (Q3), or 75-percentile, which represents the distribution of the middle portion of the data set. Kurtosis measures the probability of data points falling in the histogram tail, with platykurtosis indicating a short and broad tail with a flatter peak (kurtosis<0) and leptokurtosis indicating a long and narrow tail with a sharp peak (kurtosis >0). Skewness measures the histogram tails' asymmetry, where a skewness>0 shows an elongated tail on the right-hand side, while skewness<0 demonstrates an elongated tail on the left-hand side.

Any randomness in the data can also be investigated using histogram entropy, though care must be taken while calculating histogram entropy because it is affected by the histogram bin width size, which in turn determines the probability values inside that bin, as seen in equation 1.4. The z-score measures how far a specific value (e.g. median) deviates from the mean based on the standard deviation unit (equation 1.5). This standardises the distribution and enables the comparison of various ROIs with different kinds of variability. As a hallmark for heterogeneity inside the ROI, the slope of the cumulative distribution function (CDF) curve can be used as it accumulates all the probabilities less than or equal to a specific value, see equation 1.6.

$$\text{Entropy} = -\text{sum}(\text{probabilities}.* \log2(\text{probabilities})) \qquad \text{Equation 1.4}$$

$$\text{Zscore} = (\text{value} - \text{mean})/\text{SD} \qquad \text{Equation 1.5}$$

$$\text{CDF slope} = (95 - 10)/(\text{tumour 95tile} - \text{tumour 10tile}) \qquad \text{Equation 1.6}$$

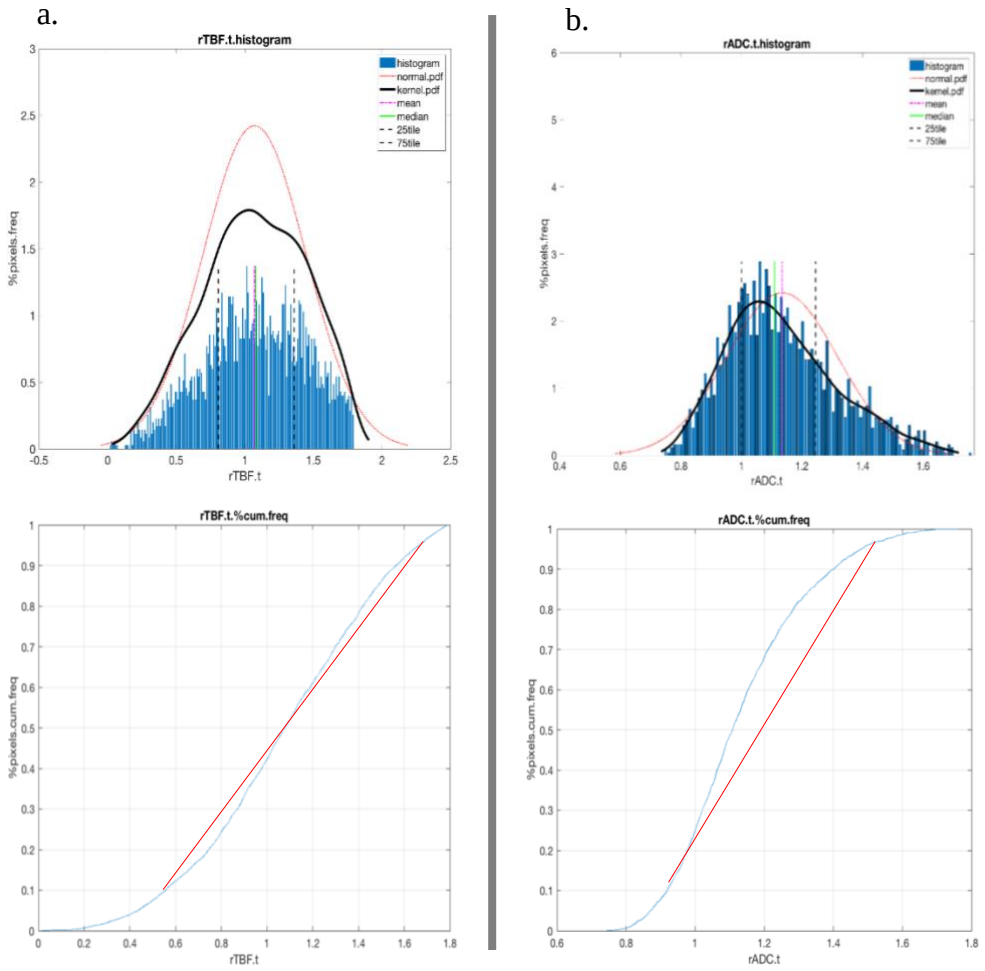


Figure 1.9 show the histograms and cumulative distribution functions (CDF) curves of (a) perfusion and (b) diffusion MRI maps for whole tumour volume of glioblastoma. The black curve demonstrates the histogram's probability density function (PDF). The red line on the CDF curve represent the slope between the 10th and 95 percentiles. The perfusion histogram (a) expresses a left skewed peak, while the diffusion histogram (b) shows a right skewed peak, both expressing the glioma's malignant transformation.

Combining the estimated values from the histogram analysis enables better characterisation of various image textures, just as identifying someone's face by combining their facial features is easier than recognising just one feature. This is the core idea of radiomics[60,63,64], which posits that combining radiological features can help to predict diseases. Similarly, just as different people have varying abilities to recognise facial features and match them to a person, different MRI methods offer different abilities to characterise diseases. A large number of

quantitative data are extracted from histogram analysis, most of them are highly correlated (creating hence *redundancy* in the final diagnostic or prognostic model) and/or could not distinguish the data according to the predefined classes (*irrelevant*). Both contribute to the so-called *model overfitting*, which is the noise that negatively impacts the performance of the model and, hence reduces its accuracy. Feature reduction is thus a major step in radiomics in the search to enhance model accuracy and limit the complex elements involved in prediction or classification tasks.

The final step in a radiomic pipeline is statistical analysis[63,64]. This step investigates the relationship between the final features set (input) and the predefined diagnostic class labels (outcome), such as the WHO grades of gliomas, using a regression model or machine learning (e.g. support vector machine). The performance of the regression model, which in turn reflects the radiomics performance, and can be reported using the so-called confusion matrix, (Figure 1.10). It is a table that sorts the predicted classes against their true classes. Each row expresses the instances in predicted classes while each column shows the instances in the true classes. This table, thus, depicts the true-positive (TP), true-negative (TN), false-negative (FN) and false-positive (FP) results. These four outcomes of the confusion matrix allow calculation of the summary statistics (equations 1.7 to 1.12) that describe the histogram parameter performance in classification. The F1-score reflects the classification accuracy for imbalanced sample size in each group, and the weighted average F1-score represents the overall model accuracy taking into account the contribution of each group to the final prediction accuracy. The positive predictive value (PPV) estimates the probability that a patient with a positive test result truly has the disease, while the negative predictive value (NPV) estimates the probability that a patient with a negative test result truly does not have the disease; sensitivity thus measures the true positive rate, representing the percentage of patients who test positive and do have the disease, while specificity is the true negative rate, as it represents the percentage of patients who test negative and who do not have the disease.

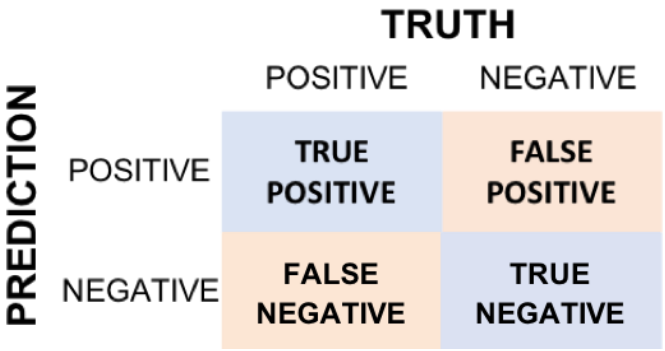


Figure 1.10 Confusion matrix, and its four outcomes.

$$F1 = 2 \text{ x } \frac{\text{precision x recall (or sensitivity)}}{\text{precision} + \text{recall}}$$ Equation 1.7

$$\text{weighted average F1} = \frac{\text{sum (F1n x sample size)of the classes}}{\text{total sample size}}$$ Equation 1.8

$$\text{PPV or precision} = \frac{TP}{TP + FP}$$ Equation 1.9

$$NPV = \frac{TN}{TN + FN}$$ Equation 1.10

$$\text{sensitivity or recall} = \frac{TP}{TP + FN}$$ Equation 1.11

$$\text{specificity} = \frac{TN}{TN + FP}$$ Equation 1.12

Chapter 2

Part of this chapter has been published in, Contrast Media & Molecular Imaging, 2018 [65].

2 Methodology 58

2.1 Tracer kinetics and tissue concentration-time curve basic concepts 58

2.2 MRI perfusion 58

2.2.1 Non-diffusible (exogenous) tracer 61

Arterial input function (AIF) 62

a. Dynamic susceptibility contrast-enhanced MRI (DSC) 64

b. Dynamic contrast-enhanced MRI (DCE) 67

2.2.2 Diffusible endogenous tracer 69

a. ASL main principle 69

b. Arterial spin labelling approaches 71

i. Continuous arterial spin labelling (CASL) 72

ii. Pulsed arterial spin labelling (PASL) 74

iii. Pseudo-continuous arterial spin labelling 79

c. Technical considerations for CBF quantification 81

D. CBF quantification models 83

i. T1 model 84

ii. General kinetic model 85

d. Improving the SNR 90

i. Hardware considerations 90

ii. Acquisition parameters and special techniques 91

e. Further parameters beyond the CBF 93

2 Methodology

2.1 Tracer kinetics and tissue concentration-time curve basic concepts

This chapter describes the main MRI methods for measuring perfusion, also called blood flow, or by definition the delivery of nutrients to the capillary bed. The central volume principle[66,67] describes the direct relationship between the cerebral blood flow (CBF), cerebral blood volume (CBV) in a given volume of tissue and the average time for the tracer to pass through the capillary bed in it, known as mean transit time (MTT), is defined as such (Eq.2.1):

$$MTT = \frac{CBV}{CBF} \qquad \text{Equation 2.1}$$

For a freely diffusible tracer, as its concentration slowly raises in the tissue, through arterial blood flow delivery, the tissue concentration (C_T) raises first with a negligible outflow by the venous blood flow. With time, the tracer molecules wash out via the venous blood flow until both arterial blood concentration (C_A) and venous blood concentration (C_v) of the tracer reach equilibrium: $C_A=C_v$. The higher the inflow rate, defined by local cerebral blood flow (f), the quicker the arterial and venous concentration will reach the equilibrium. MTT can also be seen as the required time to reach that equilibrium value.

The distribution of the tracer between tissue and blood after they reach equilibrium is described by the volume of distribution (λ). Therefore, for a freely diffusible tracer, which diffuses into all spaces of the tissue, λ=1. However, for an intravascular tracer, which will be distributed through a small fraction of the tissue space, λ < 1. Since the intravascular tracer is restricted inside the healthy brain to the intravascular compartment, assuming an intact blood-brain barrier (BBB), λ=CBV. Consequently, for the same blood flow, the MTT of an intravascular tracer is much shorter than the MTT of a free diffusible tracer. These physiological parameters (f and λ) can be estimated from a concentration-time curve that mirrors the kinetic of the chosen tracer.

In tracer studies, the essential measurements are tissue concentration over time $C_T(t)$ and the tracer concentration in arterial blood $C_A(t)$. Importantly, the system is assumed to be in steady-state , defined by a constant blood flow and neuronal activity. Models

of the tracer kinetics are generally connecting the $C_A(t)$ as an input and $C_T(t)$ as output. Kinetic models include several physiological parameters. Measuring these parameters is the goal of the kinetic analysis.

The ideal arterial concentration-time curve $C_A(t)$ is perfectly rectangular with a clear border between the bolus wash-in, plateau (equilibrium) and wash-out[68]. In other words, the C_A increase immediately, stay constant for a while then drop to zero. Non or little of the delivered tracer to the tissue volume cleared from the capillary bed. Subsequently, the $C_T(t)$ curve will increase due to the delivered blood volume to the tissue volume that carries tracer (Eq.2.2):

$$\Delta C_T = f C_A \Delta t \qquad \text{Equation 2.2}$$

The C_T curve initial slope is in this case proportional to *f* and the arterial concentration. As time goes on, the portion of the early delivered agent start to wash out from the tissue through venous clearance. If the arterial input is long enough, the rate of tracer wash-in by arterial blood flow is matched with the rate of clearance by venous blood flow, which is the equilibrium value, and the C_T curve reach the plateau. At equilibrium and since the intravascular tracer distribute through a part of the healthy brain tissue, $C_A=C_v=C_T/\lambda$. As the C_A concentration fall to zero, the C_T concentration will decrease with time. Assuming the tracer outflow rapidly from the tissue via the venous blood, the C_v concentration and C_T concentration remain in equilibrium, while the total concentration decreasing, which means the C_v equal to C_T/λ while C_T decrease. This clearance result in exponential decay of the C_T curve (Eq.2.3):

$$C_T(t) \sim e^{-ft/\lambda} \qquad \text{Equation 2.3}$$

Thus, each part of the C_T curve is sensitive to different physiological parameters; the initial slope sensitive to f, the plateau to λ and clearance to f/λ as a ratio.

A complete measurement of the C_T curve is rarely measured. Instead, various MRI perfusion methods focus on different aspects of the curve. In addition, the essential difference between the kinetic of diffusible and intravascular tracers impact the sensitivity of the curve to the physiological parameters. In a healthy brain with an impermeable capillary network, the diffusible tracer has a larger volume of distribution (λ) compared to the intravascular agents, thus, for the same blood flow, the MTT for

intravascular agents is shorter than that required for the diffusible tracer. Since pre- and pot-equilibrium of the C_T curve are the sensitive time intervals to blood flow while the MTT is the required time to reach an equilibrium value, thus, these transition periods are shorter for the intravascular agent. Subsequently, measuring the flow using the intravascular tracer is a matter needing to be deal with and overcome (i.e. a very narrow and sharp bolus that is shorter than MTT is needed and a fast-dynamic MRI to track every detail of the bolus kinetics).

In practice, the C_A curve is not a rectangular function, and the regions between the washing, plateau and washout are less clearly marked, which makes it difficult to directly relate the parts of the C_T curve to the physiological parameters. Thus, a more general mathematical model (Equation 2.4) has been developed to make it possible to draw a general conclusion about how blood flow, blood volume and MTT impact the C_T curve from an arbitrary C_A curve.

Indeed, the tissue C_T curve at a specific time point (t) can be defined as the result of the delivered agent up to that time via the local CBF (f) (see Equation 2.2), plus the remaining agent from the previous delivery (residue function, (R(t))). This represents the local impulse response function, which is the product of the local blood flow and the residue function [f.R(t)]. Thus, the tissue concentration at time t is the convolution of the input function with the impulse response function (Eq.2.4):

$$C_T(t) = f\, C_A(t) * R(t) \qquad \text{Equation 2.4}$$

Where $*$ represents convolution, C_T the tissue concentration-time curve, C_A arterial concentration or arterial input function (AIF), f the local CBF, R(t) is the residue function.

The residue function is a time function, defined as the agent that is delivered at t=0 and remains still there at time t. Assuming each molecule of agent has the same probability of entering to and washing out of the capillary bed, at the onset R(t=0) = 1, because there is no time for the delivered tracer molecules to leave; then, it decreases with time and become zero as the tracer completely washed out. Thus, R(t) represent the transit time distribution of the tracer molecules through a tissue, hence, its integral equal to MTT= λ/f.

The R(t) function hides the complexity of the transport and uptake of the agent. Thus, different agents have their own R(t) function with specified assumptions. For example, by assuming the tissue as a single well-mixed compartment, R(t) is described by an exponential decay (Eq.2.5):

$$R(t) \sim e^{-ft/\lambda} \qquad \text{Equation 2.5}$$

Understanding the R(t) function allows grasping two important characteristics of the local impulse response function: first, its initial amplitude is equal to f since R(t=0)=1. Second, the integral of the impulse response function is λ since the impulse response function is [f.R(t)] and the integral of R(t) is MTT= λ /f.

2.2 MRI perfusion

Measuring CBF can be achieved even using exogenous CA, such as paramagnetic Gd, or an endogenous agent (e.g. labelled blood water).

2.2.1 Non-diffusible (exogenous) tracer

In order to measure an accurate tissue C_T curve that is sensitive to the haemodynamic parameters using the exogenous tracer, there is need for rapid imaging with high temporal resolution, a narrow bolus and a measurement of the local arterial C_A curve, C_A (referred to as arterial input function (AIF)). The measurement of the intravascular agent kinetics is easily able to provide CBV, unlike CBF measurements. Mathematically, the integral of the convolution of two functions is equal to the product of their separate integrals. Thus, from Eq.2.4 we have Eq.2.6 and 2.7; Consequently, the integral of C_T curve over time is λ multiplied by the integral of the input function $C_A(t)$. This measurement is directly proportional to CBV, lacking for global scaling factor via AIF.

$$\int CT(t)dt = \int f.R(t)dt \;.\int CA(t)dt \qquad \text{Equation 2.6}$$

$$\int CT(t)dt = \lambda \int CA(t)dt \qquad \text{Equation 2.7}$$

In contrast, the estimation of CBF is a matter needing to be deal with and overcome. As previously mentioned, MTT is very short in the instance of intravenous CA. Thus, a

very narrow and sharp bolus that is shorter than MTT is needed. This narrow bolus produces rapid changes in the arterial concentration preceding the action of tissue concentration equilibrium. Thereafter, the tissue C_T curve becomes sensitive to the local CBF and reflects its speed. As a matter of fact, a fast-dynamic MRI is necessary to track every detail of the bolus kinetics, and hence reflect the local haemodynamics.

MRI introduces the agent concentration changes over time as dynamic signal intensity (SI) variation. The measured SI can be evaluated even as it is (non-parametric/model-free) or by parametric modelling. Assessment of the measurements from the SI time curve provide semi-quantitative indices that correlate with the haemodynamic changes arising from specific physiological conditions, but not reflecting them directly. In contrast, using the parametric modelling approach, where the SI is converted into a relaxivity-time curve, where the linear model is used to translate the relationship between the contrast concentration changes and the accompanying SI changes, hence producing a tissue C_T curve which is as close as possible to reflect linearity between tissue signal and CA concentration. Fitting the kinetic model with specific underlying assumptions into that curve enables the kinetic parameters that directly relate to the physiological changes to be estimated.

Arterial input function (AIF)

The tissue C_T curve is a result of both the AIF and the magnetic properties of the imaged tissue. To establish reliable tissue perfusion measurements (e.g. local CBF, CBV and mean transit time) the impulse response function needs to be estimated without the AIF. This is accomplished mathematically by deconvolution of the AIF from the tissue C_T curve. Thus, the actual shape of the arterial C_A curve needs to be estimated. Following CA injection in the antecubital vein, the bolus disperses as it is transported away from the injection site to the heart and toward the other organs, including the brain. This transport introduces a delay between the injection time and bolus entry into the scanned tissue. The bolus is delivered to the capillary bed as several instantaneous doses as if they are from multiple instantaneous injections. The measured AIF, therefore, represents an arbitrary shape that can be thought of as being the sum of these instantaneous injections, which helps to correct for the injection profile and the bolus delay and dispersion.

It is difficult, however, to estimate the actual AIF shape due to problems including: a) the signal changes with the bolus concentration non-linear inside large vessels that have a high CA concentration level, and b) the partial volume effect (PVE) when the AIF is calculated from small vessels near the tissue of interest to correct for bolus delay and dispersion.

For the dynamic susceptibility contrast (DSC), the best estimates of AIF are from voxels near the tissue of interest and totally outside of the arteries[69,70]. This AIF estimate represents a compromise between the bolus delay and dispersion and the PVE, while approximates a linear relationship between the contrast concentration and the corresponding tissue relaxation. AIF measurements that have some degree of inaccuracy still provide a good approximation for its shape, which means that the perfusion maps (e.g. rCBF, rCBV) provided are relative. It is worth mentioning that brain tumours studies typically express perfusion values as the ratio between the obtained value from relative perfusion maps from the lesion to that obtained from the contralateral normal brain tissue; this is named as the normalised relative value (nrCBF, nrCBV).

In contrast to DSC, DCE usually involves a compromise to generate dynamic data with adequate spatiotemporal resolution. Thus, besides the aforementioned factors that impact the accuracy of the AIF measurement, in DCE, the low temporal resolution results in additional bias in the AIF measurement. This is because slow data acquisition fails to capture the rapid changes in the arterial blood concentration and hence does not account for the mean transit time of the tracer to pass through the capillary bed. There are many approaches that have been used to assess AIF for DCE, and so far the best choice was found to be that determined for each patient individually based on reference from a venous structure [71], such as sagittal sinus, which is also reduces the PVE that noticed in small arteries. The underlying assumption behind this selection is that the area under the resulting curve from the venous selection is the same as the area under the AIF.

The suitable voxels for AIF are characterised by fast bolus arrival, a steep slope (sharp uptake), a tight peak, and large maximum peak and area under the curve (long washout). These suitable voxels can be selected manually or identified automatically.

Manual selection needs to be done by an experienced user and is subjective, which may affect the reproducibility of the measurement. On the other hand, an automatic method thus may overcome the limitations of the manual method, as it uses a searching algorithm to identify all suitable voxels.

The MRI perfusion methods based on bolus injection relate the changes in SI around the capillary network (due to the perturbations of the local magnetic field) to the relaxation changes inside the vasculature due to the passage of the CA. The CAs are magnetically susceptible materials that reduce the relaxation times, including T2, T2* and T1. Consequently, result in a signal reduction on T2-W and T2*-W images and signal enhancement on T1-W image. The former method is known as dynamic susceptibility contrast-enhanced MRI (DSC), whereas the latter as dynamic contrast-enhanced MRI (DCE). A detailed explanation of each method is presented in the following sections.

a. Dynamic susceptibility contrast-enhanced MRI (DSC)

DSC enables the capture of changes in the concentration of the injected CA over time. This manifests as a loss of signal due to the induced susceptibility effects, using dynamic T2*-W images. Thus, DSC-MRI does not measure the CA concentration directly, but indirectly via the induced changes in the T2* relaxation time due to the transition of the magnetically susceptible bolus. This provides the SI time curve, which follows the associated drop in signal over time with the increased CA concentration, and its passage through the capillaries of the target tissue. This curve used to be interpreted using summary parameters[72] (model-free), including for example, bolus arrival time (BAT), time to peak (TTP), or maximum peak (MP). Even though these model-free DSC parameters are calculated directly from the SI time curve without extensive post-processing, they are semi-quantitative and are interpreted without a physiological basis. Moreover, these parameters are affected by the physiological variations (e.g. cardiac output, vascular status and route to the tissue), and the manner in which the study was carried out (e.g. the injection rate, CA injected volume and MRI acquisition parameters).

In contrast, calculated model-based DSC parameters[73] can be interpreted on a physiological basis. Quantification of these parameters requires measurement of the

AIF. The SI time curve is transformed into a relaxivity time curve using the linear relationship between T2* and the CA concentration, where a reduction in T2* from its original value (T20*) is attributed to an increase in CA concentration, see Eq.(2.11) below, that derived from Eqs. (2.8 to 2.10):

$$S(t) = S(0)\, e^{-TE/T2*} \qquad \text{Equation 2.8}$$

$$R_2* = \frac{1}{T2*} \qquad \text{Equation 2.9}$$

$$R_2* = R_{20}* + r_2*.C(t) \qquad \text{Equation 2.10}$$

$$R_2* = -\frac{1}{TE} \ln\left(\frac{S(t)}{S(\text{Spre-bolus})}\right) \qquad \text{Equation 2.11}$$

Where S(t), is the SI at time t; $S_{pre\text{-}bolus}$ or S(0) is the SI prior to bolus arrival; R_2* is the transverse relaxation with the CA; $R_{20}*$ is the transverse relaxation without the CA; r_2 is the relaxivity of the used contrast agent; C(t), is the intravascular concentration of the CA.

The resulting area under the relaxivity time curve represents the CBV value. In order to fit the general kinetic model in Eq.(1), a suitable AIF needs to be defined. This provides calibrated measurements to the global perfusion and hence can determine the true tissue perfusion, based on the deconvolution analysis of the relaxivity time curve and the AIF curve. From this, the CBF and mean transit time (MTT) maps can be derived. MTT is defined as the value of the elapsed average time of the CA bolus within the volume (voxel) of the vasculature. Also, from the central volume principle[66,67], it is defined as the ratio of CBV to CBF.

Besides the general problems that accompany MRI perfusion methods that use intravascular agents (e.g. injection profile, dispersion and delay), which can be resolved using an AIF, DSC-MRI has its own sources of error. The DSC-MRI mathematical model assumptions involve a negligible T1 effect, and the BBB is intact. Thus, when the T1 effect ceases to be negligible (i.e. when it is expressed as an

increase instead of signal loss), the perfusion measurements' accuracy is distorted. The T1 effect can be removed by using multi-echo sequences[74,75]. Still, most DSC-MRI studies are performed using a single echo sequence, where the validity of the negligible T1 effect assumption depends on the sequence parameters,[76] particularly TR, TE and flip angle (FA). TR is recommended to be 1.5 s or less to avoid compromising SNR and temporal resolution and keep the T1 effect negligible. TE should be 25-35 ms at 3T, as a trade-off between signal drop (perfusion signal) and SNR. FA should be between 60°-70° to balance the desirable SNR and reduced T1 effect. An alternative approach to reduce the T1 effect is to inject a small pre-loading dose of CA (pre-load) [77], which is typically, one-quarter of the single dose, injected 5-10 min before the DSC-MRI[76].

The T1 effect can become a major source of error in instances of severe BBB dysfunction[77]. The tissue C_T curve of the DSC possesses an extra component, which is the recirculation part[68]. This appears as a small dip following the partial return of the curve to the baseline, reflecting the CA re-entering the target tissue. This recirculation introduces an overlap with the first passage of CA, but still the integral of this curve is proportional to the CBV, as the BBB is intact, and the CA is still circulating inside the vasculature. However, in case of a tumour, where the BBB disrupted, the CA leaks into the endothelium and remains there; this is referred to as the leakage effect. This results in erroneous measurements of CBV and an increase in the T1 effect. A useful approach to overcome this problem is to consider the early part of the tissue concentration-time curve (bolus first-pass) by using a gamma variation function[78] and/or preload[76].

Gradient echo (GRE) DSC-MRI is considered as a better choice for tumour imaging[76]. That is because the sequence is sensitive to a broad range of vessel sizes, including abnormal angiogenesis caused by tumours. However, this broad sensitivity could introduce macrovascular contamination and result in CBF overestimation. For this reason, Bayesian models[79], which are used in physiological microvascular mode, have been used to overcome this source of error.

b. *Dynamic contrast-enhanced MRI (DCE)*

Due to its higher spatial resolution and immunity to susceptibility artefact, DCE is sometimes favoured over DSC. In addition, DCE provides more interpretable physiological parameters that reflect the composition of tumour tissue, such as extravascular-extracellular space (EES) volume and microvascular permeability, by fitting the acquired dynamic data into a pharmacokinetic model (PK)[80]. Similarly, to DSC, DCE depicts the changes in the concentration of CA over time as a signal enhancement, due to CA-related shortening of the T1 of the tissue, using rapid repeated (dynamic) cine imaging. As the CA passes through the tissue of interest, it results in SI changes over time (S(t)), hence the relaxivity time curve can be determined. Quantifying perfusion metrics directly from the relaxivity time curve[80] based on a non-parametric model, such as the onset time (T0), time to peak (TTP) and maximum time (Tmax), is straightforward. Although these parameters can be derived simply, they do not directly reflect the underlying physiological process.

On the contrary, parameters that reflect the functionality of the vascular tissue more accurately can be estimated using the parametric approaches. The first step to getting quantitative physiological parameters from the acquired-DCE raw data is to translate the SI changes ($R_1(t)$) over time to tissue C_T curve, Eqs. (2.12, 2.13).

$$R_1(t) = R_{10} + r_1.C(t) \qquad \text{Equation 2.12}$$

$$R_1 = \frac{1}{T1} \qquad \text{Equation 2.13}$$

Where R_1, is the longitudinal relaxation with the CA; R_{10} is the longitudinal relaxation without the CA; r_1 is the relaxivity of the used contrast agent; C(t), is the concentration of the CA in both intravascular and EES.

Estimation of the tissue C_T curve from the relaxivity time curve requires knowledge of the two variables in Eq.(2.8), R_{10} and R_1. First, the original tissue T_1 (T_{10}) can be estimated from a separate pre-contrast sequence[81] or by using an assumed baseline T_{10}[82]. Second, the signal enhancement (T_1 reduction) is not as straightforward as that in DSC, because it depends on the sequence used; the most commonly used

sequence is the fast-low angle shot (FLASH) spoiled gradient echo Eq.(2.14). This sequence is able to provide high temporal resolution but compromises the SNR and the spatial resolution. The knowledge of these parameters besides the FA that is used in the acquisition provides a clear relationship between the signal enhancement and the CA concentration.

$$S(t) = S0 \frac{\left(1-e^{-\left(TR/T1\right)}\right)\sin(\theta)}{1-e^{-\left(TR/T1\right)}\cos(\theta)} \qquad \text{Equation 2.14}$$

This equation assumes $TE<<T_2^*$ (heavy T_1-W) and hence neglects the T_2^* effect, whereas S0 is the relaxed SI ($TR>>T_1$, and FA=90°), θ is the implemented FA, and S(t) is the SI at time t.

Once the tissue C_T and the AIF curves are determined, fitting a PK model to the estimated C_T curve is the second step towards determining accurate parameters that have physiological meaning. PK models describe the tissue C_T as various rates and volumes related to the used model. There are several PK models, and the selection depends on the required indices for the specific clinical application. In addition, the estimation of certain microcirculatory indices depends on the acquisition conditions, in particular, the temporal resolution and the total acquisition time[22]. This is because depicting tissue perfusion requires rapid imaging, while assessment of extravasation (CA leakage) requires long scanning times. Thus, it is important to adapt the selected PK model to the scan timing used[83] while designing the acquisition protocol according to the desired microvascular indices.

Almost all studies investigating DCE parameters in gliomas use compartmental models, particularly the modified Tofts and Kermode (mTK) model[84–86], due to their simplicity. The compartmental models describe the CA kinetics and exchange between the intravascular and EES spaces while assuming each of them as a compartment. These two compartments are well mixed; that is the arrival of the CA in each of them is instantaneous, hence at any given time, they have a uniform concentration of CA, and the outflow of CA for any of the compartments is directly proportional to its concentration. The mTK[87] model accounts for the intravascular contribution to the tissue concentration to produce four maps: 1) fractional plasma volume (Vp, in %), 2) fractional volume of the EES (Ve, in %), 3) forward transfer constant from blood

plasma, or permeability constant (Ktrans, in min^{-1}), 4) and reverse transfer constant to the blood plasma, or reflux constant (Kep, in min^{-1}). This model cannot clearly separate the intravascular transport of the tracer molecules relative to their exchange process between intravascular and EES. Thus, the Ktrans from the mTK model incorporate both plasma flow (F) and vessels permeability surface (PS).

Advanced models, such as the adiabatic approximated tissue homogeneity (AATH) model also known as St Lawrence and Lee model (L&L), allowed separate estimation of F and PS which respectively reflect the transport of the tracer molecules intravascularly and between the intravascular space and the EES; hence they reflect the underlying physiology more accurately and enable additional physiological parameters to be better estimated[80]. The L&L model describes the intravascular transport of the tracer molecules as a plug-flow, where the tracer travels through the capillary, defined as a tube, by flow and its concentration gradient specified along the z-direction, while the EES is treated as a well-mixed compartment. Besides, the tracer concentration in the EES compartment change in an adiabatic (slow) manner with respect to the intravascular change rate. This model produces eight maps: 1) The extraction fraction (E%), is the extracted fraction of the CA from the intravascular to the EES at the venous capillary level; 2) plasma flow (Fb); 3) mean capillary transit time (TC, min); 4) separate estimation of permeability (PS, min^{-1}); 5) influx with a clearance of the CA to the EES from the venous capillary end (Ktrans, min^{-1}); 6) Kep, 7) Ve and 8) Vp. However, this model demand higher data quality[88] (fast imaging and high SNR), which limits their use.

2.2.2 Diffusible endogenous tracer

a. ASL main principle

The aim of ASL is to measure CBF of a brain tissue voxel in a target image slice. Figure 2.1 illustrates the concept behind ASL. Fulfilling this, a 180 RF-pulse in implemented at the labelling plane; this is positioned prior to the target image slice and inverting the magnetic field of the blood water. The tagged bolus is moving into the tissue of the brain at a rate that is equivalent to the CBF. After sufficient time delay, which is also referred to as post labelling delay time (PLD) or inversion time (TI), this labelled magnetic blood water attains the slice of interest in a period indicated as transit time or the arterial arrival time (AAT). The slice of interest is that on the imaging plane,

which is acquired as a raw data. Afterwards, the perfusion measurements can be extracted by conducting the experiment again but without inverting the magnetisation of the blood water. Following the same PLD/TI, the control image is collected. Then, the ASL difference image can be calculated by removing the static tissue by subtracting the control- from the tagged- image. The variation between the signal values indicates the magnetically-labelled blood water following its transfer at TI. Hence this signals difference directly proportional to the local CBF[89].

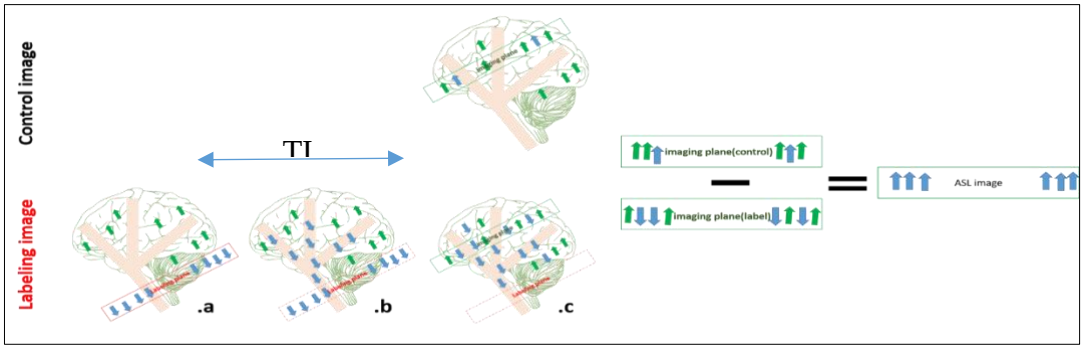


Figure 2.1 The primary mechanism employed in ASL is illustrated here. The diffusible tracer utilised is the magnetically labelled blood water. (a) the labelling plane applied (proximal) against the imaging plane where the flowing protons labelled employing inversion (blue arrows); (b) the labelled water exits the labelling plane during TI and initiates its distribution into adjacent tissue at the image plane; (c) acquisition of the labelled image. The control image is obtained without the labelling pulse to facilitate the extraction of the labelled blood water from the static tissue (green arrows). The tagged protons are the output once the labelled image is subtracted from the control image where the outcome signal differences are in direct proportion to the local perfusion.

In an ideal ASL experiment, the control image has fully relaxed blood water magnetisation ($M0a$). Once an inversion pulse is applied the blood water gains magnetisation ($-M0a$) that reaches the labelled image by the regional blood flow (f) during TI. Thus, the delivered blood volume is $f.TI$, and hence, the delivered tagged bolus into the labelled image is shown as:

$$-M0a\,.f.TI \qquad \text{Equation 2.15}$$

The difference in the signals between the longitudinal magnetisation of the control- and label- image (ΔM) is:

$$\Delta M = 2M0a\,.f.TI \qquad \text{Equation 2.16}$$

Considering 100ml of labelled blood delivered to the target tissue at the normal blood flow (60ml/min/100g of tissue) during time TI/PLD ~1 second, thus, the delivered magnetisation volume is $f.TI = 1ml$. Which mean the relative variation in the magnetisation would be ~1%. This example demonstrates one primary disadvantage of the ASL, which is the poor signal-to-noise ratio (SNR), an issue to be detailed and analysed later. As the variation in the signal is approximately 1%, it is vital that this is associated with the signals of the perfusion rather than parasitic signals. Thus, the same static tissue signals should be found for both the control- and labelled- image to ensure efficacious quantification.

b. *Arterial spin labelling approaches*

Rather than employing standard non-adiabatic RF-pulses, ASL employs adiabatic pulses. The former is implemented in an orthogonal manner to the primary magnetic field (static field (B_0)), where the resultant flip angle proportional to the employed fixed B_1 amplitude during a period of time. In contrast, the flip angle of the adiabatic pulse manipulated by a simultaneous change of the amplitude and frequency of the B_1 field using properly modulation functions. The B_1 field applied at a frequency (f_1) that gradually change from a large positive value (above-resonance) decreases to zero (on-resonance) and ends at a small negative value (below-resonance). Since the B_1 field starts with above-resonance frequency, this means the applied frequency not equal to the Larmor frequency (f_0) ,i.e. off-resonance, the B_0 at the rotating reference frame (B_z) not completely disappeared ($B_z = (f_0 - f_1) / \gamma = B_0 - (f_1/\gamma)$ since $f_0 = \gamma B_0$). This produces an effective field (B_{eff}) that orient at an angle to B_0 and B_1, where the NMV precess around it. Thus, the simultaneous variation of the frequency and amplitude of the applied B_1 change the orientation of the B_{eff} from below-resonance to resonance and beyond-resonance and hence invert the NMV, which is known as adiabatic inversion. There are benefits to this form of pulse regarding the degree of tolerance to B0 inhomogeneity, and it is also less sensitive to B1 miscalibration. Moreover, the

adiabatic inversion facilitates the formation of a sharp slice profile, hence enables a clear arterial bolus width[90].

As previously mentioned, both the control and labelled images must have the same static spins to ensure accurate extraction of the CBF. Nevertheless, the longitudinal magnetisation of the static tissue spins that in the labelling image are impacted by the magnetisation transfer effect (MT), as they experience off-resonance pulse. In such cases, the control image will have more static spins than the labelled image, as illustrated in Figure 2.2. As a result, the static tissue would generate biased images for the ASL difference images, and this leads to over-estimation of the CBF. To avoid this, the same MT-effect must be generated in the control images without tagging of the arterial blood water by applying the same inversion pulse but distal to the investigated slice. In such situations, the MT-effect that are produced through the labelling image, using the labelling pulse, cancelled out by that intentionally induced in control image[91–93].

The strategies used for ASL labelling are then classified into continuous arterial spin labelling (CASL)[89,94,95] and pulsed arterial spin labelling (PASL)[96–99]. The two strategies are different at the basic level regarding the extent location and the labelling plane's lasting duration.

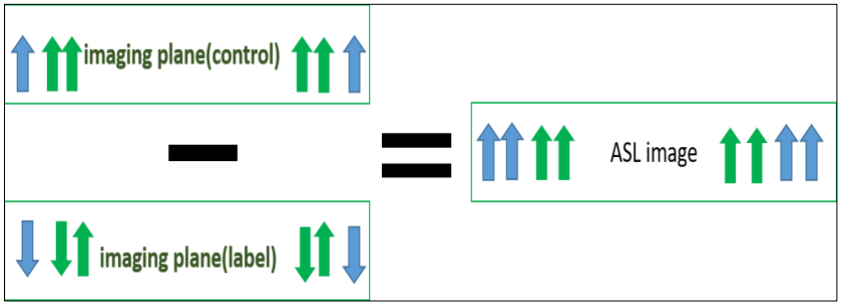


Figure 2.2 Illustrates the influence of the MT-effect of the perfusion assessment. The static tissue in the labelling image is decreased as a result of this impact and causes overestimation of the perfusion measurement. The blue arrows indicate tagged blood water and the green arrows the static tissue.

i. Continuous arterial spin labelling (CASL)

Using the CASL approach, the labelling plane is positioned at the level of the feeding artery (neck), where the flowing blood water that passes through this plane is constantly inverted. It has been called continuous inversion because of the RF-pulse

lasting for a longer duration than that of a typical RF-pulse (approximately 1 to 3 seconds)[100]. As noted earlier, ASL employs the adiabatic inversion method, CASL provides a continuous supply of the inverted arterial blood water by flow-induced-adiabatic inversion pulse, not the standard adiabatic pulse (explained above). The flow-induced- adiabatic inversion pulse, rather than using the frequency and amplitude modulation, it employs constant gradient through the flow direction accompanied by simultaneous constant RF pulse that applied over a small spatial location (tagging plane). Thus, the moving arterial blood water under a constant gradient induces changes in the local resonance frequency (above-resonance), as these blood water approach the tagging plane at Larmor frequency (on-resonance) they rotate to the transverse plane, as they move away from the tagging plane, they induce further changes in the local resonance frequency (below-resonance) resulting in blood water inversion[94]. This leads to a constant stream of labelled blood water while the gradient and the RF are turned on.

In early applications, a standard CASL control image was obtained through employing symmetrical labelling plane to that used in the labelled image but distal to the image plane (Figure 2.3) to compensate for the impact of MT[92]. However, this allowed management of the MT effect only for one image slice. To generate an equal MT effect per image a symmetrical positioning for the inversion plane around the imaging slice is required for the control- and label-image hence this method cannot be used for multi-slices acquisition.

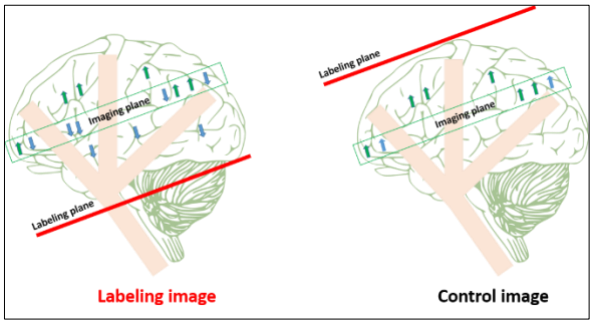


Figure 2.3 Illustration of early implementation of CASL, where the green arrows depict the static tissue and the blue arrows the labelled blood water.

For multi-slice acquisitions, two methodologies have been recommended to manage the MT effect. The initial one was proposed by Silva et al. and requires the use of two RF coils; one coil to generate a continuous adiabatic inversion (is formed of a small surface coil used for labelling and that is positioning at the level of the carotid artery)

while the other coil (head coil) is used a receiver coil for the image acquisitions[101]. Since the labelling coil and hence the transmitted RF labelling pulse physically away from the imaging slice, the MT effect can be avoided. Consequently, there is no need for the application of a distal labelling plane to the imaging plane for the control image. This method is also beneficial as it ensures that one artery alone can be labelled and so it is like territory mapping of perfusion[102]. Nevertheless, the specific absorption rate (SAR) generated by this strategy is high at the local level and would necessitate the use of extra hardware that is not normally accessible in setups for MRI scanners.

Consequently, Alsop and Detre recommended a second methodology to compensate for the MT effect in multi-slices image acquisition[103]. Creation of double inversion planes by implementing an inversion plane upstream that consists of a sine-modulated RF pulse accompanied by constant gradient identical to that of the labelling image. This would then result in the inversion of the flowing spins that cross the initial inversion plane and then be un-inverted when moving through the second. Theoretically, this would not produce any tagging effect while generates the same MT effects by using the same general RF power like that for the labelled images. Nevertheless, in clinical practice, this method lowering of the labelling efficacy, because not all the inverted spins would be un-inverted, thereby leading to a loss of the real labelling signals during the subtraction[103].

ii. Pulsed arterial spin labelling (PASL)

The PASL methodologies were developed to overcome the issues caused by CASL, including the raised SAR and the RF duty-cycle requirement because of the used of continuous RF pulses. Instead of using a continuous RF pulse to invert the flowing spins, PASL employs a slab-selective adiabatic inversion pulse, the standard one, that is positioned upstream relative to the imaging plane over a wide spatial band covering the feeding arteries. The labelled image can then be acquired following a determined delay time to enable the labelled flowing spins to perfuse into the imaging plane. To be able to generate a precise ASL difference image, the control image should be compensated for the MT effect without tagging of the flowing spins, and that to produce identical static spins to that in the labelling image. There are three primary PASL labelling approaches each employing a different labelling and MT effect compensation strategies.

The initial PASL technique was Echo-Planar Imaging and Signal-Targeting with Alternating Radiofrequency (EPISTAR)[97]; this is illustrated in Figure 2.4. For the labelling image, a spatial selective labelling pulse is used over a slab of 10cm, positioned 1cm upstream of the imaging slice. Owing to the tiny space employed between the imaging- and labelling plane, and because the slice profile of the inversion plane is not precisely rectangular, the imaging plane may be impacted by contamination generated from the tagging plane. Employing a saturation slab on the plane for imaging prior to the image acquisition helps to eliminate such contamination.

There are some challenges with generating a control image. Like the CASL, for the control image, an inversion slab that is the same as that used in labelling can be implemented 1cm distal to the imaging plane to mitigate the induced MT effect in the labelling image. Owing to the closeness of the inversion slab to edge of the image plane, the labelled venous blood water is generated and perfused into the imaging plane from above it. Such labelled venous spins lead to the ASL difference image having focal dark spots. Despite this technique mirroring the CBF variation, the precise association between the variation in CBF and the change in the intensity of the signal is unclear. Also, the magnetisation of the tagged bolus reduced and the SNR accordingly, as it experiences T1a relaxation over the long travelling time to the imaging slice[97].

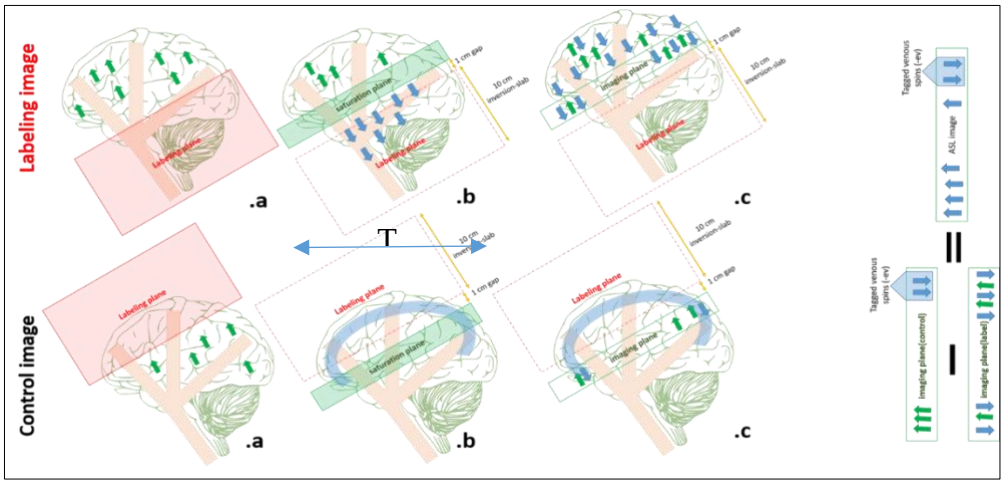


Figure 2.4 Illustration of echo-planar imaging and signal-targeting with alternating radiofrequency (EPISTAR). The tagged image is procured by: (a) employing a large labelling slab proximal to the imaging plane; (b) then application of a saturation slab to the imaging plane to eliminate the contamination from the labelling plane; (c) acquisition of the labelled image following delay time (TI) at which point the tagged blood water was travelled from the labelling plane, dispersed through the vascular system and perfused into the tissues of the image plane – this is the arterial arrival time (AAT). A control image obtained by: (a) employing a similar large labelling slab to that used for the labelling image but distal to the imaging plane; (b) then saturation slab to the imaging plane; (c) acquisition of the control image following the same TI where the venous spins that are labelled go into the control image. The ASL difference image (control - label) involves the labelled venous spins, which possess a negative sign from the subtraction and hence are appear as focal dark spots. The green arrows depict the static spins and the blue arrows the tagged blood water.

To avoid the drawbacks of the EPISTAR labelling method, Kim devised an alternative PASL strategy in 1995[98]. The strategy is referred to as flow-sensitive inversion pulse (FAIR), and it operates by subtracting a pair of inversion images. The first inversion image has a slice-selective inversion pulse (control image), while the second image does not (label image), as shown in Figure 2.5. The label image is acquired following a delay time (TI) with a non-slice selective inversion pulse. Resultantly, every spin inside and outside the image plane is inverted, and the size of the physical RF coil will determine the RF pulse's spatial extent. A benefit of this tagging method is that blood flowing into the imaging slice from various directions can be tagged. Thus, the venous blood flow is also tagged and enter the tagged image and generate artefacts, which appear as bright spots in the difference image rather than the dark spots as in EPISTAR. Furthermore, given that no gap exists between the image slice and the label plane, it is not necessary to perform a saturation pulse prior to acquiring the image.

The acquisition of the control image obtained using a similar inversion pulse as that used for the labelling image but on a selected slice. Therefore, similar to the entry slice

phenomena (define as a flow-related enhancement when unsaturated/fully relaxed blood water flow into the observed tissue volume that is partially saturated), the replacement of the inverted blood water with relaxed ones occurs as a result of blood flow, whereas the static spins stay inverted. Given that the same inversion pulse is employed in both the tag and control images, the MT effect typically cancels each other. The rectangular nature of the slice selective profile is not exact, and so it should be larger than the image slice. In this way, the image is acquired at the centre of the slice selective inversion, where clean uniform inversions exist across the whole image slice[98,104].

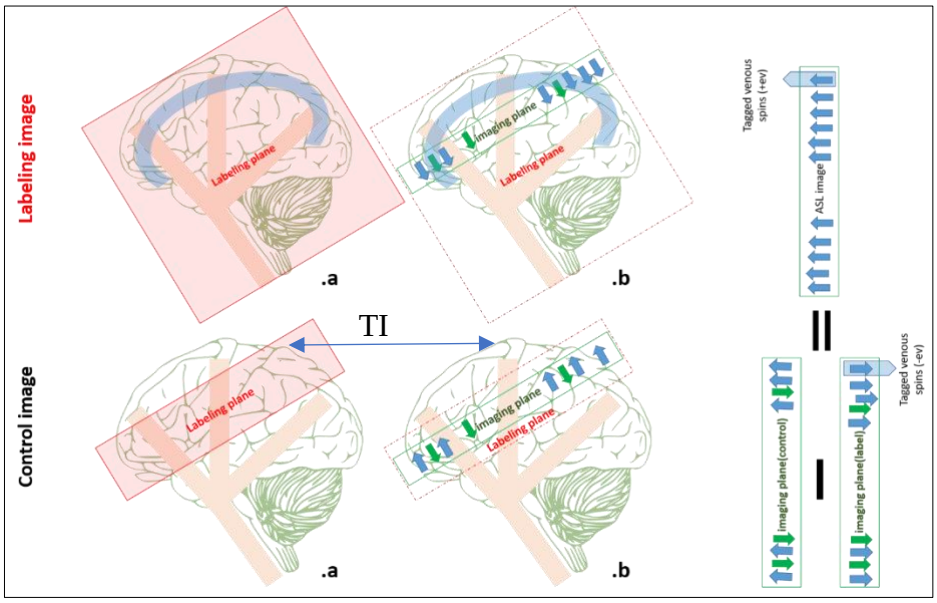


Figure 2.5 Flow-sensitive inversion pulse (FAIR) mechanism. Acquisition of labelling image needs a large inversion slab, (a) during application of non-selective slice pulse; and (b) the tagged image is acquired following a delay time (TI), and contains tagged flowing spins from veins and arteries. The acquisition of a control image needs the similar inversion pulse among a selected slice, where (a) the chosen slice that larger than the image slice, and (b) the control image is acquired following similar TI. The tagged venous spins appear as bright spots on the ASL difference image (control – label). static tissue spins are illustrated with green arrows while the tagged blood water is shown in blue arrows.

Proximal inversion with a control for off-resonance effect (PICORE) is another PASL technique that is derived from the EPISTAR method. PICORE method uses the same tagging method as that employed in EPISTAR[99] (Figure 2.6). For the control image, however, a shifted RF pulse (off-resonance) is used without a gradient (non-selective slice pulse). This off-resonance RF pulse has the same offset as that employed for the labelled image; this means that the control image will generate the same MT effect

without inversion. Furthermore, this off-resonance RF pulse does not result in spins tagging; consequently, none of the venous blood flow tagged as it happens among the EPISTAR- and the FAIR. This is an advantage for the PICORE- over both EPISTAR- and FAIR-labelling approaches.

As discussed above, each of the techniques is not entirely immune to issues such as sensitivity to the inflow from the distal end of the image slice, control image acquisition, and tag profile. In this case, selecting one technique rather than another should be based on the geometry of blood supply. If the blood supply enters the slice of interest from a direction that is known, the PICORE technique is the most reasonable to use. This stems from the fact that it generates no venous tagging. By contrast, if the direction of blood supply is not known (e.g., in a watershed area, which is the region that receives dual blood supply from furthest distal arterioles of two large arteries), FAIR is likely to be the most conservative method. This is because it tagging the arterial blood entering the slice of interest from various direction.

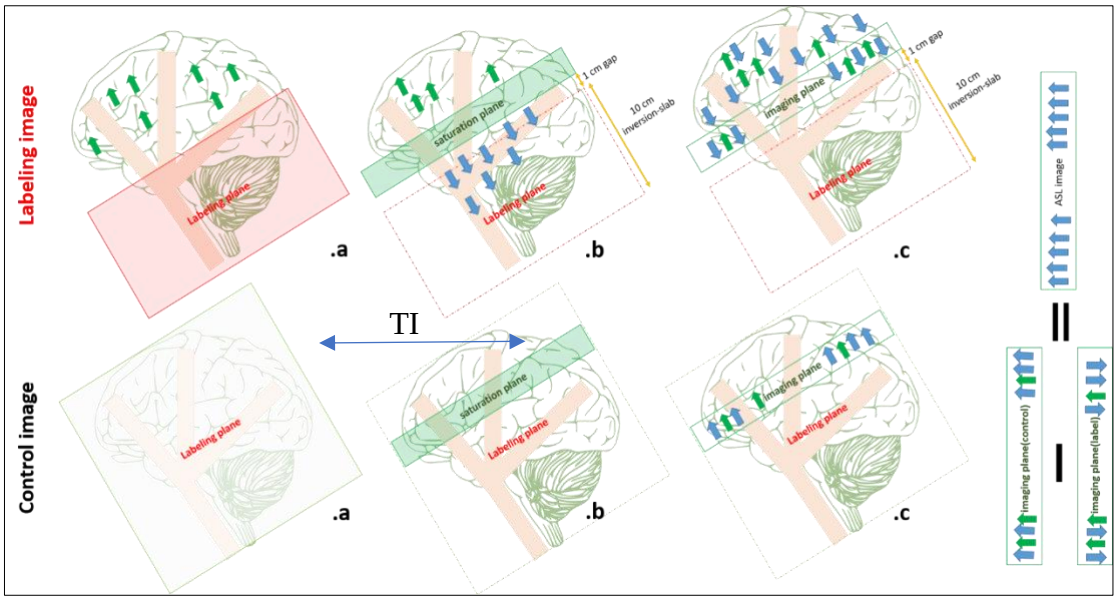


Figure 2.6 Mechanism for proximal inversion with a control for off-resonance effect (PICORE) mechanism. Acquisition of labelling image occurs with large inversion slab (a) during proximal tagging; (b) saturation slab is applied to imaging plane to eliminate tagging contamination; and (c) tagged image acquired following a delay time (TI), where the tagged blood water exists the labelled plane and begins diffusing from intravascular vessels into tissues at the image plane during arterial arrival time (AAT). Acquisition of control image occurs with (a) shifted RF pulse with zero gradient; (b) saturation slab; and (c) final acquisition of control image following TI. The ASL difference image (control – label), dissimilar to those associated with EPISTAR and FAIR, does not involve tagged venous spins. Static tissue spins are illustrated with green arrows. Tagged blood water is shown in blue arrows.

iii. Pseudo-continuous arterial spin labelling

In recent years, pseudo-continuous arterial spin labelling (PCASL)[95], which is a modified version of CASL that address its limitations, has been adopted. The PCASL labelling approach imitates the impact of the used continuous RF pulse in CASL by utilising long series of slice-selective RF pulses (~mean 1.5μT)[100] accompanied by a train of gradient pulses that is unbalanced (i.e. the net gradient area among time TR not zero) with a small mean (~1mT/m)[100]. Similar to the CASL, PCASL uses flow-induced adiabatic inversion, which needs to be successfully achieved for efficient labelling. This influenced by the flow velocity, the geometry of flow, and the inhomogeneity of the B1 and B0 field, where all of them impact the induced changes in the local resonance frequency of the flowing blood water. In PCASL phase error can be accumulated, due to the aforementioned factors, during the gaps between the RF pulses and result in degradation of the tagging efficiency. Thus, the spacing between the RF pulses should be as small as possible, recommended to be 1ms from the centre of one pulse to the centre of the next pulse. Also, in order to avoid aliased[95] tagging plane due to the periodic RF pulses, the slice profile of the RF pulses should be narrow by increasing the selectivity of the RF pulse employing a strong gradient (~10mT/m).

As indicated in Figure 2.7, the amplitude of the refocusing portion of the gradient pulses increase (unbalanced), which means that the flowing spins gain additional phase. To modify the phase of the RF pulse to become on-resonance with the accumulated phase of the flowing spin, the phase of the n RF pulse (Φn) must be $\Phi n = \gamma\, n\, G\, T\, Z$, where γ represents the gyromagnetic ratio, G represents the average gradient (approximately 1 mT/m in the labelling plane), T represents the spacing of the RF pulses, and Z represents the distance from the gradient isocentre to the specific point where the application of the labelling plane's gradient pulse takes place. For the control image, unlike the label image, the employed gradient pulses train are balanced, accompanied by alternating RF pulses (Figure 2.7). The balanced gradient generates a zero-average gradient between the pairs of RF pulses, while the alternating RF pulses result in a zero-average B1. Thus, these pulses produce the same level of RF power used for the labelling image but without labelling effect. As a result, the control image has the same MT effect as the labelling image[95]. Since the utilised gradient pulses for the labelling

and control images are different (unbalanced- vs balanced- gradients, respectively), this called unbalanced control in the literature.

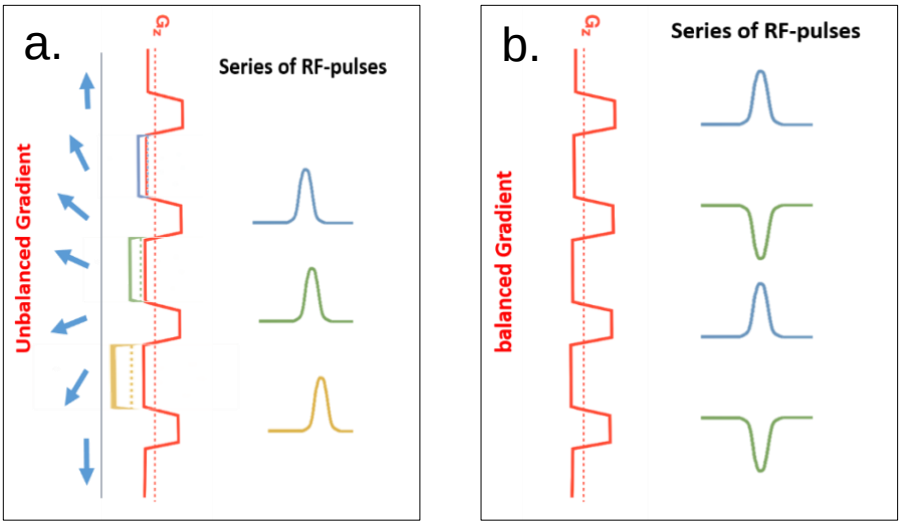


Figure 2.7 illustration of the labelling plane of pseudo-continuous arterial spin labelling (PCASL). The unbalanced gradient generates approximately 1 mT/m on average, while the RF pulses generate a B1 approximately 1.5μT on average. The blue arrows indicate the flowing blood water prior to and following inversion. Varying colour gradients and RF pulses reflect increases in amplitude and modified RF pulse phases, respectively. b) control pulse sequence of pseudo-continuous arterial spin labelling (PCASL). The balanced gradient generates zero on average, as well as the alternating RF pulses generate a B1 that is also zero-average. The green and blue RF pulses represent the alternating attitude of the RF pulses that lead to a zero-average B1.

PCASL surpasses PASL since it generates a greater SNR. The elevation in the SNR is because both PCASL and CASL are associated with longer labelling times and smaller transition zone width compared to PASL. The smaller transition zone width for CASL/PCASL relates to the fact their labelling plane is situated at the distal end of the PASL labelling slab (Figure 2.8). Consequently, PASL inverts the entire slab at the same time, while CASL/PCASL tags the flowing blood as it crosses the tagging plane. In other words, the tagged blood from CASL/PCASL does not take as long to move from the tagging slab to the image of interest (i.e., transition zone width). As a result, CASL/PCASL provide a larger volume of the tagged blood that experience a lower-level magnetisation loss from T1-relaxation compared to PASL, accordingly greater SNR[100].

PCASL is preferable to CASL due to three main reasons: a) PCASL lowers the specific absorption rate (SAR) as it uses a long series of slice selective RF pulses instead of long RF pulse for CASL; b) PCASL overcomes the MT effect because it is employing a strong selective gradient during the RF pulses; thus, they become off-resonance

relative to the imaging plane. c) RF pulses in PCASL can be generated using readily available RF amplifier systems[100,105], unlike the requirements of the continuous RF pulse used in CASL. This pulse type, particularly short and highly powered, requires a modified RF amplifier that cannot be accessed in almost all contemporary clinical imaging systems. Therefore, PCASL is recommended[100] by the International Society for Magnetic Resonance in Medicine (ISMRM) and the European ASL in Dementia Consortium.

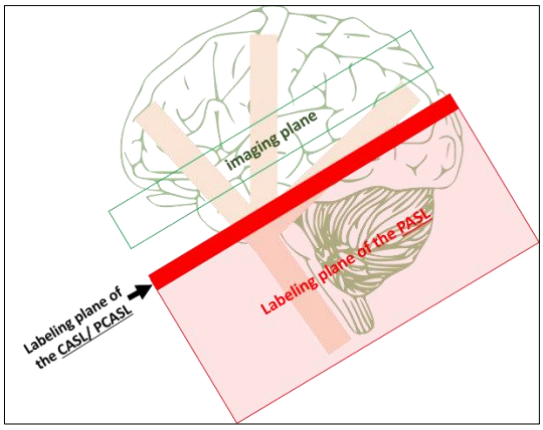


Figure 2.8 Difference between tag planes of CASL/PCASL and PASL.

c. *Technical considerations for CBF quantification*

A range of intrinsic parameters (physical and physiological) impact ASL image quality and can disrupt the link between the perfusion value and the obtained signal. Therefore, to select suitable extrinsic parameters and a labelling approach, it is necessary to consider and understand these intrinsic parameters. Such parameters include MT effect, transit delay (flow velocity-dependent), local T1 relaxation time, arterial bolus width and labelling efficiency. MT effects, as well as the strategies for minimising these effects in specific labelling approaches, have already been explored above.

Arterial transit time (also referred to as the arrival time) denotes the period needed for the labelled blood water to enter the imaging plane. This parameter is influenced by factors such as the size, tortuosity, and length of the vessels (Figure 2.9). The intensity of ASL signals tends to change based on the arterial transit time, and this is the case even when the same blood perfusion is provided. To reduce the sensitivity of the CBF quantification to arterial transit time, post labelling decay (PLD) was introduced[106], to give the necessary time for the tagged bolus to enter the imaging plane. During the

PLD period, some tagging is lost via T1-decay (Figure 2.9). Nevertheless, transit time and T1-decay are on the order of seconds, meaning that T1 and transit time are competing constant times that must be taken into consideration when determining the PLD.

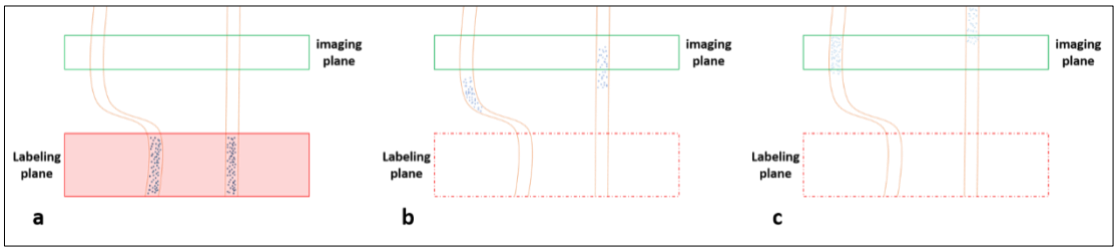


Figure 2.9 A drawing showing the impact of the vessel structure on the arterial arrival times, (a) during labelling, (b) during initial transit time after labelling, and (c) during second transit time after labelling. Lightening of shade indicates that the tagged bolus has lost tagging via T1-decay.

It is necessary to define the arterial bolus duration (bolus width) in order to add a delay to the pulse sequence and enable the tagged bolus to arrive. The bolus width duration is defined as the overall time taken for the tagged blood water to exit the tagging plane. In CASL/PCASL, the bolus width duration is considered equal to the RF pulse duration. In this context, the tagging band is a thin line with continuous labelling (long RF pulse), where the inversion of the flowing spins occurs when this line is crossed. Hence, after the RF pulse terminates, the tagged blood exits the labelling line. Resultantly, if the CBF increases, the amount of tagged blood water crossing the labelling plane increases. This is reflected in the ASL image as increased signal intensity. The converse is true for a decrease in the CBF.

Contrastingly, in PASL, the bolus width duration is not well-defined. The tagging band, which operates over a specific volume with a short duration, generates a fixed volume of the tagging blood water, irrespective of the variation of the CBF. To differentiate between low and high CBF, it is possible to apply a saturation pulse over the labelling plane following blood water labelling, referred as the initial inversion time (TI1), thereby cutting off the tagging tail (Figure 2.10). This generates a defined tagged bolus not in space but in time. Followed by delay time (TI) for the tagged bolus to reach the imaging plane. Hence, if the CBF increases, an increase will be observed in the volume of labelled blood that left the labelling plane prior to the cut-off. This method is referred to as quantitative image of perfusion using a single subtraction (QUIPSS II)[107]. Yielding

a quantitative flow image with this method relies on the satisfaction of the following conditions: firstly, the saturation pulse needs to be applied prior to the moment before all the tagged spins have exited the tagging band, so that TI1<tagged bolus; and secondly, the delay following the saturation pulse must be sufficiently long for every arterial tagged spin to arrive at the tissue voxel of the image plane so that TI-TI1>transit delay. A comparable technique employed to enhance the perfusion measurement's accuracy is referred to as QUIPSS II with thin slice TI1 periodic saturation (Q2TIPS)[108]. The method substitutes the single saturation pulse with a thin slice period train of saturation pulses at the tagging slab's distal end.

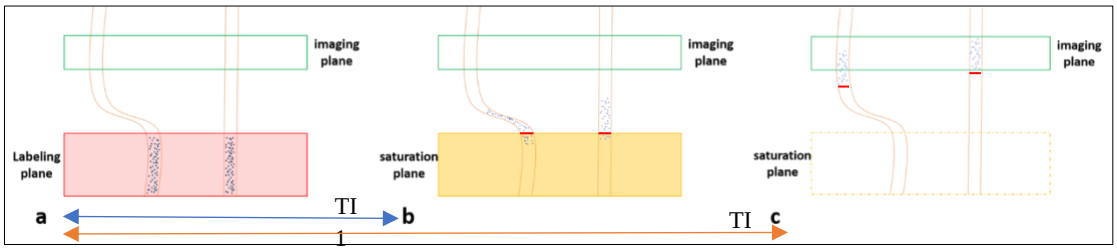


Figure 2.10 Illustration of QUIPSS II technique and its conditions. This technique generates a defined tagged bolus for the PASL not in space but in time and hence provide a quantitative flow image. Here, both conditions of the QUIPSS II technique are satisfied, the first condition is (from a to b) TI1<tagged bolus, and the second condition is that TI-TI1 is greater than the transit delay.

Immediately following a perfect inversion pulse, the difference in arterial blood magnetisation is 2M0a. However, this is not the case in real-practice settings because different ASL tagging approaches contain a various degree of successfully achieved inversion (labelling), which in turn impact the labelling efficiency. Therefore, the labelling efficiency (α)[109,110] is defined as the ratio between the ASL difference image that with imperfect inversion and the perfect 2M0a. Every labelling approach offers a contrasting ratio for labelling efficiency, and this can be taken into consideration when calculating the CBF, as shown in the forthcoming section.

D. CBF quantification models

The ASL images acquired as raw data (source images) need to be processed to quantify the CBF. Quantitative CBF map is estimated from the source images using a mathematical model. This model is taking into consideration both relaxation and kinetics of the labelled protons and provides a reliable quantitative perfusion measurement.

i. T1 model

Detre and Williams[89] modified the Bloch equation for the longitudinal magnetisation by adding in the flow effect employing Fick's principle, which stipulates that variation between the delivery of the substance present in blood and its clearance is proportional to the local flow:

$$\frac{dMt(t)}{dt} = \frac{Mt0 - Mt(t)}{Tt1} + f\,M\mathrm{a(t)} - f\,Mv(t) \qquad \text{Equation 2.17}$$

Where $Mt0$ is the tissue's longitudinal magnetisation at equilibrium, while the time-dependent longitudinal magnetisations for the tissue, the venous- and arterial- blood, are $Mt(t)$, $Mv(t)$ and $M\mathrm{a(t)}$, respectively. The terms $f\,M\mathrm{a(t)}$ and $f\,Mv(t)$ represent the magnetisation bolus transition which enters and leaves the tissues capillary bed through the arteries and veins, respectively. This is referred to as the T1-quantification model[89].

This model involves a number of assumptions that the magnetisation bolus will not be affected by the longitudinal relaxation time of the blood as the water instantaneously exchange between the tissue and blood and hence the transit time is negligible $M\mathrm{a(t)}$, therefore, will be unchanged. Also, the brain tissue can be thought of as a well-mixed compartment and the labelled bolus completely extracted thus:

$$M\mathrm{a(t)} = Mv(t) = Mt(t)/\lambda \qquad \text{Equation 2.18}$$

λ is the distribution volume between the blood and the tissue and illustrates the tracer distribution during equilibrium between the tissue and the blood.

Moreover, the longitudinal magnetisation of the tagged blood flow experiences a reduction with apparent longitudinal relaxation time $T1app$ depending on the blood flow and the T1 relaxation time of the tissue (Tt1); hence the $T1app$ magnitude is determined by:

$$\frac{1}{T1app} = \frac{1}{Tt1} + \frac{f}{\lambda}$$ Equation 2.19

As the labelling continuously performed the magnetisation evolves into a steady-state; thus, the perfusion can be expressed by:

$$f = \frac{\lambda}{T1app}\left(\frac{Mcontrol - Mlabel}{2\,Mcontrol}\right)$$ Equation 2.20

ii. General kinetic model

As previously mentioned, the difference ASL image express ~1% signal variation. It is vital that this signal variation reflects the quantitative perfusion rather than parasitic signals due to confounding factors. There are numbers of confounding factors that could be a primary source of systematic errors in the ASL experiment and hence need to be considered for quantitative reflection of perfusion. These confounding factors include the transit time, exchange the labelled water between capillaries and tissue, and its effect on the T1 relaxation time and the incomplete extraction of labelled water from the capillaries. The T1 quantification model is based on a restrictive assumption that the delivered labelled blood water immediately mixes and exchanges with the tissue's water. This model is not considering the systematic errors that may occur in ASL measurements. On the other hand, the general kinetic model[111] reproduces the T1 model while relaxes the assumption to account for the confounding factors into account. It describes the changes in the signal of ASL over time using three-time functions with assumptions, to allow quantification of the systematic error magnitude.

The difference signal ($\Delta M(t)$) can be considered as the tracer concentration (labelled blood water or magnetisation) that is transported by arterial flow to the voxel at a time (t). The general kinetic model in eq.2.4 developed to represents the ASL agent concentration precisely. In ASL, the tissue concentration at a time point (t) is a result of the delivered magnetisation by the arterial flow and its removal by the venous blood flow and the longitudinal relaxation. Thus, in the general kinetic model, besides the typical time functions ($C_A(t)$ and R(t)), there is an additional time function called

magnetization relaxation function m(t), which demonstrates the carried fraction of the original longitudinal magnetisation by the labelled bolus that remains at the time (t).

In case of an ideal ASL experiment, where the blood water is fully inverted in the labelled image and fully relaxed in the control image, the arterial input function ($C_A(t)$) equals to one. Figure 2.11 illustrates the assumptions of the $C_A(t)$, it will be zero prior to the initial transit time (Δt) as there is no tagged blood , greater than zero following Δt and as long as the continuous arrival for the bolus duration length ($\Delta t + \boldsymbol{\tau}$), and return to zero as the bolus terminates. The delivered magnetisation $C_A(t)$ is reduced from one because its longitudinal magnetisation is partially relaxed (blood T1a rate) during the transit time . The following equation highlights the delivery function:

$$C(t) = e^{-t/T1b} \quad \text{PASL}$$

$$C(t) = e^{-\Delta t/T1b} \quad \text{CASL}$$

Equation 2.21

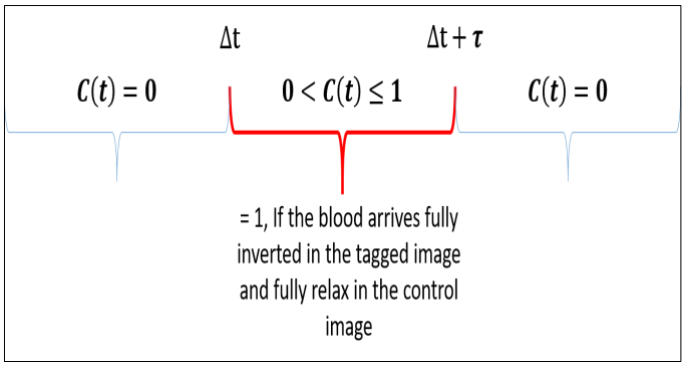


Figure 2.11 An illustration of C(t), the delivery parameter which does not equal zero at Δt <t< $\Delta t + \boldsymbol{\tau}$.

Assuming the labelled water clearance follows the rule of a “single well-mixed compartment”, the R(t) is described by an exponential decay. Single well-mixed compartment kinetics means that the exchange between the different sub-compartments found in these tissues is fast with constant concentration through the tissue as function of time . Assuming a constant concentration ratio between the venous and the tissue zones, the residue function can be described as:

$$R(t) = e^{-ft/\lambda}$$

Equation 2.22

Once the tagged bolus is extracted from the vascular space to the tissue space, the magnetisation is further reduced by the T1 relaxation time of the tissue (Tt1):

$$m(t) = e^{-t/Tt1} \qquad \text{Equation 2.23}$$

Consequently, the adapted general kinetic model for ASL according to the aforementioned assumptions (eq.2.21 to 2.23) is:

$$C(t) = \begin{cases} 0 & 0 < t < \Delta t \\ e^{-t/T1b} \text{ PASL} & \Delta t < t < \Delta t + \boldsymbol{\tau} \\ e^{-\Delta t/T1b} \text{ CASL} & \\ 0 & \Delta t + \boldsymbol{\tau} < t \end{cases} \qquad \text{Equation 2.24}$$

$$R(t) = e^{-ft/\lambda}$$

$$m(t) = e^{-t/T1}$$

Here, Δt is the initial time of arrival of the labelled blood, $\boldsymbol{\tau}$ is the bolus width or the period of the arterial bolus, and $\Delta t + \boldsymbol{\tau}$ indicates the ongoing delivery of the tagged blood to the imaging slice with a length equal to the bolus width.

According to the general kinetic model (eq.2.4) for any agent that is not metabolized, the tissue concentration curve $Ct(t)$ is represented by the convolution (*) of the arterial input function with the local impulse function. Since the ASL difference image (ΔM) represents the amount of the tracer concentration delivered by the arterial blood flow at a time (t) in an ideal ASL experiment, the arterial input function is $2M0a\,.f.dt.$; hence the arterial input function is:

$$\Delta M = 2\text{M0a}\, f\, C(t) \qquad \text{Equation 2.25}$$

The impulse response function at a time (t) is defined as the delivered magnetisation by the local blood flow f adding the fraction of the remaining magnetisation from the previous delivery, which is the product of the local blood flow, the residue function and the magnetisation function [f.R(t).m(t)].

Therefore, the general kinetic model for quantification of the cerebral perfusion based on the labelled magnetisation as freely diffusible is represented by:

$$\Delta M(t) = 2M0a\, f\, C(t) * [r(t)m(t)] \qquad \text{Equation 2.26}$$

Here the convolution is presented by * with the brackets referring to integral of the time dimension (the arterial bolus duration).

Applying the assumptions from eq.2.24 to eq.2.26 gives the following solution for PASL:

$$\Delta M(t) = \begin{cases} 0 & 0 < t < \Delta t \\ \dfrac{-2\, M0a\, .\alpha\, .f}{\delta R}\, e^{-R1a\, t}\left(1 - e^{-\delta R(t-\Delta t)}\right) & \Delta t < t < \Delta t + \boldsymbol{\tau} \\ \dfrac{-2\, M0a\, .\alpha\, .f}{\delta R}\, e^{-R1a.\Delta t}\left(1 - e^{\delta R(t-\Delta t)}\right) & \Delta t + \boldsymbol{\tau} < t \\ \quad .\, e^{-R1app\,(t-(\Delta t+\boldsymbol{\tau}))} & \end{cases} \qquad \text{Equation 2.27}$$

The inversion efficiency is represented by α to account for the fact that the blood water may be not fully inverted in the labelled image or not fully relaxed in the control image. The inversion efficiency is defined as the ratio between the ASL difference image that with imperfect inversion and the perfect 2M0a; R1α is the rate of blood relaxation and $R1app$ is the apparent tissue relaxation rate, where $\delta R = R1a - R1app$. Figure 2.13 illustrates how each section of the general kinetic model's solution occurs through a specific time period.

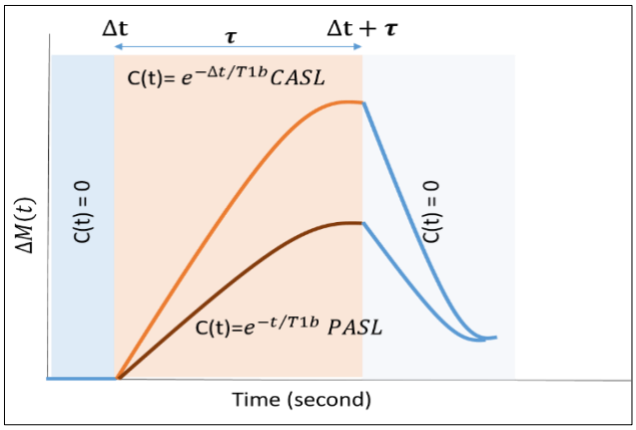


Figure 2.12 This is an illustration of the general kinetic model curve which shows the variation between CASL and PASL regarding the delivery bolus C(t). Here CASL is represented by the higher curve and PASL by the lower. The various area of the general kinetic model, as time function, are also represented by the blue, red and light blue boxes

Various models can be used to obtain approximations of quantitative perfusion. For the sake of simplicity, the kinetic model presented earlier has been recommended in the white paper[100] (it is a summary for the recommended ASL implementation with the consensus of the International Society for Magnetic Resonance in Medicine-(ISMRM) and the European ASL in Dementia consortium) with particular assumptions that measure the CBF from the obtained ASL at single delay time. First, as long as the post labelling delay time is longer than the arterial arrival time (AAT), the entire labelled bolus is delivered to the target tissue. Second, because the half-life of the magnetisation bolus is relatively short beside it diffuse into the tissue space that with a large water component, this bolus is kept within the tissue with no or negligible outflow. Third, the labelled magnetisation decay can be described by the T1a due to the small variation between the blood- and tissue- T1 relaxation, 1.6s and 1.3s at 3T[112,113], respectively. Accordingly, the measurement of CBF per voxel can be done performed for PCASL as:

$$\mathrm{CBF} = \frac{6000.\lambda.(\mathrm{SI}\ control - \mathrm{SI}\ label).e^{\frac{\mathrm{PLD}}{\mathrm{T1a}}}}{2\ .\alpha\ .\mathrm{T1}a\ .\mathrm{SI}PD\ .(1-\mathrm{e}^{\frac{-\tau}{\mathrm{T1}a}})}\ [ml/100\ g/min] \qquad \text{Equation 2.28}$$

And for QUIPSS II PASL as:

$$\mathrm{CBF} = \frac{6000.\lambda.(\mathrm{SI}\ control - \mathrm{SI}\ label).e^{\frac{\mathrm{TI}}{\mathrm{T1a}}}}{2\ .\alpha\ .\mathrm{TI1}\ .\mathrm{SI}PD}\ [ml/100\ g/min] \qquad \text{Equation 2.29}$$

This equation contains constants values, quantified values per subject and fixed time periods values for every acquisition. Constants values are the following: the factor of 6000 used to transform CBF units from ml/g/s to ml/100g/min; λ the coefficient for water blood/water partition being defined as 0.9 ml/g; α the labelling efficiency that is reliant of the labelling strategy being for PASL α = 0.98 and for PCASL α = 0.85[90]; $T1_a$ the blood relaxation time, which is equal to 1650ms at 3 Tesla[102,104]. The measured values are the signal intensity of the label- and the control-image, SI_{label} and $SI_{control}$, respectively; the obtained proton density image without labelling pulse to scale the ASL difference image and enable the measurement of absolute CBF SI_{PD} . Regarding the fixed duration values : $\boldsymbol{\tau}$ is the duration of the tagging, for the PCASL recommended to be 1800ms, while for the PASL is recommended to be 800ms, which define using QUIPSS II or Q2TIPS as aforementioned; Symbolise the delay time for PCASL as PLD and for the PASL as TI.

d. Improving the SNR

Although a proportional relationship exists between the SNR and image quality, a critical problem associated with ASL images is the inadequate SNR. It is typically the case in MRI that one of two approaches are used to increase the SNR either by strengthening the received signals, or by reducing the level of noise. These issues methodologies can be considered in ASL by opting for taken into account the SNR impact factors, which include the hardware employed, the acquisition parameters, and the use of special techniques.

i. Hardware considerations

Given the strong magnetic field used in MRI, signals increase in a quadratic way while noise increases in a linear way[114]. Therefore, the acquisition of ASL images using a strong magnetic field will elevate the SNR due to the high magnetisation and the prolongation of the T1 relaxation time of blood[115]. Nevertheless, a range of limitations arises due to the high magnetic field. The heterogeneity of the magnetic field increases in line with its strength and elevated field strength coexists with B1 increase and the consequent rise in RF power disposition. Although these limitations must be considered, the use of a strong magnetic field is valuable in addressing the weaknesses associated with low field strength[116]. Additionally, coil type influences the

strength and quality of the signals that are received. Multi-channel receiver coils are frequently recommended[100] because they facilitate the reception of data from a range of channels, thereby increasing the SNR in ASL. Furthermore, they enable the acceleration of the image acquisition using parallel imaging.

ii. Acquisition parameters and special techniques

Increasing the voxel size leads to a higher SNR as larger voxels contain a greater number of protons . Nevertheless, large voxel size corresponds to lower spatial resolution, which can blur image details and lead to partial volume effects. In ASL images, these effects lead to perfusion cross-contamination, which may lead to underestimates or overestimates of the grey matter (GM) and white matter (WM) perfusion, respectively[117]. This is due to the fact that WM has lower perfusion (the literature indicates that WM-CBF is 10-33% of GM-CBF[118]) and longer transit time compared to GM. For that reason, WM generates lower perfusion signals when compared to GM. As a compromise the voxel size is advised to be 3-4mm in-plane and 4-8mm through-plane[100].

Acquisition of repeated and subsequently averaged control and labelled images, thereby elevating the averaging number, can also strengthen ASL signals. However, this prolongs the scanning time and gives rise to motion artefacts. These artefacts can be reduced by lowering the scanning time by using parallel acceleration via under-sampling the K-space. Still, this approach can lower the SNR, and so a moderate acceleration of 2-3 is suggested[100]. Furthermore, the induced motion artefacts can be mitigated by employing background suppression during image acquisition and by acquiring the repeated labelled and control images in an interleaved way. This pattern lowers the misregistration artefact associated with image subtraction. Motion artefact can be further reduced during pre-processing via co-registration.

As noted above, it is possible to improve ASL signals by using low spatial resolution (large voxel) and increasing the averaging. Nevertheless, when these SNR optimisation techniques are applied, signal fluctuation is generated from physiological noise and motion artefacts. Static tissues dominate the source of the noise and artefact, meaning that the SNR can be increased by reducing the static tissue. Noteworthily, this can be achieved by employing inversion pulses following the

labelling pulse, and this process is known as background suppression (BS)[119,120]. It is critical to recognise that, following the labelling pulse, the inverted spins return to equilibrium as well as the difference value (control – label) decays towards zero . Therefore, the application of BS pulses occurs in a global way and does not influence the difference value. The BS efficiency is typically 95%, the implication of which is that each inversion pulse will lead to the distortion of the ASL signals by around 5%[120]. Furthermore, the use of BS in multi-slice imaging that involves multi-excitation pulses per repetition time (TR), will have varying effectiveness across the slices. Contrastingly, the efficiency of the technique is highest in imaging methods that rely on a single excitation per repetition time (e.g. single-slice 2D, single-shot 3D, or segmented 3D). Regarding segmented-3D imaging modalities, these play a key role in improving the SNR while eliminating ghosting artefacts (i.e., motion-related artefacts that arise from inter-segment phase inconsistencies).

The volume of labelled blood water that reaches the imaging plane is the variable that has the largest effect on ASL image quality. In the context of CASL/PCASL, the volume is calculated based on the labelling duration, which itself is limited by the T1 relaxation of blood (T1a). For much longer labelling duration than T1a, the benefit of generating newly labelled blood is counteracted by signal loss due to relaxation of the previously labelled spins. The corresponding loss of SNR and long TR necessitate the use of a greater number of image averages to achieve the same SNR; as a result, the experiment is not practical. To strike a compromise between maximising SNR, reduced power deposition, and preserving the SNR for long ATT, whilst balancing signal loss due to T1a decay, the recommendation is to use a labelling duration of 1,800ms at 3T[100].

On the other hand, the tagged bolus volume of the PASL is based on the labelling slab size. Therefore, the slab width should be as large as possible; however, this is limited by several factors. First, an increase in slab thickness entails an increase in the width of the transition zone; hence during the long transitions, the tagged spins will experience signal loss. Second, the RF transmit coil's size restrict the thickness of the labelling slab. Third, the B1 from the RF transmit coil falls off far from the iso-centre, meaning that the blood water at the coil edges will only undergo partial inversion (thus influencing the degree to which labelling is efficient). These partially inverted spins will

require a longer time to washout from the labelling slab. As a result of this extended washout, a long TR is needed prior to the next tagging pulse implementation, which lowers time efficiency. A reasonable compromise, as identified in the literature, is to use a slab thickness of 15-20cm[100].

e. *Further parameters beyond the CBF*

Acquisition of ASL images at only one point in time offers absolute quantification but the approach is sensitive to transit time. As aforementioned, it is possible to use a post labelling delay in CASL/PCASL to lower this sensitivity, while in PASL, QUIPSS II or Q2TIPS can be employed. Other perfusion parameters can be extracted from the ASL images. By obtaining ASL at several post labelling delay time points, information related to flow and transit time can be collected and quantified beside the CBF, specifically by fitting the obtained images at several time points to the general kinetic model[121–125].

Utilisation of ASL model-free permits arterial blood volume (aBV) measurements[126], namely by acquiring the ASL image at several time points in the presence and absence of a crusher gradient. With this approach, it is possible to estimate the local arterial input function (AIFs) by subtracting the obtained perfusion images in the presence and absence of the crusher gradient. Resultantly, deconvolution of the tissue perfusion signals curve with the AIFs curve can yield the residue function (i.e., the amplitude that reflects the CBF), and an estimate of aBV can be produced by integrating the AIFs.

Chapter 3

(Part of the content of this chapter has been published in Oncotarget in 2019[45])

3 The value of Arterial Spin Labelling in Adults Glioma Grading: Systematic Review and Meta-analysis 95

Abstract 95

3.1 Introduction 96

3.2 Material and methods 96

3.2.1 Literature search and selection 96

3.2.2 Statistical Methods 100

3.3 Results 101

3.3.1 Eligible ASL studies 101

3.3.2 QUADAS-2 assessment 101

3.3.3 Differentiation between HGGs and LGGs 102

3.3.4 Differentiation of grade II from grade III gliomas 103

3.3.5 Differentiation of grade II from grade IV gliomas 103

3.3.6 Differentiation of grade III from grade IV gliomas 103

3.3.7 HROC curve analysis of gliomas grading 106

3.3.8 HROC curve analysis for differentiation between HGGs and LGGs 108

3.4 Discussion 111

3.5 Conclusion 116

3 The value of Arterial Spin Labelling in Adults Glioma Grading: Systematic Review and Meta-analysis

Abstract

This chapter sought to assess the diagnostic efficacy of ASL when employed in grading of gliomas in adults. Following a systematic review of two databases, MEDLINE and EMBASE, the articles extracted were assessed against the inclusion criteria resulting in 18 articles being selected. QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2) was implemented to examine each study's quality. From these studies, the extracted qualitative values were used to conduct a meta-analysis employing a random-effect model together with forest plot modelling, and joint sensitivity and specificity. Curve analysis using the hierarchical summary receiver operating characteristic (HROC) was also performed. Differential grading of high-grade gliomas (HGGs) from low-grade gliomas (LGGs) is possible using the absolute tumour blood flow (TBF). Absolute TBF can also distinguish between grade II and IV tumours but not between grades II and III or between grades III and IV. Contrastingly, distinguishing HGGs from LGGs and between the various glioma grades is possible using relative TBF (rTBF). Optimum outcomes for glioma grading were achieved using maximum rTBF (rTBFmax). The analysis of specificity and sensitivity also mirrored these findings; with rTBFmax scoring most for the efficacy of differentiation between glioma grades. Differentiating between LGGs and HGGs (-1.46 (-2.00, -0.91), p-value<0.001) and between grade II and III tumours (-1.39 (-1.89, -0.89), p-value<0.001), yielded on average the same approximate rTBF effect size. Nevertheless, between grades III and IV tumours, the effect size was reduced (-1.05 (-1.82, -0.27)), p<0.05). These findings were in keeping with those of specificity and sensitivity analysis. Accordingly, the findings of this meta-analysis indicate ASL to be beneficial for differential grading of gliomas and in particular when employing rTBFmax.

3.1 Introduction

Glioma grading as stipulated by the WHO impacts both patient prognosis and treatment choices. MRI has an important role to play in all tumour management stages including diagnosis, treatment and follow-up. T1-weighted postcontrast (T1-c) imaging enables highlighting areas of disruption in the blood-brain barrier (BBB)[32,33], which are linked to higher-grade gliomas (HGGs) as specified by the WHO. Nevertheless, enhanced T1-c can be misinforming as some low-grade gliomas (LGGs) exhibit contrast enhancement, while lack of enhancement can also be noticed in some high-grade gliomas (HGGs)[34]. ASL compared to the other contrast-enhanced perfusion MRI techniques uses tagged blood water as a diffusible tracer, hence allows the estimation of absolute cerebral blood flow (CBF). In addition, ASL removes the need to employ an exogenous contrast agent; consequently, it can be repeated (e.g. for therapy monitoring. Finally, ASL insensitivity to alterations in vessel permeability can indicate the vascular density of tumour via the estimated perfusion[39].

Absolute tumour blood flow (TBF) and normalised or relative TBF (rTBF) are the hallmark output values of quantitative ASL. Studies have demonstrated the benefits of these values in differentiating between LGGs and HGGs[43,127,128] and in differential grading of gliomas[129]. Contrastingly, various studies have also indicated opposing findings[84,130,131]. The aim of this chapter is, therefore, to systematically review and perform a meta-analysis of all these inconsistencies through the reported findings and to present evidence in support of ASL's diagnostic precision in preoperative differential grading of gliomas.

3.2 Material and methods

3.2.1 Literature search and selection

PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidelines[132] for systematic reviews were applied in this chapter to answer the research question 'What is the diagnostic value of arterial spin labelling (ASL) in the differentiation of glioma grades in adult patients?'. The PICO (Population/Intervention/Comparator/Outcomes) categories were used to establish the keywords employed in the search. These were linked using the Boolean operators 'AND' between PICO fields and 'OR' within PICO fields. Each keyword was classed by

a corresponding concept: The first concept – P for Population: glioma OR neuroglia OR oma; the second concept – I for Intervention: arterial spin OR artery spin; the fourth concept O for Outcome: diagnosis OR grading OR differentiate. To make the search more comprehensive, the third concept (C for Comparator) was excluded, and this led to the joining of the first two steps and the last step as [(glioma OR neuroglia OR oma) AND (arterial spin OR artery spin) AND (diagnosis OR grading OR differentiate)]. The systematic review took place in June of 2018, and this involved mining the EMBASE database and the OVID MEDLINE ® database. A total of 111 articles from EMBASE and 48 from MEDLINE were found that agreed with the keywords used, equating to 159 articles. Elimination of duplicates left 122 articles, from these only the reports that were using ASL grading of untreated gliomas in adults were included. PRISMA was employed for indicating appropriate articles, as illustrated in Figure 3.1. 104 articles that did not relate to the research question were excluded, leaving 18 articles deemed eligible and matched the determined inclusion criteria. Examples for the reports that did not adhere to the inclusion criteria included those using ASL in animal models[133], employing ASL to examine treatment efficacy[134,135] or implemented in a group of paediatric patients[136].

Out of the 18 studies, only seven reported the TBF values as mean and standard deviation as well the cut-off and its corresponding diagnostic specificity and sensitivity rates. Of the other articles, eight stated the mean and standard deviation and three reported other values (cut-off values, specificity and sensitivity rates). The articles reported TBF in various terms and the findings of the studies were retitled and entered into the analysis in the following manner:

Absolute or relative to normal-appearing white matter was indicated as TBFmean/rTBFmean when the region of interest (ROI) segmented was based on the conventional images, and the whole tumour volume was included.

Maximum TBFmax/rTBFmax terms were used to indicate the estimated measurements of the highest signals from tumour ROI on the perfusion map. Certain studies detailed both the maximum and the mean TBF/rTBF values.

The power of the evidence from this meta-analysis may be restricted by the extensive variability in the included articles, including the diversity in the studied samples,

employed labelling approach and the undertaken post-processing. Accordingly, two researchers acting autonomously employed the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2) tool[137] to examine the bias risk associated with the selected articles and the extent of their applicability.

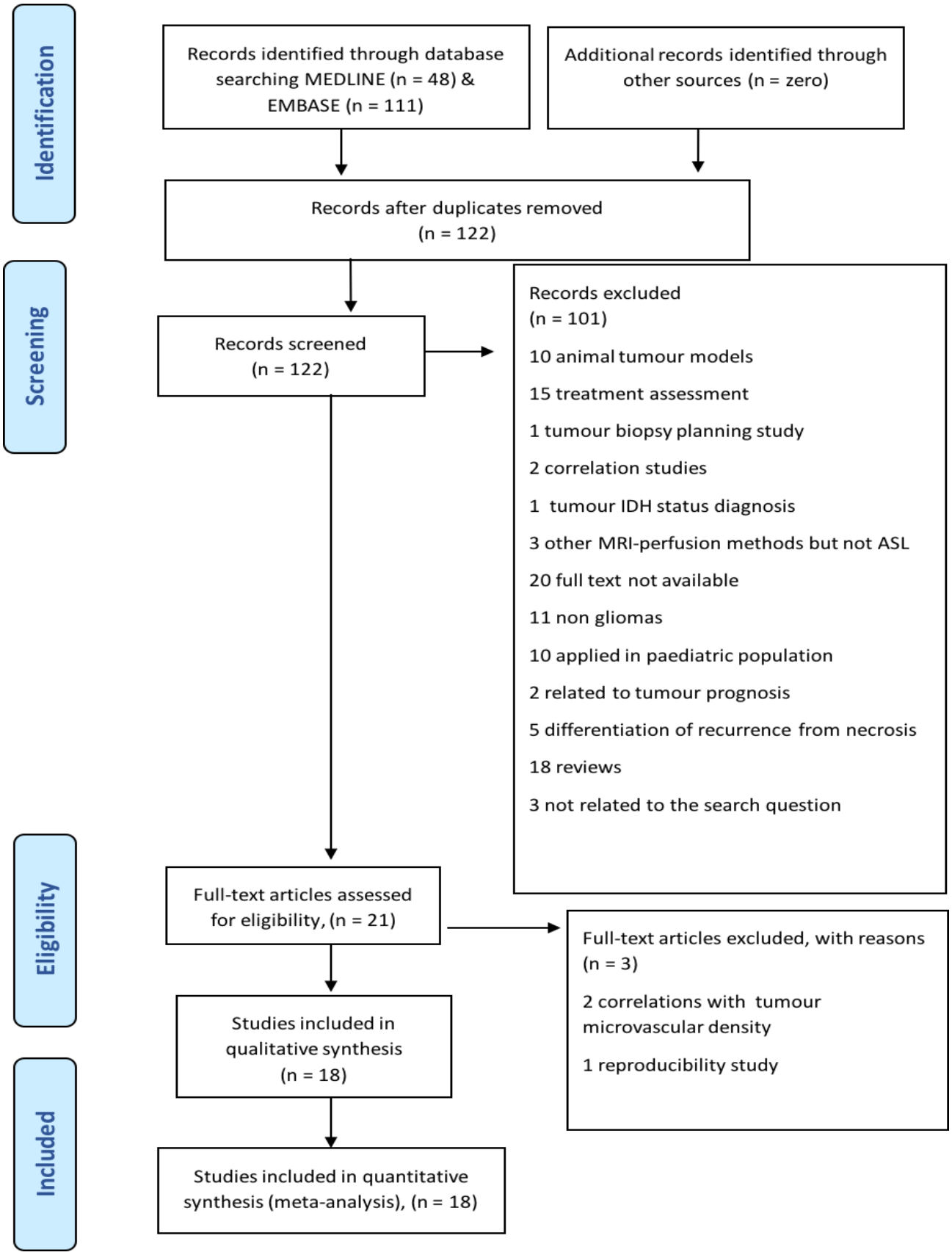


Figure 3.1 Illustration PRISMA flow chart for the selection process among the identified research studies.

3.2.2 Statistical Methods

The gliomas grades categories' comparisons considered were LGGs against HGGs, grade II against grade III, grade III against grade IV, and grade II against grade IV. All extracted outcomes were on a continuous scale. From every study, the sample size, mean and standard deviation (SD) were obtained. Some articles only reported the data range; for this one-fourth of the full range was assumed to be the SD. From the included articles, those where no variability measure , such as the interquartile range, the range or SD were indicated followed exclusion from this analysis (e.g. article 14, although it supplied the ROC analysis). Some articles reported the same index from both the full sample and the sub-group (e.g. study 8, Table 2A in the supplementary data). In addition, other articles reported the finding from the whole sample employing either different ASL methods (e.g. article 16, Table 2A in the supplementary data) or different image analysis techniques (e.g. article 18, Table 1A in the supplementary data). In such instances, the different data sets from each article were assumed as individual study data to prevent the study weighting being doubled.

To establish whether the obtained values among the studies varied significantly, the Chi-square test for heterogeneity was employed. Heterogeneity was also quantified using the I^2 statistic value. The Chi-square test[138] is calculated as a weighted sum of the squared difference between the effect of each individual study and the pooled effect among the studies. It assesses whether the observed variation in the meta-analysis is due to chance and determines if this is statistically significant. I^2 [138,139] describes the percentage of variability in the estimated effects of studies in the meta-analysis that are attributed to heterogeneity rather than chance. It has been developed to quantify the inconsistency between studies outcomes (heterogeneity). I^2 calculated as:

$$I^2 = 100\% \cdot \frac{Q - df}{Q} \qquad \text{Equation 3.1}$$

Where Q is the chi-squared statistic and df is its number of degrees of freedom. If the I2 value is above 50%, this is considered as evidence for substantial heterogeneity.

The pooling of the findings from the various articles was then performed. The standardised mean difference (SMD) rather than the raw mean difference was calculated because the obtained values among studies were on different scales. All

meta-analyses were performed using a random-effects model which did not consider the heterogeneity between the various studies. Funnel plots were used to assess bias in publication, and these assessed relationships between the SMD or the estimated effect size and its standard error (degree of uncertainty).

The diagnostic performance of the obtained ASL's values was assessed using receiver operating curve analysis methods (ROC), such as sensitivity and specificity in distinguishing between LGGs and HGGs . The analysis strategy employed was like that outlined by Li et al.[140,141], which models both specificity and sensitivity jointly, due to the inverse association between these two outcomes. The approach taken here was to fit a two-level mixed logistic regression model, with independent binomial distributions for the true positives and true negatives conditional on the sensitivity and specificity in each study, and a bivariate normal model for the logit transforms of sensitivity and specificity between studies. Such a strategy generates pooled approximations of specificity and sensitivity together with associated confidence intervals (CI) for each. Moreover, a hierarchical summary receiver operating characteristic (HROC) curve was produced. This analysis was conducted by employing the "*metandi"* command in STATA 15 (StataCorp LLC, College Station, TX, USA).

3.3 Results

3.3.1 Eligible ASL studies

The implemented ASL acquisition method was used to group the various articles, including continuous -ASL (CASL), pseudo-continuous-ASL (PCASL) and pulsed-ASL (PASL). These are detailed in Appendix A (Tables 1A-3A). Tables 1A-3A also illustrate the gliomas histological grades and types, the ASL parameters, the reported ASL estimates, and the corresponding p-value of the differential test between LGGs and HGGs.

3.3.2 QUADAS-2 assessment

After the QUADAS-2 investigation, five articles had been found to have a lower level of bias risk and concerns regarding applicability. The output from the QUADAS-2 investigation is illustrated in Figure 3.2. The bias risk graph expresses four domains with minimal risks being associated with both the reference standard domain and the flow and timing domain. Bias could be introduced in most studies through the index

test domain as the reviewers of the study either were not blinded to the standard reference (high) or the authors did not mention this piece of information (unclear). The fourth domain was the patient selection domain, where about a quarter of the studies that were chosen for the meta-analysis did not employ either a random or consecutive means of sampling (high) and approximately a quarter did not detail their participant sampling method (unclear). On the other hand, the applicability graph has three domains. There was a low risk regarding the index test domain and the reference standards domain since all the examined articles employed ASL as the index test and histopathological assessments as a reference standard. Conversely, the patient selection domain had scored high by approximately 10% due to one article accounting for residual gliomas in the analytical procedure and another one including a wide age range from paediatric patients to adults.

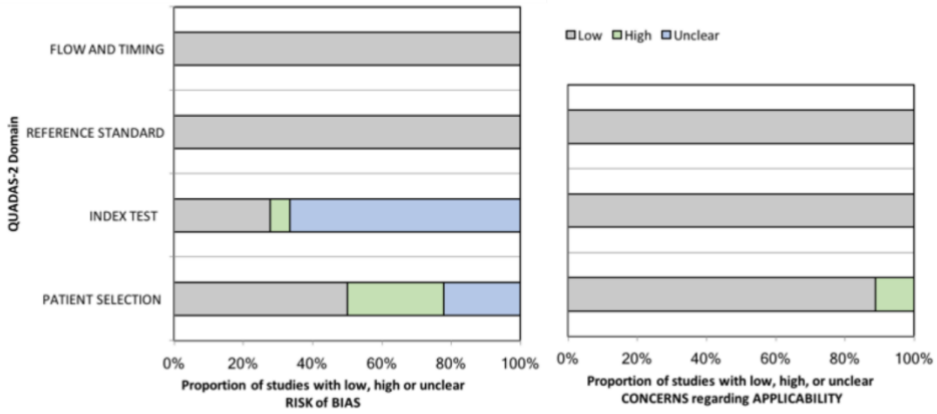


Figure 3.2 Findings of the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) among the investigated studies.

3.3.3 Differentiation between HGGs and LGGs

The selected articles demonstrated high levels of heterogeneity except for TBFmean (Table 3.1). This was noted as they scored a high I^2 value and the Chi-square test was statistically significant. The obtained ASL measurements were significantly lower for the LGGs compared to the HGGs, where the greatest effect size was noted for rTBF. Here the SMD between LGG and HGGs for rTBF was 1.5 SD. Moreover, the extensive heterogeneity in the study was confirmed by the findings of the rTBF funnel plots with values being outside the confidence limits at both extremes. Nevertheless, publication

bias was not overt as the SMD standard error was not linked to the effect size (Egger test, P-value = 0.17, 0.72, 0.24 for rTBF, rTBFmean and rTBFmax, respectively). Rather, the TBF funnel plots indicated an absence of publication bias as most points fell within the limits of confidence (Egger test, P-value = 0.05, 0.43, 0.10 for TBF, TBFmean and TBFmax, respectively). The findings are shown in Appendix A (Figure 1A-6A).

3.3.4 Differentiation of grade II from grade III gliomas

Table 3.1 illustrates a summary of the analysis results. The rTBF heterogeneity between the studies was low, and the rTBF values significantly lower for grade II compared to grade III gliomas. The mean rTBF value was 1.4 SDs lower in grade II than in grade III. Regarding absolute TBF, only two articles reported this and the heterogeneity level between them was high with I^2 at 66%. The absolute TBF noted to have a trend of lower mean value in grade II gliomas by 0.9 SDs compared to grade III. As all points were within the limits of confidence in the funnel plot, asymmetry in the findings was absent indicating no publication bias (Egger test, P-value = 0.46 for rTBF). The graphical illustrations of these findings can be found in Appendix A (Figures 7A and 8A).

3.3.5 Differentiation of grade II from grade IV gliomas

Table 3.1 summarises the findings of the ASL-derived parameters for grade II vs. grade IV gliomas. These findings demonstrated extensive heterogeneity. The ASL biomarkers revealed significantly lower estimates in grade II tumours. The rTBF showed the highest effect size, which was 2 SDs lower in grade II than in grade IV. The TBF effect size was 1.4 SDs. The generated funnel plot for the rTBF was symmetrical, which suggests absence of publication bias (Egger test, P-value = 0.25 for rTBF). The findings are graphically shown in Figures 9A and 10A (Appendix A).

3.3.6 Differentiation of grade III from grade IV gliomas

Table 3.1 illustrates the values of the obtained ASL biomarkers in grade III and IV tumours. A significantly high level of heterogeneity across the 6 studies, which provided rTBF, is indicated by I^2 equal to 69%. The pooled result of the rTBF values was significantly lower in grade III compared to grade IV. The SMD between these two grades was 1SDs, this is slightly less than the mean difference between grade II and

grade III. The pooled results of TBF demonstrated little heterogeneity between the investigated studies and a trend of lower TBF for grade III gliomas (p=0.08). Additionally, the funnel plot for the TBF was asymmetrical with the articles with smaller sample sizes indicating greater SMD values due to the greater standard errors. The reverse was true with studies with large sample sizes having smaller effects as their SEs were low (Egger test, P-value =0.04 for TBF). This, however, was not the case for the rTBF (Egger test, P-value =0.19 for rTBF). Some suggestion of publication bias was noted for the two ASL values as graphically shown in Appendix A (Figures 11A and 12A).

Table 3.1 Comparison of the differences in ASL-related biomarkers between HGGs and LGGs and between different glioma grades.

biomarkers	Number of studies	Total sample size		Heterogeneity		Effect size		Egger test to evaluate publication bias
		LGG	HGG	P-value	I^2	SMD (95% CI) (*)	P-value	P-value
rTBF	15	237	323	<0.001	86%	-1.46 (-2.00, -0.91)	<0.001	0.17
rTBF mean	9	142	192	<0.001	86%	-1.53 (-2.26, -0.79)	<0.001	0.72
rTBF max	6	95	131	<0.001	87%	-1.36 (-2.23, -0.49)	0.002	0.24
TBF	11	155	219	0.002	64%	-0.82 (-1.20, -0.45)	<0.001	0.05
TBF mean	4	51	70	0.50	0%	-0.61 (-0.99, -0.23)	0.002	0.43
TBF max	7	104	149	<0.001	76%	-0.96 (-1.53, -0.39)	0.001	0.10
		II	III					
rTBF	4	62	48	0.26	25%	-1.39 (-1.89, -0.89)	<0.001	0.46
TBF	2	43	21	0.09	66%	-0.90 (-1.85, 0.04)	0.06	(-)
		II	IV					
rTBF	4	62	61	<0.001	87%	-2.07 (-3.38, -0.76)	0.002	0.25
TBF	2	43	32	0.01	84%	-1.44 (-2.76, -0.12)	0.03	(-)
		III	IV					
rTBF	6	54	69	0.006	69%	-1.05 (-1.82, -0.27)	0.008	0.19
TBF	4	27	40	0.64	0%	-0.45 (-0.95, 0.05)	0.08	0.04

(*) SMD calculated as the difference between LGGs ASL-parameters and HGGs ASL parameters (usually higher than the LGG counterparts)

(-) No pooled results due to the low number of included studies

3.3.7 HROC curve analysis of gliomas grading

Table 3.2 illustrates the diagnostic performance of ASL in diagnosing the various grades of glioma. Sensitivity higher than 90% was noted in distinguishing between grade II from grade III gliomas. The sensitivity decreased for distinguishing between grades III and IV – though it was still indicated as high. These findings support the previous evidence that the effect size between grade III and IV gliomas was smaller than that between grade II and III gliomas. The specificity of ASL for each glioma grading task, however, was less than 70%. True specificity and sensitivity were indicated to be highly variable on analysis of the 95% confidence interval and prediction regions. Figure 3.3 below illustrates the HROC plots. The Appendix A (Table 4A) details the values for the sensitivity, specificity, positive and negative predictive values (PPV and NPV, respectively) for different cut-off employed for grading using the ASL parameters from the included studies.

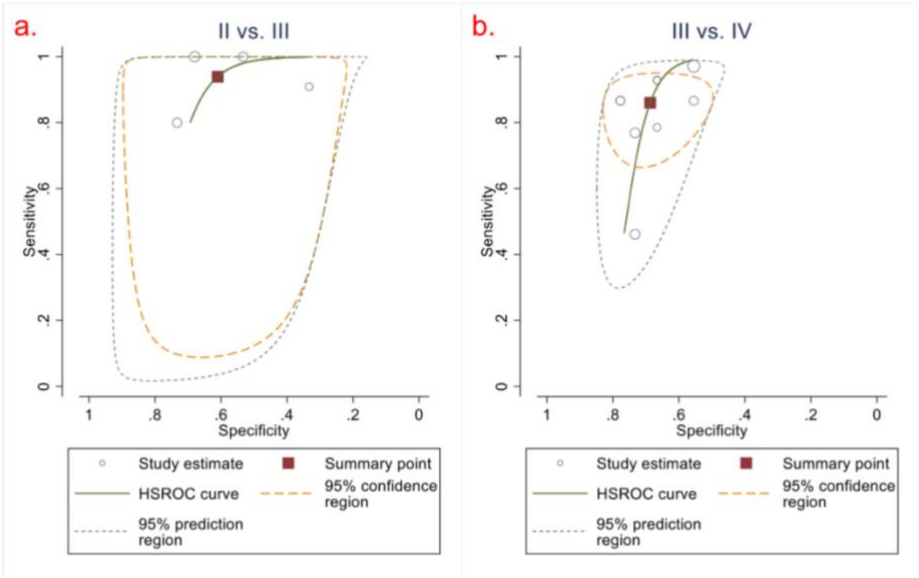


Figure 3.3 The HROC plot expressing the sensitivity and specificity summary point (square) and its the 95% CI (green curve) for the rTBF from ASL to allow differential grading of gliomas: (a) grade II and grade III (94%, CI (75%, 99%)) and (61%, CI (48%, 73%)), respectively; (b) grade III and grade IV (86%, CI (75%, 93%)) and (69%, CI (57%, 79%)), respectively.

Table 3.2 Diagnostic performance of the ASL in discrimination between glioma grades.

Glioma grading	Number of studies included	Total sample size		Sensitivity (95% CI)	Specificity (95% CI)	AUC (95% CI)
II vs. III	4	52	64	94% (75%, 99%)	61% (48%, 73%)	0.76, (0.72, 0.79)
II vs. IV	3	(*)	(*)	(*)	(*)	(*)
III vs IV	9	148	75	86% (75%, 93%)	69% (57%, 79%)	0.75, (0.71, 0.79)

(*) No pooled results due to the low number of included studies

3.3.8 HROC curve analysis for differentiation between HGGs and LGGs

The overall findings indicated relatively high levels of specificity and sensitivity (~85%) by employing ASL biomarkers to differentiate between LGGs and HGGs. There was a similar degree of sensitivity for the ASL-derived measures being investigated. The specificity showed more variability, ranging from 79% for rTBFmean to 92% for rTBFmax. Wide variability was also indicated for the specificity and sensitivity values by the prediction region and the 95% confidence interval. Table 3.3 shows these findings and Figure 3.4 illustrates the HROC plots. Table 5A in Appendix A details the values for the sensitivity, specificity, PPV and NPV for different cut-off values employed in the differentiation of gliomas grades using the ASL biomarkers from the analysed studies.

Table 3.3 Diagnostic performance of the ASL imaging biomarkers in stratifying the tumours between HGGs and LGGs.

Biomarker	Number of studies	Total sample size		Sensitivity (95% CI)	Specificity (95% CI)	AUC (95% CI)
		LGGs	HGGs			
All	17	206	397	86% (78%, 91%)	84% (76%, 90%)	0.91, (0.89, 0.93)
TBF	1	(*)	(*)	(*)	(*)	(*)
rTBF	16	181	370	86% (77%, 91%)	84% (76%, 90%)	0.91, (0.89, 0.94)
rTBF max	5	76	122	85% (69%, 94%)	92% (80%, 97%)	0.95, (0.93, 0.97)
rTBF mean	8	80	188	84% (71%, 92%)	79% (66%, 88%)	0.87, (0.84, 0.90)

(*) No pooled results due to the low number of included studies

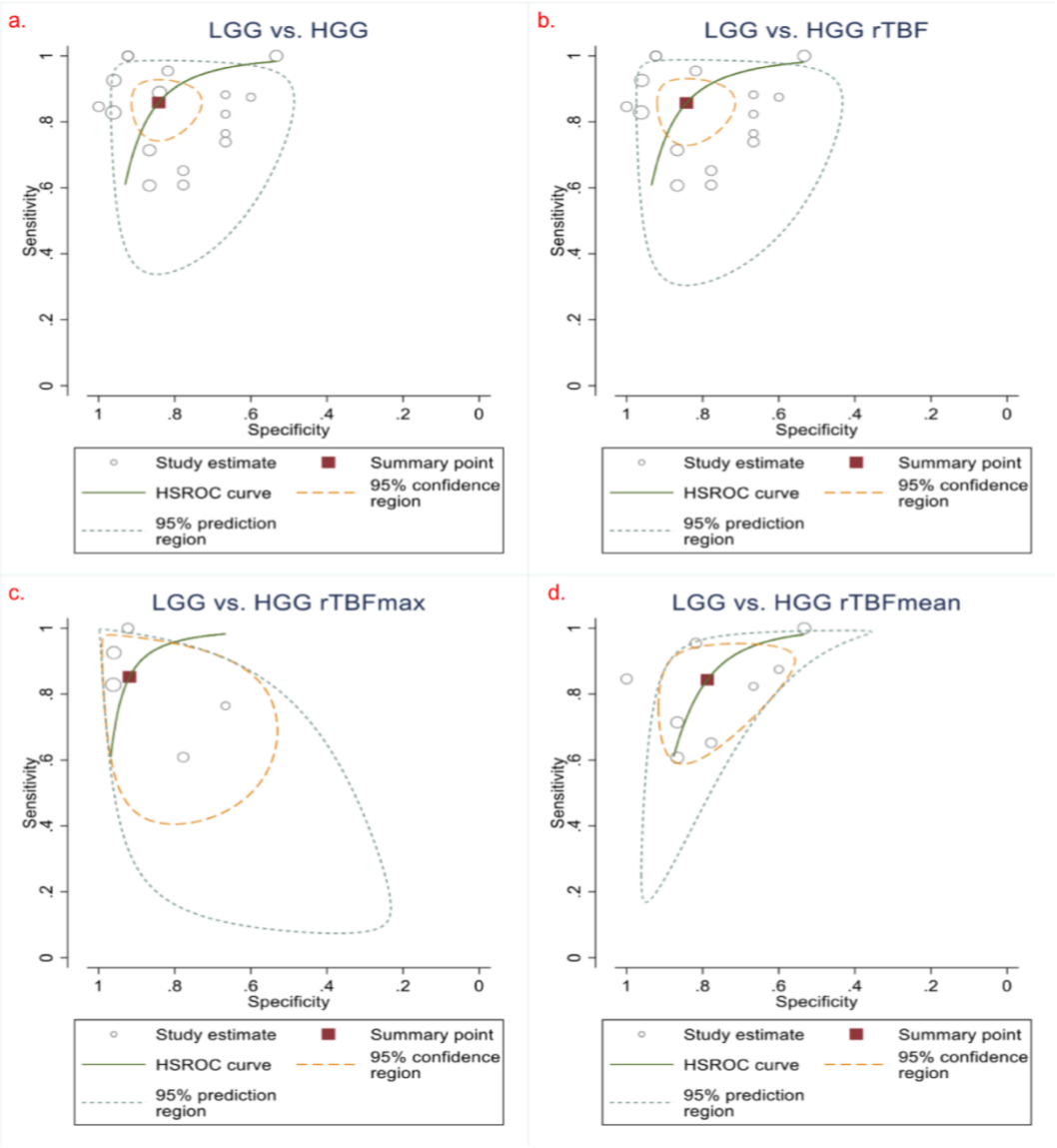


Figure 3.4 The HROC plot shows the sensitivity and specificity summary points (square) together with the 95% CI (green curve): (a) Derived from all the ASL obtained measurements to discriminate HGGs from LGGs (86%, CI (78%, 91%)) and (84%, CI (76%, 90%)) respectively: (b) rTBF for differential between LGGs and HGGS (86%, CI (77%, 91%)) and (84%, CI (76%, 90%)) respectively; (c) rTBF mean for differential between LGGs and HGGS(84%, CI (71%, 92%)) and (79%, CI (66%, 88%) respectively; (d) rTBF max for differential between LGGs and HGGS (85%, CI (69%, 94%)) and (92%, CI (80%, 97%)), respectively.

3.4 Discussion

The meta-analysis findings reported here demonstrate that absolute TBF can be employed for differentiation between LGGs and HGGs and between grade II and IV gliomas. Conversely, distinguishing between grade II and III, and between grade III and IV gliomas was not possible using absolute TBF. On the other hand, rTBF was found to be more efficient than absolute TBF in differentiation between LGGs and HGGs with an acceptable degree of diagnostic performance. Overall, for gliomas grading, the rTBFmax index expressed the best diagnostic performance. Likewise, the findings for the sensitivity and specificity analysis revealed the same results with the rTBFmax providing the highest sensitivity and specificity. The effectiveness of rTBFmax for gliomas grading was also reported by Fudaba et al., where rTBFmax provided higher sensitivity and specificity than rTBFmean[142]. The calculated rTBF effect size was relatively the same between LGGs and HGGs and between grades II and III gliomas (-1.46 (-2.00, -0.91), p-value<0.001; -1.39 (-1.89, -0.89), p-value<0.001, respectively). However, smaller effect size was noted for distinguishing between grade III and grade IV gliomas (-1.05 (-1.82, -0.27)), p<0.05) in keeping with the findings of the specificity and sensitivity analysis.

A previous meta-analysis[143], published in 2017, are in line with our own results. That meta-analysis assessed the ability of ASL perfusion values from nine studies to distinguish HGGs from LGGs. Unlike our meta-analysis, the authors included studies from wide age groups (adult and pediatric). They concluded that, although aTBF demonstrated a significant increase in HGGs than in LGGs, rTBF provided a larger effect size, and it can thus be considered as a better index to differentiate HGGs from LGGs.

The variation across the included studies was clearly observed after the systematic review of the literature and by the heterogeneity analysis. This may be attributed to examining various histological types of glioma, variations in the ASL employed labelling approach and the implemented sequence parameters (e.g. post-labelling-delays PLDs), as well as the used method in processing and analysing the images. Interestingly, all articles that indicated ASL as incapable of differentiating between LGGs and HGGS were articles that examined various types of gliomas[84,130,131]. This sounds rational as low-grade oligodendrogliomas are linked to higher than grade II

gliomas perfusion levels, which would result in diagnostic bias[43,144]. One study only examined oligodendrogliomas, noting that it was difficult to distinguish between the different WHO grades when using ASL[145]. Contrastingly, a handful of studies looking at astrocytomas demonstrated ASL's ability to differentiate between LGGs and HGGs[146–148] as well predict the astrocytoma grades with a high level of specificity and sensitivity[129,142]. Nevertheless, further studies indicated ASL's ability to efficiently distinguish between LGGs and HGGs when examining a mixture of these glioma types[43,128] being also able to grade them[127]. The recent WHO classification guidelines for gliomas categorise gliomas using the genetic/molecular grouping such as the isocitrate-dehydrogenase-(IDH) mutation status and the 1p/19q mutation co-deletion[2]. The investigated articles in this meta-analysis did not consider genetic or molecular information. This would contribute to the demonstrated variation in the findings as perfusion values for IDH-mutation gliomas are reported to be lower than that for IDH-wild gliomas[84,149].

The varying acquisition sequence implementation and labelling strategies employed in the analysed studies introduced variation in the ASL measurements. Three ASL labelling techniques were regularly used in these studies; namely continuous-ASL (CASL), pseudo-continuous-ASL (PCASL) and pulsed-ASL (PASL). The latter is the most widespread method employed as it is easily accessible, has a low specific absorption rate (SAR) and powerful efficiency of labelling[150] for a wide array of blood velocities. However, the PASL technique has a lower signal-to-noise ratio (SNR) in comparison to other ASL labelling techniques. On the other hand, CASL has a lower labelling efficiency and a greater SAR compared to PASL. Finally, PCASL carries the benefits of both PASL (low SAR) and CASL (high SNR), although, its efficiency of labelling is less than PASL. The PCASL technique was reported to have the best reproducibility when investigated in healthy individuals[151] as compared to the other ASL labelling methods.

Among the included studies, all studies that utilised the PASL approach, except two studies with negative findings[84,152], indicated it as capable of differentiating between LGGs and HGGs[43,142,146–148,153,154] and between gliomas grades[129]. Studies that employed CASL were merely promising in distinguishing LGGs from HGGs[144,145,152]. The current literature presents conflicting findings regarding the efficacy of PCASL

acquisition; it has been reported as efficient in distinguishing between LGGs and HGGs[128] and in grading[127]; however, some studies did not support these results[84,130,131]. Such discrepancy can be attributed to PCASL's low efficiency of labelling, which can be different across patients and scanners. Another significant acquisition parameter that could contribute to the variation of the measurements is the post labelling delay (PLD), also called inversion time (TI), since most of the involved studies obtained ASL at a single PLD/TI. Acquisition of ASL at a single PLD/TI minimises the sensitivity to the blood transit time though it does not eliminate it. This is because PLD/TI time selection needs to be a compromise between the various AATs across brain regions and magnetisation loss due to T1-decay, which are both on the order of seconds. Thus, while the PLD/TI time is required to be sufficiently long to allow the labelled blood bolus enough time to move through from the labelling plane to the target tissue of the imaging plane, it must also be sufficiently short to preserve the magnetisation label from T1-decay. Due to neo-angiogenesis in tumours that impacts the AAT, choosing an appropriate PLD/TI when examining tumours is a challenging task. Furtner et al. noted that the most efficacious TI for establishing differences between low- and high-grade astrocytomas was at 370ms[146]. Similar or lower sensitivity and specificity were reported from other studies obtained ASL images at single delay time (range from 1200-1900ms)[128,129,142,147,152,153]. Unlike using a single PLD/TI time, obtaining the ASL measures at multiple PLD/TIs [106] reduces the sensitivity to the bolus arrival time, hence obtaining more reliable CBF estimates. Cebeci et al. used 8 TIs to discriminate HGGs from LGGs[43] and Yang et al. in astrocytomas obtained 16 TIs, the latter not only differentiated HGGs from LGGs but also enabled glioma grading, as the precision of the estimated CBF is expected to increase with multiple PLD/TI points delay[129].

The quantified CBF estimates among the studies were heavily skewed due to the non-standardised analysis and selection of ROI. Extensive use of normalised TBF values (rTBF) was noted as these reduce the degree of scattering within a group, which arises from haemodynamic and age-related differences[148]. Moreover, the elevation of intracranial pressure usually noted in gliomas patients tends to decrease the global CBF, and accordingly, the precision of the estimated local TBF. Several studies, however, argued that the noise from the used internal reference tissue for normalising

the TBF introduces variation across observers[42,155]. This could explain the wide range of the rTBF values in HGGs. Nevertheless, the rTBF was considered more reliable than the absolute TBF for differentiating between LGGs and HGGs[43,148], and this has also been confirmed by this meta-analysis.

Previous ASL studies in gliomas have used different brain tissues as an internal reference for normalisation, including the grey matter (GM)[156], the white matter (WM)[153], the mean of both[154], or the contralateral normal tissue mirrored to tumour[148]. Interestingly, the mirror ROI was more efficacious than either WM or GM for normalisation[157], which is perhaps due to the approximately similar distance from the labelling plane. Nevertheless, most studies employed the contralateral normal-appearing white matter as an internal reference[127,129,142,145,147,153,158]. It is necessary to highlight that both the reliability and validity of the perfusion measurements may be impacted by the image used for tumour region-of-interest (ROI) delineation. Various studies select the tumour ROI following a visual examination of the maximal signal intensity on either the M_0 image or the ASL-subtracted image, while others used the conventional MRI image. Regarding the generation of the tumour mask based on the contrast-enhanced T1 weighted images, it should be noted that the enhanced tissue represents the disruption of the blood-brain barrier (BBB) and does not necessarily reflect the increased perfusion[159].

In line with this meta-analysis results, TBFmax has been reported to be more precise than TBFmean for tumour characterisation[160,161]. This is attributed to the fact that HGGs are heterogeneous in nature; hence the TBFmax will be representative of the most anaplastic tumour part. Moreover, the TBFmax as a biomarker is better in comparison to TBFmean, which is influenced by partial volume averaging. Histograms analysis[60,162] that explore the tumour heterogeneity are probably able to differentiate the gliomas grades with high diagnostic accuracy and reproducibility.

Various studies have assessed the usability of ASL and its role in routine brain tumour diagnosis comparing it against well-established MRI methods such as MR spectroscopy (MRS)[142,145,154] and MRI-diffusion techniques[127,128,153] as well as against other MRI perfusion techniques, including dynamic susceptibility contrast-enhanced [43,148,160] and dynamic contrast-enhanced (DCE)[84,130]. Currently, DSC is the most

frequently employed MRI perfusion technique in the clinical practice for brain tumours assessment. Nevertheless, for patients who could not tolerate high-rate contrast injection or have a contraindication to contrast agents, ASL might be beneficial. Not to mention, ASL is a true alternative perfusion technique to avoid permanent gadolinium depositions in case of multiple MRI studies, if the latter is of clinical concern [163]. The studies that investigated both ASL and DSC in glioma grading focused on the use of ASL as an alternative or surrogate of DSC[42,43,148], for example, by assessing the interchangeability or the agreement of the estimated perfusion metrics from each assessment, instead of suggesting which method is superior. Such studies, though, exhibited the non-inferiority of ASL. Warmuth et al. reported a strong positive correlation between the rTBF measurement from ASL and DSC[148]. Another study reported a high correlation between the estimated rTBFmean and rTBFmax from both DSC and ASL[42]. Cebeci et al. found a moderate but significant correlation between the rTBF values from PASL and DSC[43]. These studies provide preliminary evidence that ASL can be used as a non-invasive alternative to DSC addressing the shortcoming of non-diffusible tracer (gadolinium) in DSC that leaks out of disrupted BBB and may lead to wrong estimation of the rCBV measurements[164].

Two studies that were assessed in this meta-analysis investigated the association between DCE and ASL in gliomas[84,130]. These studies used PASL and PCASL and proposed that ASL was not efficacious in the grading of gliomas and indicated poor to moderate correlation between DCE and ASL. This may be due to differences in ROI selection method, the different pathophysiological basis and meaning of the derived DCE and ASL biomarkers, and HGGs sample size in each study.

To be able to use ASL as an alternative to contrast MRI perfusion techniques (DCE and DSC), its inter-observer variability and the reproducibility in brain tumours must be examined. Numerous research studies have indicated good inter-observer variability employing PASL[129,147], PCASL[127,131] and QUASAR (quantitative STAR labelling of arterial regions) [161]. Moreover, Hirai et al. indicated excellent intra- and inter-rater reproducibility of ASL in gliomas using both maximum and mean TBF [161].

With respect to the use of ASL for tumour imaging in the future, PCASL was employed by Yoo et al. in a recent publication to investigate any association between genetic

biomarkers and perfusion values in HGGs[165]. The findings indicated that overexpression of the epidermal growth factor receptor (EGFR) was significantly correlated with absolute TBF and rTBF. Moreover, another study by Yamashita et al. demonstrated that both rTBF and TBF were significantly higher in glioblastoma (GBM) patients with IDH-wildtype status compared to those with IDH-mutation[149]. Besides, Brendle et al. indicated that in contrast to DCE, ASL could be used to stratify astrocytomas based on the IDH-mutation status[84].

This meta-analysis does have its limitations. First, almost all the assessed perfusion indices were heterogeneous among the selected articles. This is expected as MRI acquisitions tend to show variability across centres or even different vendor platforms in a single centre. This degree of heterogeneity was accounted for in the analysis using a random effect model. The second limitation was the small sample size, with 18 included studies. However, this reflects the strict methodological standards employed in line with our objectives.

3.5 Conclusion

This meta-analysis aimed to shed light into the diagnostic performance of ASL in glioma grading and demonstrated the suitability of ASL-derived perfusion metrics in glioma grading. rTBFmax showed the best diagnostic and staging performance. Hence, ASL metrics' capacity as imaging biomarkers can be routinely used for the characterisation and staging of gliomas at baseline, with implications for treatment selection and possibly surveillance imaging. However, further research with larger numbers of patients and well-defined tumour subtypes groups, including their molecular information, is needed to refine TBF-related threshold values that allow higher diagnostic and prognostic accuracy and are essential for the wide dissemination of the technique.

Chapter 4

4 Repeatability of Perfusion Measurements in Gliomas Using Pulsed and Pseudo-continuous Arterial Spin Labelling (ASL) MRI 118

4.1 Introduction 119

4.2 Methods 120

4.2.1 Patients 120

4.2.2 ASL acquisitions 120

4.2.3 CBF maps calculation 122

4.2.4 Image analysis 125

4.2.5 Statistical analysis 126

4.3 Results 128

4.4 Discussion 141

4.5 Conclusion 143

4 Repeatability of Perfusion Measurements in Gliomas Using Pulsed and Pseudo-continuous Arterial Spin Labelling (ASL) MRI

Abstract

Investigate the repeatability of perfusion measures in gliomas using pulsed- and pseudo-continuous- arterial spin labelling (PASL, PCASL) techniques, and evaluate different regions-of-interest (ROIs) for relative tumour blood flow (rTBF) normalisation. Repeatability of cerebral blood flow (CBF) was measured in the Contralateral Normal Appearing Hemisphere (CNAH) and in the brain tumours (rTBF and absolute-TBF (aTBF)). rTBF was normalised using large/small ROIs from the CNAH. Repeatability were evaluated using intra-class-correlation-coefficient (ICC), Within-Coefficient-of-Variation (WCoV) and Coefficient-of-Repeatability (CR). PASL and PCASL demonstrated high reliability (ICC > 0.9) for CNAH-CBF, aTBF and rTBF. PCASL demonstrated more stable signal-to-noise ratio (SNR) with a lower WCoV of the SNR than that in PASL. PASL and PCASL showed higher WCoV in the tumour (aTBF, rTBF) than in CNAH CBF, and higher WCoV for rTBF than for aTBF when normalised using a small ROI. The lowest CR was observed for rTBF normalised with a large ROI. PASL and PCASL showed similar repeatability for the assessment of perfusion parameters in patients with primary brain tumours as previous studies based on volunteers. Both methods displayed reasonable WCoV in the tumour area and CNAH. PCASL's more stable SNR in small areas (caudate) is likely to be due to the longer post-labelling delays.

4.1 Introduction

Cerebral blood flow (CBF) quantification using arterial spin labelling (ASL) has been suggested to provide an imaging biomarker that can be used in both glioma grading [45,143] and for evaluation of tumour progressions following radiotherapy [166–168]. However, fundamental differences between ASL implementations (e.g. different acquisition sequences and spin labelling approaches) are known to affect the sensitivity of CBF estimation due to variation in measurement precision. Other potential sources of errors can also include scanner instability and short-term physiological fluctuations in tumour perfusion. Repeatability and reproducibility analysis are useful tools to validate the precision of the quantitative measurements in clinical trials and enable exploitation of their benefits and integration into clinical practice [169].

A previous study investigating the reproducibility of quantitative STAR labelling of arterial region (QUASAR) [170] reported within-subject standard deviations (WS-SD) due to scanner instability of 3.1 ml/100g/min, increasing to 4.3 ml/100g/min after repositioning, and to 5.3 ml/100g/min including long term physiological variations over separate measurements. The scanner instability was thus considered as the main cause of quantification error, since each further increase in WS-SD is smaller than the initial value. This result was confirmed by another study that compared the test-retest reproducibility of pulsed-ASL (PASL), continuous-ASL (CASL) and pseudo-continuous-ASL (pCASL) approaches [151]. However, one of the main issues which limits generalizability of these studies is that they have only reported ASL reproducibility in healthy volunteers [151,170]. A recent study which evaluated the intra-session reliability of PCASL in 6 patients with glioblastoma (GBM), showed high reliability in both the normal-appearing grey matter (NAGM) (ICC>0.90, WCoV 3.40%-7.12%) and tumour (ICC 0.98; WCoV 4.91%) [171].

ASL measurements in brain tumours studies are usually expressed as absolute tumour blood flow (aTBF) or relative (or normalised) tumour blood flow (rTBF). rTBF has been shown to be more reliable than aTBF when distinguishing between high- from low-grade gliomas (HGGs and LGGs, respectively) [43,45,148]. Using a large ROI as the reference region for normalization provides excellent reliability, but can introduce inaccuracies due to the variability of the bolus arrival time (BAT) throughout the ROI [172]. Thus, the choice of tissue type to act as an internal reference and the size of the

ROI used for this purpose is likely to add bias and affect the variability and repeatability of rTBF.

The aim of this chapter is to assess the repeatability of both pulsed ASL (PASL) and pseudo-continuous ASL (PCASL) with 3D GRASE and to decide which internal reference, both in terms of the ROI size and region of placement, is better for normalising the rTBF.

4.2 Methods

4.2.1 Patients

Forty adult glioma patients (WHO histological glioma grades II-IV; aged 43±14 years) were recruited in this study. All participants provided signed informed consent, and the study was approved by the local institutional research board and the ethics committee. Sixteen subjects were scanned using PASL (9 HGGs, 7 LGGs); this data was retrospectively re-analysed from a previous study comparing perfusion measurements derived from ASL, dynamic susceptibility contrast (DSC) and dynamic contrast-enhanced (DCE) in clinical neuro-oncology [173]. An additional 24 subjects were prospectively identified and scanned using PCASL (9 HGGs, 15 LGGs). Both groups of patients were scanned on clinical 3T MR scanners (Siemens Healthineers, Erlangen, Germany).

4.2.2 ASL acquisitions

PASL data were acquired with a flow-sensitive alternating inversion recovery (FAIR) ASL labelling approach using a segmented 3D GRAdient- and Spin-Echo (GRASE) readout (8 shots; for other sequence parameters, see Table 4.1). For the calculation of CBF maps in absolute biological units (ml/100 g/min), images were acquired using the same acquisition scheme without background suppression and with three saturation recovery times (1s, 2s, and 4s). M_0 and T_1 were estimated by fitting these saturation recovery images using NiftyFit, an open-source software developed by the Centre for Medical Image Computing at University College London [174].

PCASL data was similarly acquired using a segmented GRASE readout, except with four shots (Table 4.1). For CBF quantification, a proton density (PD) image was

acquired as the first volume of the series, with the same imaging readout but without labelling or background suppression.

Table 4.1 Imaging parameters of the investigated 3D ASL protocols		
	PASL (Siemens product)	PCASL (Siemens WIP)
Scanner	Siemens 3T Skyra	Siemens 3T Prisma
Number of segments	8	4
TR/TE (ms)	3500/12	4600/13.28
Bolus duration (ms)	1000 (estimated)*	1800
Post-labelling delay time (PLD/TI) (ms)	1800	1800
FOV (mm)	240 × 240	220 x 220
Matrix size	64 × 52	64 x 60
Voxel size (mm^3)	3.75 × 4.63 × 5.25	3.44 x 3.67 x 4
Slice thickness (mm)	5.25	4
Slices per slab	20	28
Acquisition Time	3min 52s	5min 15s
Refocusing flip angle (°)	130	130
Turbo / EPI factors	15/13	28/15
Background suppression scheme (BS)	2 Inversion pulses with timings calculated to null static tissue magnetization with T1=700ms and 1400ms 100ms before acquisition	same as for PASL
Number of repeats/averages	4	8
Echo spacing	0.54 ms	0.51 ms

* Note that due to a technical issue, the bolus duration was estimated for the PASL datasets.

4.2.3 CBF map calculation

After applying motion correction to the raw data using MCFLIRT [175] (from FMRIB Software Library (FSL)), ASL difference images (control – label) (ΔASL) were calculated. This resulted in four ΔASL for PASL, and six ΔASL for PCASL, because the first two measurements were removed from all PCASL data sets due to the presence of CSF artefacts. CBF maps are typically calculated by averaging over several repeats to increase SNR. Here, rather than measuring the repeatability between the CBF maps calculated from each measurement (single ΔASL), CBF maps were generated by averaging different combinations of single ΔASL measurements (Figure 4.1), to establish a number of measurements for each patient (known as

moving block bootstrapping [176]). Repeated measurements from the same subject present as a finite time series of correlated data. Typical bootstrapping resamples measurements assuming they are independent and random, however, this assumption is violated in our study. Instead, moving block bootstrapping [176] overcome this problem as it resamples the observations as blocks instead of individual observation. This creates replicated samples and increase the accuracy of sampling distribution estimate without altering the dataset.

Consequentially, each combination (or block) involved half of the acquired ΔASL measurements. The 4 ΔASL PASL images thus produced six combinations and the 6 ΔASL PCASL images produced 20 combinations. Each of these generated combinations was then averaged, and CBF maps were calculated using the simplified general kinetic model proposed in the ISMRM (International Society for Magnetic Resonance in Medicine) Perfusion Study Group consensus paper[100] (eq.4.1 and 4.2):

For FAIR-PASL:

$$\mathrm{CBF} = \frac{6000.\lambda.(\Delta\mathrm{ASL}).e^{\frac{\mathrm{TI}}{\mathrm{T1a}}}}{2.\alpha.\mathrm{TI1}.\mathrm{M0}}\ [ml/100\ g/min] \qquad \text{Equation 4.1}$$

For PCASL:

$$\mathrm{CBF} = \frac{6000.\lambda.(\Delta\mathrm{ASL}).e^{\frac{\mathrm{PLD}}{\mathrm{T1a}}}}{2.\alpha.\mathrm{T1}a.PD.(1-e^{\frac{-\tau}{\mathrm{T1}a}})}\ [ml/100\ g/min] \qquad \text{Equation 4.2}$$

Each equation contained assumed constant values, measured values for each subject, and fixed timing parameters for each acquisition. The constant values were a factor of 6000, used to convert the CBF unit of ml/g/s to ml/100 g/min; λ, the water blood/brain partition coefficient, defined to be 0.9 ml/g; α, the labelling efficiency, dependent on the labelling approach used (in this case, for PCASL = 0.85 and for PASL = 0.98); and T1a, the blood relaxation time, which was approximately equal to 1650ms at 3T [100]. Please note that in this study, the reduction in labelling efficiency due to the use of BS pulses was not considered, as it was kept constant for each method throughout the project and as such, would not affect any calculated coefficient of variation. The measured PD (PCASL) or fitted M_0 (PASL) value was used as a scaling image for the ΔASL image, enabling the absolute quantification of CBF. The fixed duration values

were the labelling bolus duration (LD) (τ/TI1 for PCASL/PASL, respectively), and the post labelling delay (PLD)/inflow time (TI) for PCASL/PASL, respectively.

The PD for PCASL was multiplied by an appropriate scaling factor, as recommended by Alsop et al. [100] (eq.4.3) since it was acquired with TR less than 5s.

$$\mathrm{SI\_}M0 = \mathrm{SI\_}acq \,.\, \frac{1}{(1 - e^{\left(\frac{-TR}{T1,tissue}\right)})} \qquad \text{Equation 4.3}$$

The TR was 4.6s and the T1 of the grey matter (~1.3s) [112], resulting in a multiplying factor of 1.03.

PASL 4 repeated measurements (control-label)

1	2	3	4	CBF1
1	2	3	4	CBF2
1	2	3	4	CBF3
1	2	3	4	CBF4
1	2	3	4	CBF5
1	2	3	4	CBF6

CBFs from possible combination

PCASL 6 repeated measurements (control-label)

1	2	3	4	5	6	CBF1
1	2	3	4	5	6	CBF2
1	2	3	4	5	6	CBF3
1	2	3	4	5	6	CBF4
1	2	3	4	5	6	CBF5
1	2	3	4	5	6	CBF6
1	2	3	4	5	6	CBF7
1	2	3	4	5	6	CBF8
1	2	3	4	5	6	CBF9
1	2	3	4	5	6	CBF10
1	2	3	4	5	6	CBF11
1	2	3	4	5	6	CBF12
1	2	3	4	5	6	CBF13
1	2	3	4	5	6	CBF14
1	2	3	4	5	6	CBF15
1	2	3	4	5	6	CBF16
1	2	3	4	5	6	CBF17
1	2	3	4	5	6	CBF18
1	2	3	4	5	6	CBF19
1	2	3	4	5	6	CBF20

CBFs from possible combination

Figure 4.1 The permutation matrices illustrate all the possible combinations without repetition of k-element subsets (half of the repeated measurements) from a set of an element (repeated measurements, single ΔASL). The PASL has 2-elements subsets of a 4-elements set while the PCASL have 3-elements subsets of a 6-elements set. Thus, PASL dataset produces 6 combinations and the PCASL one 20 combinations of CBF maps.

4.2.4 Image analysis

All MR images were converted to NIfTI format using MRIconvert, (version 2.0.x, https://lcni.uoregon.edu/downloads/mriconvert/mriconvert-and-mcverter).

Conventional MR images (T1-weighted (T1-W), post-contrast T1-weighted (T1-C), T2-weighted (T2-W) and Fluid-attenuated inversion recovery (FLAIR)) were used as anatomical templates. ITK-SNAP[177] (version 3.6.0; www.itksnap.org) was used to segment the whole tumour volume manually on the T2-weighted images for PASL and the FLAIR images for PCASL, according to availability. 3D high resolution anatomical T1-W MPRAGE images were used for automatic segmentation of the grey matter (GM) and white matter (WM) volumes in the contralateral normal-appearing hemisphere (CNAH) with the FAST tool from the FSL library [178]. Subsequently, these CNAH GM and WM volumes were used as "large ROIs" for rTBF normalisation. The caudate nucleus (caud) in the CNAH was segmented automatically from the MPRAGE using the FIRST tool from the FSL library [178], and was utilised as an example of "small ROI" for rTBF normalisation. Care was taken to avoid the tumour if it crossed the midline to the contralateral hemisphere. In short, the overlapping area between the CNAH binary mask and the tumour binary mask was subtracted from the CNAH binary mask, to generate a tumour free binarized CNAH mask. Rigid followed by affine registration in NiftyReg [179] was used to resample all segmented ROIs to the M_0/PD images for PASL/PCASL, respectively. This is because M_0/PD images have the same resolution as the CBF maps and possess similar anatomical features to structural images, thus facilitating co-registration (see Figure 4.2).

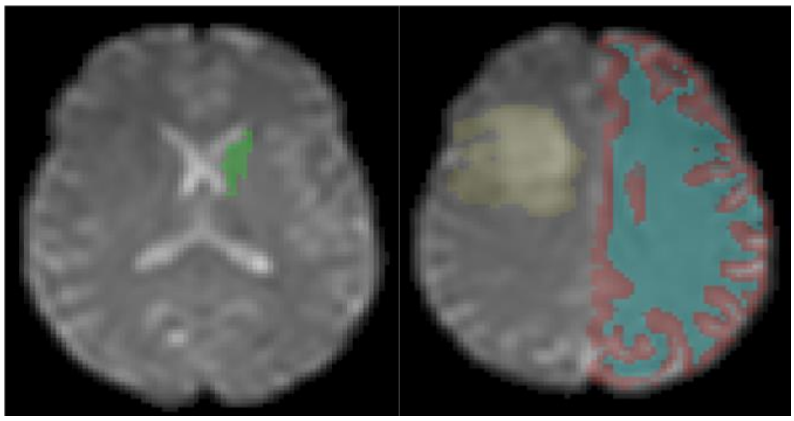


Figure 4.2 Automatically segmented ROIs in the contralateral healthy hemisphere in a patient with right frontal glioma, which was manually segmented. Green, caudate; red, GM; blue, WM; yellow, tumour.

Values were extracted from the maps as follows: i) tumour: mean aTBF and 95th-percentile (aTBF-95per); ii) CNAH: mean CBF in WM, GM, and caudate (CBF-WM, CBF-GM, and CBF-caud, respectively). The normalised tumour values relative to WM, GM, and caud (rTBF-WM, rTBF-GM and rTBF-caud, respectively) and their 95th-percentiles (rTBF-WM-95-per, rTBF-GM-95-per and rTBF-caud-95-per, respectively) were calculated. Note that the 95 percentile value of perfusion within the tumour is an important biomarker that has been shown to be relevant for diagnosis purposes in primary gliomas [45]. SNR was calculated for the ΔASL images and the calculated CBF maps in the CNAH ROIs according to the NEMA methodology [180] (eq.4.4):

$$SNR = \sqrt{2}\frac{\text{mean (ROI from (image1))}}{SD\ \text{(ROI from (image1} - \text{image2))}} \qquad \text{Equation 4.4}$$

With image1 and image2 representing either two consecutive ΔASL images or calculated CBF maps, respectively. The SNR was only calculated in the CNAH ROIs in order to avoid any variation which might result from abnormal tumour angiogenesis. Thus, the SNR from the ΔASL in WM, GM and caud were CBF_WM_SNR, CBF_GM_SNR and CBF_caud_SNR while from CBF maps were ΔASL_WM_SNR ΔASL_GM_SNR, and ΔASL_caud_SNR, respectively.

4.2.5 Statistical analysis

As this study sought to determine the extent to which the PASL and PCASL can be considered repeatable in terms of measuring aTBF/rTBF in tumour and CBF in the CNAH, repeatability was assessed separately for each ASL method. Test-retest repeatability was assessed based on the repeatability indices according to QIBA (Quantitative Imaging Biomarkers Alliance[1]), namely the intra-class correlation coefficient (ICC), the within-coefficient-of-variation (WCoV in %) and the Coefficient of Repeatability (CR).

The ICC (two-way mixed effect, absolute agreement) was calculated for each of the extracted values from the CBF maps, grouped within each patient. This ICC model

[1] http://qibawiki.rsna.org/images/8/8c/FMRITechnicalPerformanceIndices041613.pdf, accessed on June 2018.

was selected as repeated measurements cannot be assumed to be a random sample[181], and repeated measurements are meaningless without the presence of agreement. The ICC helps to determine the ratio of the true variance to the observed variance. ICC values should be interpreted as follows: <0.5 shows "poor reliability, 0.5 to 0.75 shows moderate reliability, 0.75 to 0.9 demonstrates good reliability, and >0.9 offers excellent reliability.

The within-subject coefficient of variation (WCoV in %) is another repeatability index that represents the within-patient variation and is given in (eq.4.5):

$$\%WCOV = \left(\frac{SDwithin\ subject\ measurements}{mean\ of\ the\ measurements}\right).100 \qquad \text{Equation 4.5}$$

WCoV reflects the variability within the data due to random error. Less variability implies higher reliability. The WCoV was calculated among the repeated measurements for each patient, as the within-subject variation is not expected to be uniform among tumour patients. Applying the root mean square approach [182], the WCoVs were squared for each patient to find the overall mean and 95% CI, after which the square roots of the mean and the 95% CI were calculated.

A further repeatability index is the Coefficient of Repeatability (CR), also known as the smallest real difference (SRD) [183,184]; it allows quantification of the variation (error) in the same unit as the measured value (ml/100 g/min in this case) with 95% probability. Lower CR corresponds to a more sensitive/responsive measurement. Changes less than the CR are not detectable and deemed to be due to measurement variation, meaning that the method is insensitive to changes smaller than the CR.

To estimate the CR, first, the within-subject standard deviation (WS-SD) was calculated among the repeated measurements for each patient, as it is not expected to be uniform among tumour patients. The same procedure was used as for the WCoV and the CR was calculated according to (eq.4.6):

$$CR = 1.96\sqrt{2(WS\text{-}SD)^2} \qquad \text{Equation 4.6}$$

As different sets of patients underwent PASL and PCASL acquisitions, the WCoV was used for each extracted value from the calculated maps to compare the variations between PASL and PCASL (using Mann-Whitney U tests); the variations between HGGs and LGGs (using unpaired t-tests); the variations between aTBF and rTBF, between normal CBF and aTBF, and between normal CBF and rTBF, both using paired t-tests. The IBM SPSS version 24 (*SPSS Inc, Chicago, IL*) statistical package was used for the statistical analysis.

4.3 Results

4.3.1 Intraclass correlation coefficient (ICC)

Generally, all perfusion metrics obtained from both PASL and PCASL demonstrated high reliability (ICC > 0.9) as illustrated in Table 4.2. Nevertheless, the between-patient SD of the extracted values was lower for PCASL than for PASL (Table 4.2). The tumour perfusion values using PASL demonstrated higher between-patients SD.

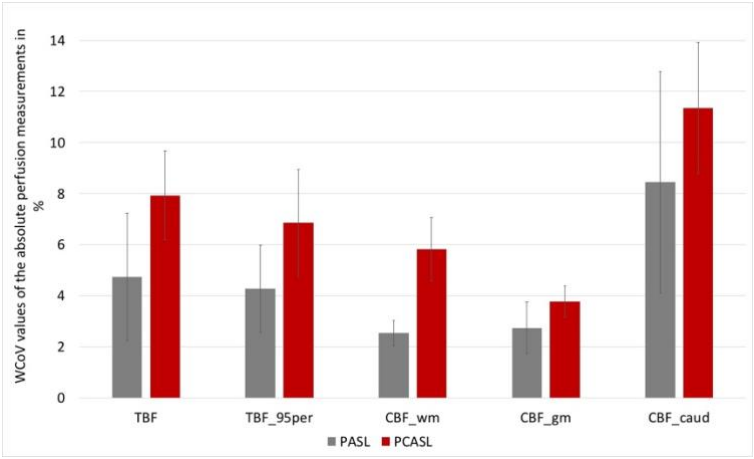
4.3.2 Within-subject coefficient of variation (WoCV) and Coefficient of Repeatability (CR)

There were no significant differences (p-value>0.05) between the HGGs and LGGs regarding the estimated WCoV of tumour perfusion values (aTBF and rTBF). WCoV of rTBF and rTBF_95per that normalised to a "small ROI" was significantly higher than WCoV of aTBF (Tables 4.2 and Figure 4.3, and Table 4.3) whereas no significant WCoV increase of aTBF compared to rTBF and rTBF_95per were observed when the latter was normalised to "large ROI". Normalisation of the tumour perfusion metrics to the whole WM or GM of the CNAH showed its ability to filter out some of the signal variations, while using the "small ROI" raised the signal variation level. Moreover, the estimated WCoV of the aTBF_95per, rTBF and rTBF-95per (normalised to caudate) were significantly higher than CBF WCoV in WM and GM. The estimated WCoV of the aTBF, rTBF and rTBF-95per (normalised to WM or GM) were not significantly different from CBF WCoV in GM and WM, whereas WCoV of the aTBF and rTBF (normalised to WM or GM) was significantly lower than the estimated WCoV of CBF in the caudate (Table 4.3). All estimated CR of the measured values are illustrated in Table 4.2 and Figure 4.4, where the lowest CR was for the rTBF normalised to the whole CNAH GM. In the PASL CBF maps, macrovascular artefacts were observed,

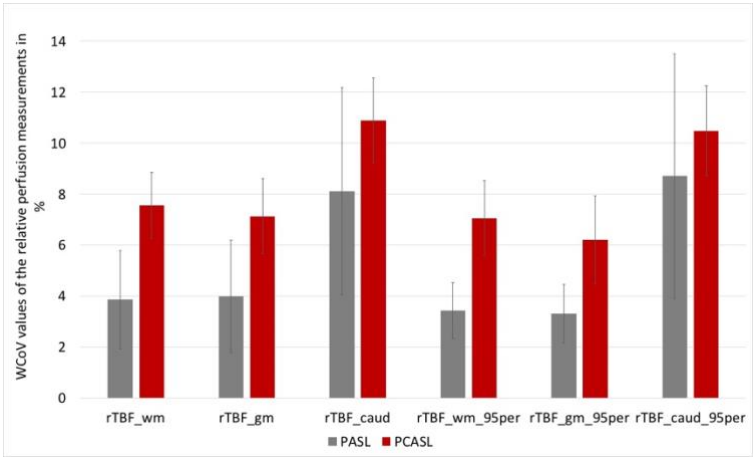
coming from accumulation of labelled spins in large vessels; Figure 4.5 shows examples of CBF maps for both PASL and PCASL.

With the imaging parameters used within this study, the mean CBF and SNR values from PCASL were lower than those from PASL (Table 4.2, Figure 4.6). Furthermore, as expected, the SNR showed significant negative correlation with the WCoV (WM, r= -0.632 p-value < 0.001; GM, r= -0.602 p-value < 0.001; caud, r= -0.463 p-value=0.003). The WCoV of the perfusion metrics was significantly higher ($p < 0.05$) for PCASL than PASL (Table 4.4). In contrast, the WCoV of the SNR was significantly lower for PCASL than PASL in the caudate, and showed the same trend, though non-significant in the WM and the GM.

a.



b.



c.

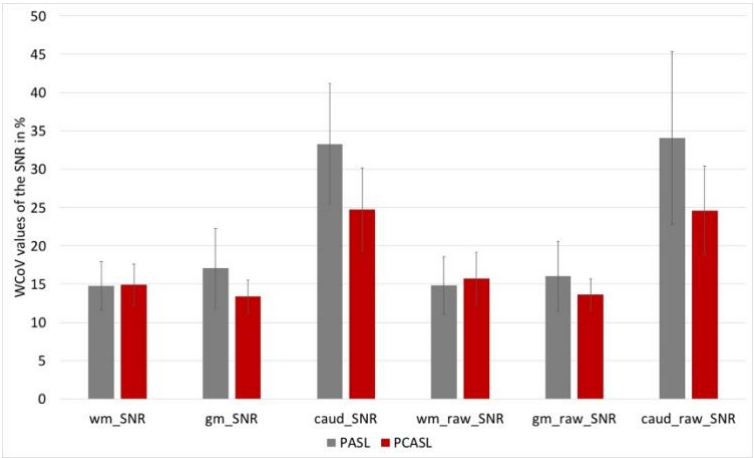
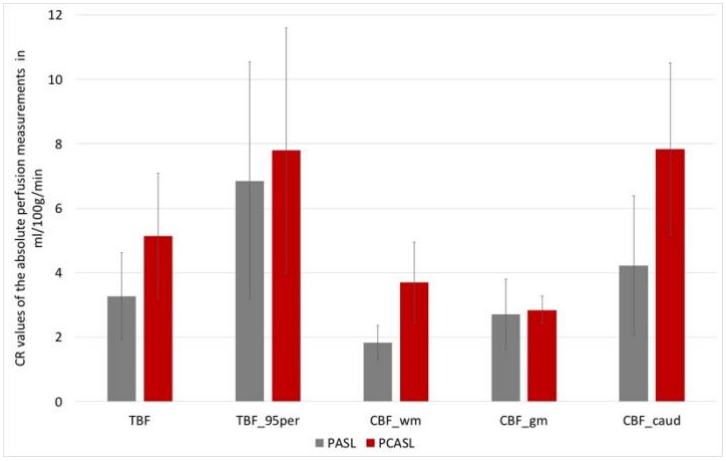
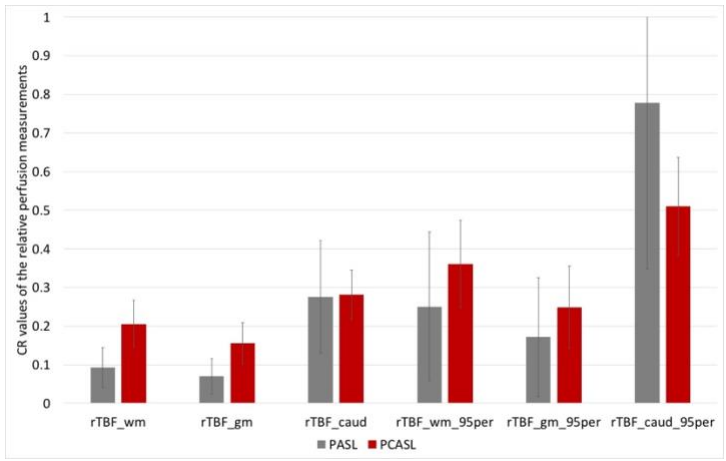


Figure 4.3 %WCoV from PASL and PCASL of (a) the absolute perfusion measurements, (b) the relative perfusion measurements and (c) the SNR from the CNAH.

a.



b.



c.

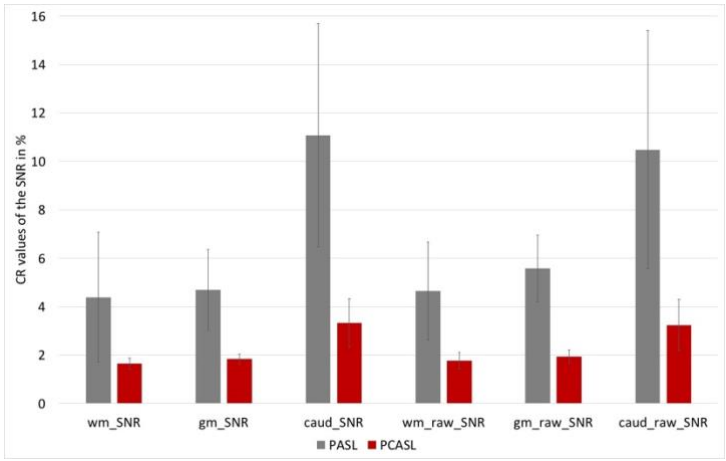


Figure 4.4 Quantified CR from PASL and PCASL of (a) the absolute perfusion measurements in [ml/100g/min], (b) the relative perfusion measurements and (c) the SNR from the CNAH.

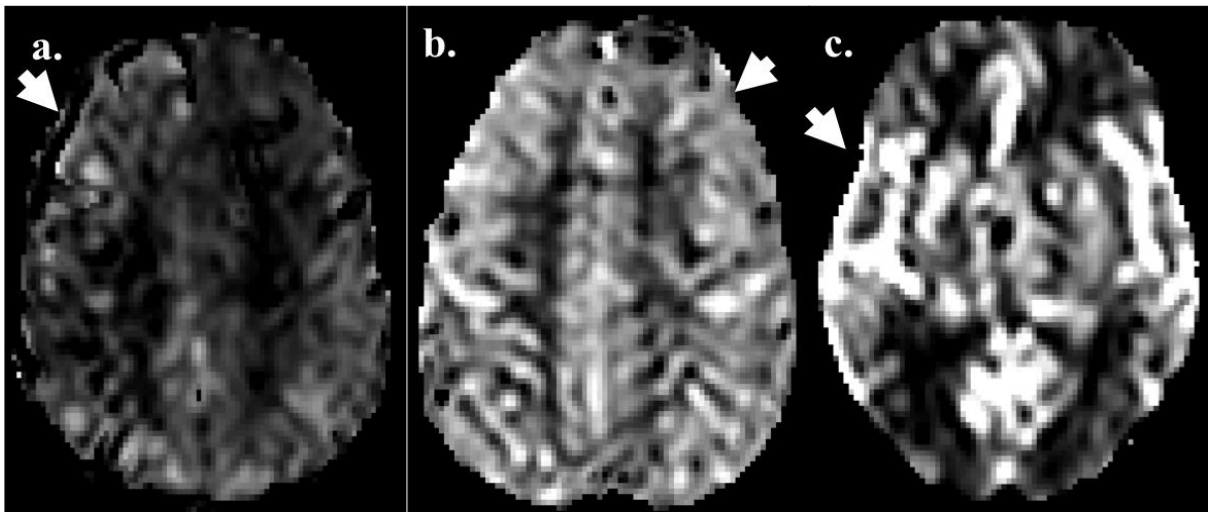
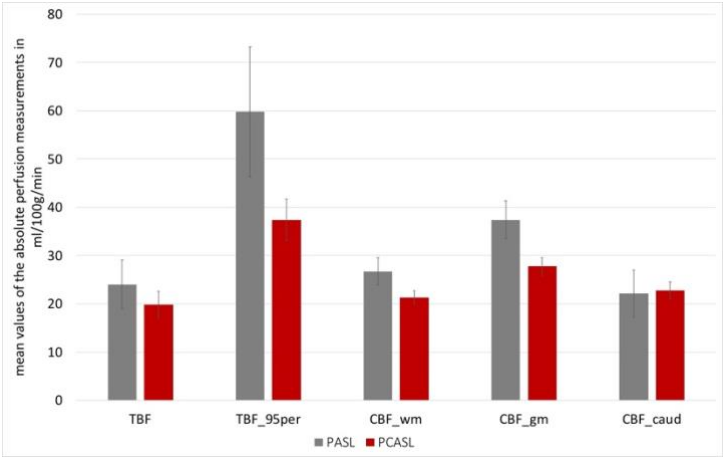
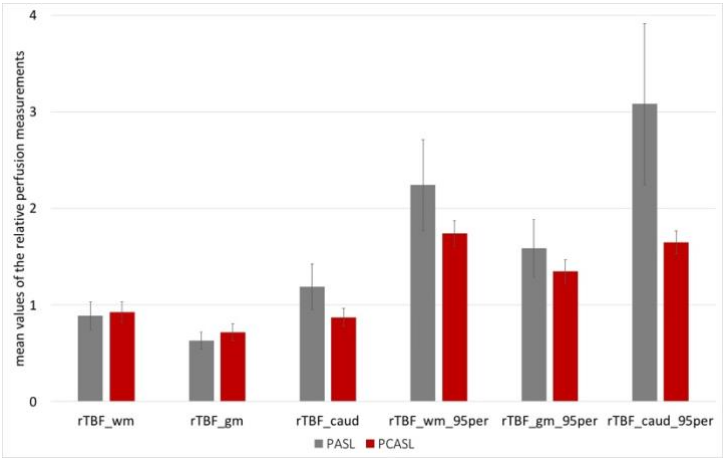


Figure 4.5 An example illustrating the relationship between the WCoV and the ASL-SNR specifically in the CNAH (indicated by arrows). (a.) PCASL CBF-map from for 56-year-old glioma patient, the tumour is located in the left hemisphere while the right hemisphere is the CNAH (arrow). The WCoV values were WM= 7.2%; GM= 5.1%; caud= 16.9%; the high WCoV values from the CNAH presented with low perfusion SNR. (b.) PCASL CBF-map from a 33-year-old glioma patient, the tumour was located at the right hemisphere whereas the left hemisphere is the CNAH (arrow). The WCoV values are WM= 3%; GM= 1.3%; caud= 9.2%; the low WCoV from the CNAH correlated with high perfusion SNR. (c.) PASL CBF-map in a 55-year-old glioma patient, the tumour was located in the left hemisphere while the right hemisphere is the CNAH (arrow). The WCoV values are WM= 3.9%; GM= 4.9%; caud= 10.12%; even though the CNAH showed small WCoV that is due to the high SNR, this is an artificial increase in SNR mostly due to accumulated labelled spins inside the macrovasculature. Note that the images are shown in grayscale to enable visual differentiation between true perfusion (as in a. and b.) and macro-vascular artefact (as in c.)

a.



b.



c.

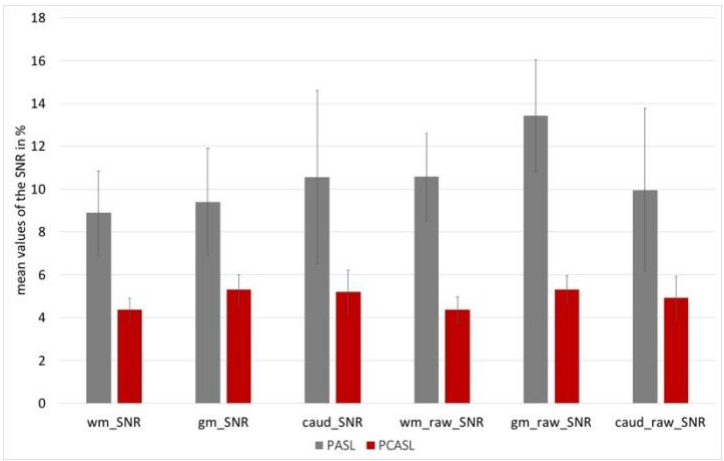


Figure 4.6 Mean values from the PASL and PCASL of (a) the absolute perfusion measurements in [ml/100g/min], (b) the relative perfusion measurements and (c) the SNR from the CNAH.

Table 4.2 between patients mean, SD and 95%CI; within patients ICC, WCoV and CR estimates in the extracted measurements from the CBF maps								
	PASL				PCASL			
	Mean ± SD (95%CI)	ICC (95%CI)	WCoV% (95%CI)	CR (95%CI)	Mean±SD (95%CI)	ICC (95%CI)	WCoV% (95%CI)	CR (95%CI)
aTBF	24.04 ± 9.98 (18.7, 29.4) [ml/100g/min]	0.998 (0.995, 0.999)	4.7 (2.2, 6.3)	3.3 (1.92, 5.017) [ml/100g/min]	19.9 ± 6.56 (16.9, 22.8) [ml/100g/min]	0.996 (0.994, 0.998)	7.9 (6.2 to 9.4)	5.1 (3.2, 6.5) [ml/100g/min]
aTBF_95per	59.8 ± 26.97 (45.4, 74.1) [ml/100g/min]	0.999 (0.997, 0.999)	4.3 (2.6, 5.5)	6.9 (3.166, 9.16) [ml/100g/min]	37.4 ± 10.49 (32.9, 41.8) [ml/100g/min]	0.996 (0.994, 0.998)	6.8 (4.7 to 8.4)	7.8 (3.9, 10.3) [ml/100g/min]
CBF_wm	26.7 ± 5.62 (23.7, 29.7) [ml/100g/min]	0.998 (0.995, 0.999)	2.5 (2, 2.9)	1.8 (1.303, 2.24) [ml/100g/min]	21.3 ± 3.34 (19.9, 22.8) [ml/100g/min]	0.992 (0.987, 0.996)	5.8 (4.6 to 6.9)	3.7 (2.4, 4.6) [ml/100g/min]
CBF_gm	37.4 ± 7.89 (33.2, 41.6) [ml/100g/min]	0.997 (0.995, 0.999)	2.7 (1.7, 3.5)	2.7 (1.64, 3.472) [ml/100g/min]	27.8 ± 4.49 (25.9, 29.7) [ml/100g/min]	0.997 (0.996, 0.999)	3.8 (3.2 to 4.3)	2.8 (2.4, 3.2) [ml/100g/min]
CBF_caud	22.1 ± 9.92 (16.8, 27.4) [ml/100g/min]	0.996 (0.992, 0.998)	8.4 (4.1, 11.2)	4.2 (2.057, 5.605) [ml/100g/min]	22.8 ± 4.43 (20.9, 24.7) [ml/100g/min]	0.980 (0.966, 0.990)	11.3 (8.8 to 13.4)	7.8 (5.2, 9.8) [ml/100g/min]
rTBF_wm	0.89 ± 0.29 (0.73, 1.04)	0.998 (0.996, 0.999)	3.86 (1.9, 5.1)	0.092 (0.040, 0.124)	0.92 ± 0.26 (0.81, 1.03)	0.996 (0.993, 0.998)	7.5 (6.3 to 8.7)	0.205 (0.144, 0.252)
rTBF_gm	0.63 ± 0.18 (0.53, 0.72)	0.997 (0.993, 0.999)	3.98 (1.7, 5.3)	0.07 (0.023, 0.096)	0.71 ± 0.22 (0.62, 0.81)	0.997 (0.995, 0.998)	7.1 (5.7 to 8.3)	0.156 (0.1025, 0.195)

rTBF_caud	1.19 ± 0.47 (0.94, 1.4)	0.993 (0.986, 0.997)	8.1 (4, 10.7)	0.27 (0.128, 0.366)	0.87 ± 0.23 (0.77, 0.97)	0.990 (0.983, 0.995)	10.9 (9.2 to 12.3)	0.281 (0.218, 0.333)
rTBF_wm_95per	2.24 ± 0.94 (1.7, 2.7)	0.998 (0.997, 0.999)	3.4 (2.3, 4.3)	0.24 (0.055, 0.36)	1.74 ± 0.32 (1.61, 1.88)	0.992 (0.986, 0.996)	7.06 (5.58 to 8.3)	0.361 (0.248, 0.445)
rTBF_gm_95per	1.59 ± 0.60 (1.3, 1.9)	0.998 (0.997, 0.999)	3.3 (2.17, 4.15)	0.17 (0.0175, 0.24)	1.35 ± 0.29 (1.22, 1.47)	0.995 (0.992, 0.998)	6.19 (4.47 to 7.5)	0.249 (0.142, 0.322)
rTBF_caud_95per	3.08 ± 1.67 (2.2, 3.9)	0.995 (0.991, 0.998)	8.7 (3.9, 11.7)	0.77 (0.347, 1.045)	1.65 ± 0.29 (1.52, 1.77)	0.979 (0.964, 0.989)	10.5 (8.7 to 11.9)	0.51 (0.38, 0.611)
CBF_wm_SNR	8.9 ± 3.94 (6.8, 10.9)	0.968 (0.936, 0.987)	14.8 (11.7, 17.4)	4.4 (1.69, 5.97)	4.36 ± 1.36 (3.79, 4.9)	0.985 (0.975, 0.993)	14.9 (12.2 to 17.2)	1.65 (1.43, 1.84)
CBF_gm_SNR	9.4 ± 4.98 (6.7, 12)	0.978 (0.965, 0.991)	17.1 (11.8, 21.01)	4.7 (3.036, 5.91)	5.32 ± 1.70 (4.6, 6.04)	0.989 (0.982, 0.995)	13.41 (11.29 to 15.24)	1.84 (1.63, 2.028)
CBF_caud_SNR	10.6 ± 8.08 (6.2, 14.9)	0.955 (0.909, 0.982)	33. 3 (25.4, 39.6)	11.1 (6.481, 14.28)	5.2 ± 2.47 (4.15, 6.24)	0.954 (0.923, 0.977)	24.7 (19.291 to 29.15)	3.3 (2.33, 4.097)
ΔASL_wm _SNR	10.6 ± 4.09 (8.4, 12.7)	0.969 (0.938, 0.988)	14.8 (10.9, 17.8)	4.6 (2.619, 6.031)	4.4 ± 1,44 (3.8, 4.99)	0.983 (0.972, 0.992)	15.7 (12.27 to 18.54)	1.8 (1.43, 2.055)
ΔASL_gm_SNR	13.4 ± 5.22 (10.6, 16.2)	0.971 (0.941, 0.988)	16. 1 (11.5, 19.5)	5.6 (4.2, 6.69)	5.3 ± 1.59 (4.6, 5.98)	0.986 (0.976, 0.993)	13.6 (11.6 to 15.4)	1.9 (1.67, 2.183)

ΔASL_caud_SNR	9.9 ± 7.61 (5.9, 14)	0.956 (0.910, 0.982)	34.1 (22.7, 42.5)	10.5 (5.6, 13.74)	4.9 ± 2.47 (3.88, 5.96)	0.957 (0.928, 0.978)	24.6 (18.8 to 29.2)	3.2 (2.187, 4.025)

Table 4.3 Paired t-test results comparing the WCoV% between aTBF and rTBF, between CBF and aTBF, and between CBF and rTBF

	PASL		PCASL	
	t	p-value	t	p-value
aTBF and rTBF-WM	1.3	> 0.05	0.28	> 0.05
aTBF and rTBF-GM	1.2	> 0.05	1.9	> 0.05
aTBF and rTBF-caud	-2.7	0.016	-3.8	0.001
aTBF and rTBF-WM-95per	1.1	> 0.05	1.2	> 0.05
aTBF and rTBF-GM-95per	1.5	> 0.05	3.6	0.002
aTBF and rTBF-caud-95per	-2.8	0.015	-2.8	0.010
aTBF-95per and rTBF-WM	0.98	> 0.05	-1.9	> 0.05
aTBF-95per and rTBF-GM	0.72	> 0.05	-1.1	> 0.05
aTBF-95per and rTBF-caud	-2.5	0.023	-4.7	<0.001
aTBF-95per and rTBF-WM-95per	1.3	> 0.05	-1	> 0.05
aTBF-95per and rTBF-GM-95per	1.5	> 0.05	1.5	> 0.05
aTBF-95per and rTBF-caud-95per	-2.6	0.019	-3.9	0.001
CBF-WM and aTBF	-1.9	> 0.05	-2.5	0.019
CBF-WM and aTBF-95per	-2.3	0.035	-0.95	> 0.05
CBF-WM and rTBF-WM	-1.7	> 0.05	-2.8	0.010
CBF-WM and rTBF-WM-95per	-1.9	> 0.05	-2.3	0.029
CBF-WM and rTBF-GM	-1.6	> 0.05	-1.7	> 0.05
CBF-WM and rTBF-GM-95per	-1.6	> 0.05	-0.18	> 0.05
CBF-WM and rTBF-caud	-3.4	0.004	-10.6	<0.001
CBF-WM and rTBF-caud-95per	-3.5	0.003	-9.5	<0.001
CBF-GM and aTBF	-2.3	0.037	-5	<0.001
CBF-GM and aTBF-95per	-2.8	0.013	-3.7	<0.001
CBF-GM and rTBF-WM	-1.8	> 0.05	-5.5	<0.001
CBF-GM and rTBF-WM-95per	-1.9	> 0.05	-5.1	<0.001
CBF-GM and rTBF-GM	-1.5	> 0.05	-4.8	<0.001
CBF-GM and rTBF-GM-95per	-1.4	> 0.05	-3.3	0.003

CBF-GM and rTBF-caud	-3.3	0.005	-10.6	<0.001
CBF-GM and rTBF-caud-95per	-3.4	0.004	-9.5	<0.001
CBF-caud and aTBF	2.9	0.010	3.7	0.001
CBF-caud and aTBF-95per	2.8	0.014	5.5	<0.001
CBF-caud and rTBF-WM	3	0.008	3.7	0.001
CBF-caud and rTBF-WM-95per	2.9	0.012	4.9	<0.001
CBF-caud and rTBF-GM	3.4	0.004	4.2	<0.001
CBF-caud and rTBF-GM-95per	3.3	0.005	5.6	<0.001
CBF-caud and rTBF-caud	0.75	> 0.05	0.024	> 0.05
CBF-caud and rTBF-caud-95per	-0.13	> 0.05	0.6	> 0.05

Table 4.4 Mann-Whitney U Test and ranks comparing the WCoV of the extracted measurements between PASL and PCASL						
	ASL_method	N	Mean Rank	Mann-Whitney U	Z	Exact Sig. (2-tailed)
wcov_TBF	PASL	16	13.94	87	-2.899	0.003
	PCASL	24	24.88			
	Total	40				
wcov_TBF_95per	PASL	16	14.56	97	-2.623	0.008
	PCASL	24	24.46			
	Total	40				
wcov_CBF_WM	PASL	16	10.38	30	-4.472	<0.001
	PCASL	24	27.25			
	Total	40				
wcov_CBF_GM	PASL	16	14.88	102	-2.485	0.012
	PCASL	24	24.25			
	Total	40				
wcov_CBF_caud	PASL	16	15.13	106	-2.374	0.017
	PCASL	24	24.08			
	Total	40				
wcov_rTBF_WM	PASL	16	11.63	50	-3.920	<0.001
	PCASL	24	26.42			
	Total	40				
wcov_rTBF_GM	PASL	16	11.38	46	-4.031	<0.001
	PCASL	24	26.58			
	Total	40				
wcov_rTBF_caud	PASL	16	12.25	60	-3.644	<0.001
	PCASL	24	26.00			
	Total	40				
wcov_rTBF_WM_95per	PASL	16	13.44	79	-3.119	0.001
	PCASL	24	25.21			
	Total	40				

wcov_rTBF_GM_95per	PASL	16	13.81	85	-2.954	0.003
	PCASL	24	24.96			
	Total	40				
wcov_rTBF_caud_95per	PASL	16	14.25	92	-2.761	0.005
	PCASL	24	24.67			
	Total	40				
wcov_WM_SNR	PASL	16	20.56	191	-0.028	0.989
	PCASL	24	20.46			
	Total	40				
wcov_GM_SNR	PASL	16	22.63	158	-0.939	0.359
	PCASL	24	19.08			
	Total	40				
wcov_caud_SNR	PASL	16	25.50	112	-2.209	0.027
	PCASL	24	17.17			
	Total	40				
wcov_raw_wm_SNR	PASL	16	19.44	175	-0.469	0.652
	PCASL	24	21.21			
	Total	40				
wcov_raw_gm_SNR	PASL	16	21.63	174	-0.497	0.633
	PCASL	24	19.75			
	Total	40				
wcov_raw_caud_SNR	PASL	16	25.13	118	-2.043	0.041
	PCASL	24	17.42			
	Total	40				

4.4 Discussion

In this study, PASL and PCASL repeatability in gliomas was evaluated using three indices, namely ICC, WCoV and CR. As both patient groups were not the same across the ASL labelling approaches, the evaluation was undertaken separately. Our findings demonstrated here that aTBF WCoV was not different than rTBF WCoV when a large ROI was used as an internal reference. This suggest that when normalizing TBF values, a large ROI is better than a small one to avoid artefactual increased in noise. In addition, the CR (or SRD) was lowest for the rTBF when normalised to the large ROIs, where a small CR indicates high sensitivity to changes [184]. In a recent study by Zhou et al. the measured WCoV was lower for large ROIs than for small ROIs, for example WCoV from temporal lobe was 4.95% while from the putamen was 7.12%. This is supported by the inverse relationship between the within-subject variation and the ROI size seen in ASL and reported in several studies [185,186]. Our result from chapter 3 and previous studies reported the high diagnostic accuracy for glioma grading using rTBF, and also the role of the rTBF in the improvement of the detection of small differences of perfusion in the case of moderate SNR [45,187]. The lack of difference in WCoV in this study between both indices might be due to the relatively artificial way of estimating repeatability indices, based on split measurements. Generally, however, in clinical studies where an internal reference is chosen, it is important to select it carefully, preferably a large ROI based on GM within the CNAH.

The reliability for both PASL and PCASL was excellent (ICC>0.9) in concordance with previous studies, where it ranged from good to excellent [151,171,188]. We believe that the high-reliability in our measurements was due to the short period between the repeated measurements, as previous ASL reliability studies showed the adverse effect of long time interval among the scan sessions [151,188]. The slight superiority of PASL over PCASL in reliability in this study is primarily due to the difference in acquisition parameters. In particular, the voxel size between both methods was ~1.7 times larger for our PASL implementation than for the PCASL one. In addition, presence within the PASL data of a slightly higher between patient SD might also contribute to the ICC increases, as is generally the case when between patient SD is greater than the within patient SD [189].

The WCoV of CBF in CNAH was found to be comparable with previous studies [151,170,185]. Higher WCoV in the perfusion metrics (both for normal CBF and tumour) was observed in PCASL compared to PASL. Probable reasons are the same as before, including intrinsic differences in the two patient groups, the acquisition parameters, and in particular the voxel size. Besides, the difference in labelling efficiency between both sequences, with PASL having a generally slightly higher efficiency than PCASL, might have also possibly contributed [190].

The higher WCoV of PCASL-based estimation of CBF does not necessarily mean the superiority of PASL over PCASL, as both labelling methods employed were applied to different patient groups. The quantified means from the CNAH (CBF-WM, CBF-GM and CBF-caud) were higher in PASL than in PCASL. In addition, PCASL has been reported to display higher SNR and reproducibility than PASL when applied to the same patient group [151]. Furthermore, in our study, the WCoV of the SNR was lower for PCASL for small regions and showed a trend towards reduced WCoV in the rest of the brain, which indicates a potential higher stability in comparison to PASL.

It is worth noting that there is an inverse relationship between the amplitude of the ASL signal and its variation (WCoV) across time [151]. As a matter of fact, even an artefactual increase in the SNR from the accumulated labelling spins in the macro-vessels will reduce the WCoV, if it can be measured reproducibly. Therefore, the decrease in the WCoV of the PASL perfusion metrics might be a consequence of the higher ASL signal due to labelled spins that have not yet reached the microcirculation. This phenomenon is also known as a macro-vascular artefact or Arterial Transit Artefact (ATA) (see Figure 4.5) and is mainly due to the delayed arrival of the labelled blood to the tissue.

This underscores the significant role of the PLD selection and its impact on the robustness of the CBF quantification. This is most critical in tumours with abnormal transit times due to the newly formed tortuous vessels. A recent ASL study suggested using phase-contrast MRI to determine the blood velocity profile and inform the selection of the PLD [191]; this might help to ensure complete delivery of the labelled bolus into the tissue of interest, however at the cost of scan time, and thus might not be practically implementable in the clinics.

The main limitation of our study is its small sample size of the repeated measurements, which has been addressed using moving block bootstrapping [176]. In this study, we focused on comparing the short-term repeatability of the different ASL acquisition techniques and did not address within-subject variation due to patient repositioning and physiological tissue perfusion perturbations and changes, which is of particular importance for longitudinal tumour studies. However, based on the findings of the large QUASAR study [170], an estimate of these additional effects could be established as a guiding rule regarding sample size calculation for future clinical trials. Most importantly, however, was the recognition that repeatability was not lower in primary brain tumour patients than in healthy volunteers, as assessed in previous studies. This is particularly remarkable due to the use of split acquisition in blocks, resulting in a generally lower average scan time than used clinically.

4.5 Conclusion

This repeatability study highlights the excellent reliability of both PASL and PCASL for neuro-oncological studies, in line with previous volunteer studies. However, aTBF/rTBF estimates tend to maintain high within-subject variability that may obscure the detection of between-subject variation with small changes in tumour perfusion. When estimating rTBF, large ROIs need to be used to filter out systemic misleading within-subject variation and thus provide more robust perfusion measurements. Finally, in this study, no large difference between repeatabilities were observed between PASL and PCASL, which might have been attributed primarily to the difference in acquisition parameters, and both methods can therefore be recommended to assess brain perfusion, with the caveat that PCASL allows for longer PLD than PASL, which therefore minimises the macro-vascular artefacts, and provide better estimates of patient perfusion.

Chapter 5

5 The Diagnostic Performance of Different Advanced-MRI Based on Radiomics for Noninvasive Histomolecular Staging of Untreated Adult Gliomas 145

5.1 Introduction 146

5.2 Material and Methods 147

5.2.1 Study design 147

5.2.2 Patients 147

5.2.3 Hand-crafted radiomics 148

5.2.4 Image acquisition and processing 150

5.2.5 Image analysis and feature extraction 155

5.2.6 Feature reduction and statistical analysis 156

5.3 Results 158

5.3.1 Histogram features 158

5.3.2 Diagnostic performance 158

ii. The significant features separately 158

iii. The combined significant features for different MRI methods 163

5.4 Discussion 168

5.5 Conclusion 174

5 The Diagnostic Performance of Different Advanced-MRI Based on Radiomics for Non-invasive Histomolecular Staging of Untreated Adult Gliomas

Abstract

Radiomics has emerged as a promising non-invasive method for glioma classification. This paper aims to explore the diagnostic performance of advanced MRI methods to classify primary gliomas. Prospective data from 32 adults with untreated gliomas, who underwent pseudo-continuous arterial spin labelling (PCASL), dynamic susceptibility contrast-enhanced (DSC), dynamic contrast-enhanced (DCE) and diffusion-weighted imaging (DWI) were post-processed. The tumours were staged according to the WHO 2016 histomolecular classification scheme. The DCE datasets were post-processed using the modified Tofts-Kermode (mTK) and the Lawrence and Lee (L&L) models. Whole tumour histogram features from the perfusion- and diffusion-weighted maps were extracted. The features have been combined using Kruskal-Wallis test, backward elimination and cross-correlation methods. Multinomial logistic regression analysis with leave-one-out cross validation (LOOCV) was undertaken on these features separately and combined in sets. The validation model with the highest Kappa represents the best-performed model and hence the optimal features' subsets. The model performance was reported using several indexes including sensitivity, specificity, F1 score, accuracy (ACC) and Cohen's Kappa (K). The diagnostic performances of the ASL, DSC and DCE for the attained features separately were comparable to combining optimal features' subsets, where the latter made it feasible to predict more classes. The optimal features' subsets from DSC showed superior diagnostic performance, with ACC for grading 84%, sub-grading 72%, IDH-status 86%, grouping 66% and sub-grouping 56%. Followed by ADC for grading (ACC=70%) and sub-grading (ACC=61%), DCE-L&L for grouping (ACC=60%). Both DCE-L&L and ASL showed the second-best diagnostic performance for IDH-status (ACC=81%). DCE-mTK surpass DCE-L&L for grading (ACC: 56% vs 72%). MRI radiomics hold promise for accurate glioma classification. ASL and ADC represent a gadolinium-free approach poses a potential alternative to DSC. Complex pharmacokinetic models yield better diagnostic performance than simplified compartmental solutions in DCE-MRI.

5.1 Introduction

The WHO 2016 classification[2] of brain gliomas mandates tumour subtyping according to molecular/genetic features including stratification by isocitrate dehydrogenase gene mutation (IDH-mut) and 1p/19q co-deletion, which are favourable for longer overall survival time[7–9]. Currently, histological biopsy remains the reference standard for identifying glioma grade and their molecular/genetic features. However, due to the method invasiveness and possible sampling errors, surrogate non-invasive imaging biomarkers are of great interest. Advanced MRI may qualify as a valid alternative to invasive biopsy for gliomas staging. Recently, several studies reported the subsidiary role of perfusion and diffusion-weighted (DWI) MRI[84,192–194] in staging gliomas according to the WHO 2016. Rarely, the results have so far captured the heterogeneity of the lesions. Instead, the results are based on univariate analyses of maximum and mean values from the hot-spot regions of interest (ROI) or the entire tumour volume. However, useful information, imperceptible by the human eye, known as the texture, can be extracted from the radiological images using histogram texture analysis[60,63,64]. The latter explores the distribution of the voxels' values, and hence tumour heterogeneity. The high dimensional data, collectively known as radiomics, generated from such analysis can subsequently be reduced, e.g. preserving its significant aspects, to ensure accuracy and limit complex elements for the prediction or classification task.

MRI radiomics has been reported as a promising non-invasive method for glioma classification[195–200]. Tian et al. [195] reported the effectiveness of radiomics for glioma grading using different MRI modalities including structural MRI, DWI and arterial spin labelling (ASL); notably, the combined multiparametric MRI radiomics demonstrated the highest grading accuracy. Other studies reported radiomics from conventional MRI as a potentially useful approach for predicting the IDH-mutation among low-grade[197–199] and high-grade gliomas[200] (LGGs, HGGs, respectively). An interesting remark is that all studies investigating dynamic contrast-enhanced (DCE) perfusion MRI in gliomas have utilised simplified compartmental models, particularly the modified Tofts-Kermode (mTK) model[84–86]. This model cannot clearly separate the intravascular transport of the

tracer molecules relative to their exchange process between intravascular and EES. Thus, the Ktrans from the mTK model incorporate both plasma flow (F) and vessels permeability surface (PS). In contrast, advanced models, such as the adiabatic approximated tissue homogeneity (AATH) model also known as St Lawrence and Lee model (L&L)[201] allowed separate estimation of F and PS which are respectively reflect the transport of the tracer molecules intravascular and between the intravascular space and the EES; hence they reflect the underlying physiology more accurately and enable additional physiological parameters to be better estimated[80,88], but demanding higher data quality.

Our study aims to expand the work of the previous radiomics studies in gliomas by investigating the diagnostic performance of multiple advanced MRI methods including the three dominant perfusion techniques. Namely, pseudo-continuous-ASL (PCASL) dynamic contrast-enhanced (DCE) using both the modified Tofts-Kermode (mTK) and the St Lawrence and Lee (L&L) models, and the dynamic susceptibility contrast-enhanced (DSC) along with DWI in order to classify gliomas using a tiered approach that considers conventional grading (*LGGs (grade-II)* and *HGGs (grade III-IV)*) and WHO 2016 histomolecular staging including glioblastoma IDH-wild type (*GBM-wt*), grade II IDH-mutant astrocytoma (*astro-mut-LG*), grade III IDH-mutant astrocytoma (*astro-mut-HG*), grade II IDH-mutant oligodendroglioma (*oligo-mut-LG*), and grade III IDH-mutant oligodendroglioma (*oligo-mut-HG*).

5.2 Material and Methods

5.2.1 Study design

The used data is from a prospective, non-blinded imaging trial approved by the National Research Ethics Service and the local Joint Research Office with written consent obtained from all recruited patients.

5.2.2 Patients

Screening and selection of the study participants were conducted in the period from January 2018 to June 2019 through the neuro-oncology multidisciplinary tumour board in our hospital. The study participants were prospectively identified as adult patients with primary non-treated gliomas and able to undergo

multiparametric MRI scans. At the same time, patients with the following conditions: brain tumours other than gliomas, contraindication to MRI and gadolinium, and previously treated brain tumours were excluded from the study.

5.2.3 Hand-crafted radiomics

This study used the conventional hand-crafted radiomic pipeline (see Figure 5.1), that consists of the following tasks: 1) image acquisition and processing including map generation and ROI segmentation; 2) feature extraction; 3) feature reduction and 4) statistical analyses and model validation using LOOCV.

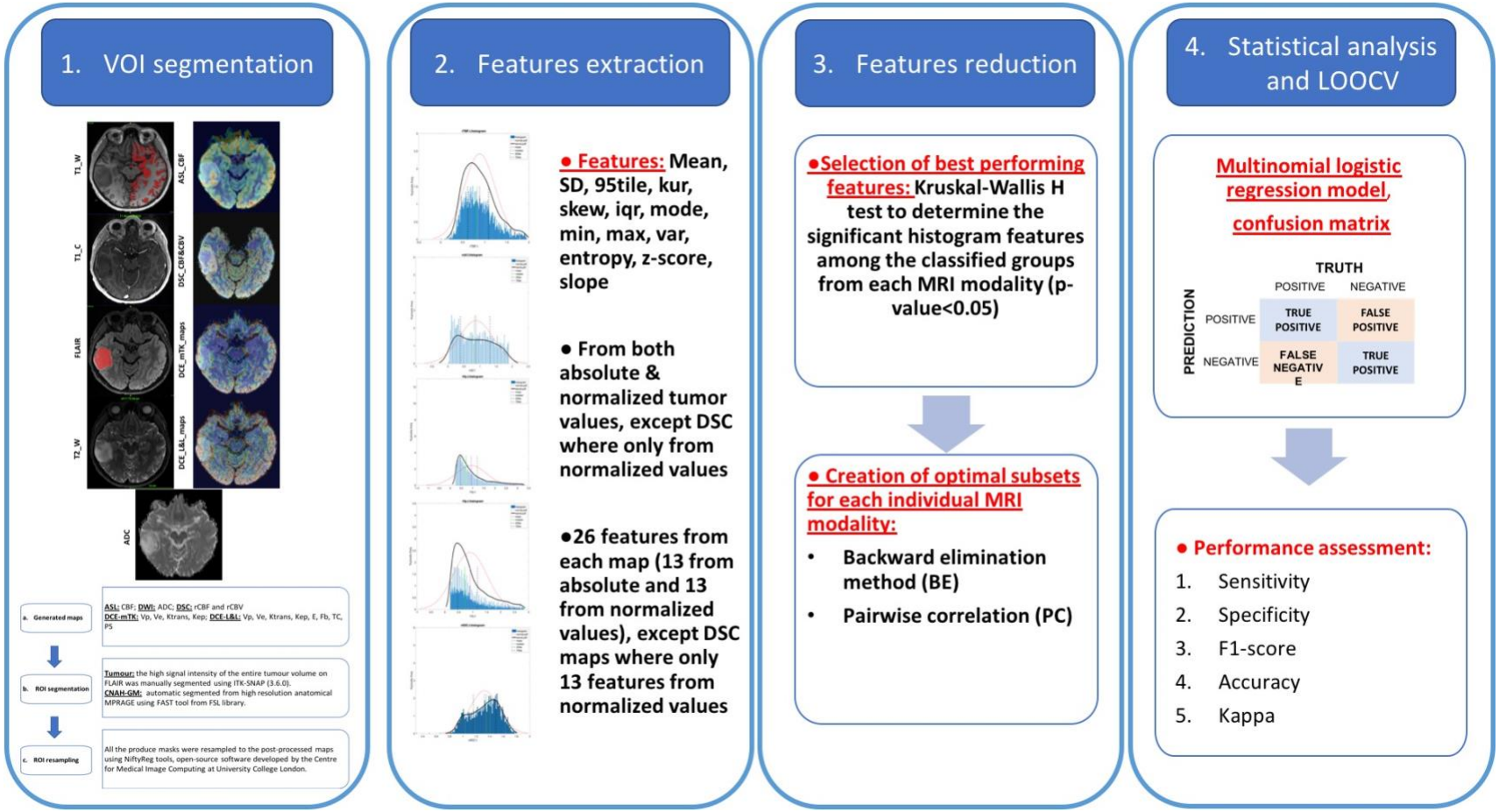


Figure 5.1 Summary of the image analysis workflow: 1) volumes of interest (VOIs) after tumour segmentation in the T2-weighted FLAIR, grey matter (GM) segmentation from the contra-lateral normal-appearing hemisphere on 3D-T1-weighted images, T2-weighted and post-contrast T1-weighted images were available as reference. The segmented VOIs were then resampled and overlaid onto the generated maps; 2) Histogram features extraction; 3) Feature reduction and determination of the optimal features' subsets; 4) Statistical analyses and model validation using leave-one-out-cross-validation (LOOCV) to enable the assessment of the diagnostic performance of the final features' subsets.

5.2.4 Image acquisition and processing

3D high-resolution T1-, post-contrast T1-, T2- and T2/FLAIR-weighted structural imaging, DWI-MRI and multimodal MR-perfusion (PCASL, DCE and DSC) were acquired on a clinical 3T MR scanner (Prisma, Siemens Healthineers, Erlangen, Germany) using a 64-channel head coil. DWI was acquired with diffusion weightings in 3 orthogonal directions at two b values (0 and 2500 s/mm^2), the choice of the b-values was justified by the increased differentiation capability for glioma grade at higher b-values[202]. Table 5.1 shows the technical details of the sequence protocol. PCASL was chosen to obtain ASL data and was guided by phase-contrast scout images to position the labelling plane where carotid and vertebral arteries run along the z-direction (see Table 5.1). A proton density (PD) image, for scaling purposes, was acquired as the first volume of the series, with the same imaging readout but without labelling or background suppression.

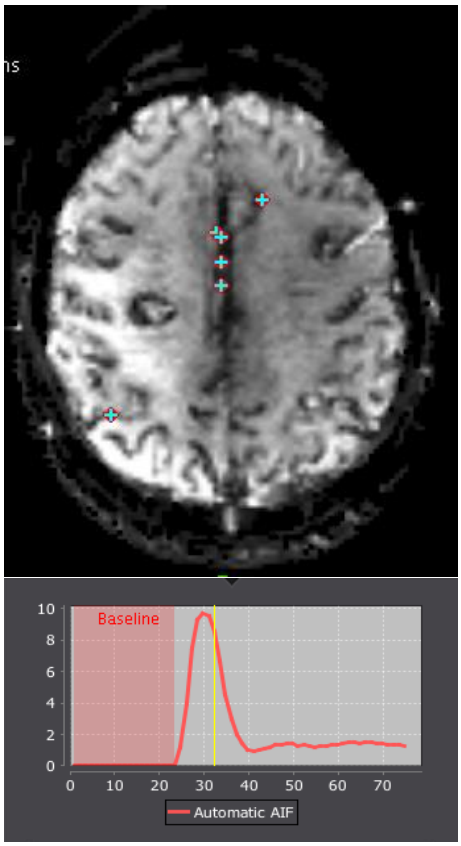
Following DWI and ASL, both contrast-based perfusion imaging modalities were obtained using body-weight adapted contrast agent (Prohance, Bracco, Milan, Italy) administered in two separated boluses using a power injector at a rate of 4ml/sec and flushed each time with 20ml of normal saline at the same injection rate. The first bolus intended for the DCE-MRI (~5 minutes acquisition time) acted at the same time as preload for the DSC-MRI[76]. Table 5.1 illustrates the sequence acquisition parameters for DCE- and DSC-MRI including the acquisition details for the multiple flip angle-based T_{10} (3, 7, 10, 13, 15 degrees) brain mapping prior to gadolinium injection.

Commercially available software (Olea Sphere 2.3, Olea Medical, La Ciotat, France) was used to process the DCE- and DSC-MRI data. DCE-MRI datasets were fitted to the modified Tofts-Kermode compartmental-model (mTK)[87], as well as to the St Lawrence and Lee (L&L)[201] distribution-model. This occurred after motion correction was performed on the dynamic data and automatic selection of the arterial input function (AIF) was performed for every patient based on a cluster analysis algorithm[203], an example of the AIF and the tumour uptake curve are illustrated in Figures 5.2 and 5.3. The mTK model produced four maps: fractional plasma volume (v_p, in %), fractional volume of the extracellular extravascular space (EES) (v_e, in %), transfer constant from blood plasma or permeability

constant (K^{trans}, in min^{-1}) and reverse transfer constant to the blood plasma or reflux constant (k_{ep}, in min^{-1}). On the other hand, the L&L model output has eight maps including extraction fraction (E, in %, denoting the extracted fraction of the gadolinium from the intravascular to the EES); plasma flow (Fb, in ml/min^{-1}); mean capillary transit time (Tc, in a sec); permeability surface product (PS, ml/min^{-1}); influx to the EES (K^{trans}, in min^{-1}); K_{ep}, v_e and v_p. The DSC datasets underwent also motion correction, spatial smoothing and automatic AIF selection for every patient based on a cluster analysis algorithm[203], an example of the AIF and the tumour uptake curve are illustrated in Figures 5.2 and 5.3. The macro-vascular contamination was accounted for by using a Bayesian model approach[79]. Leakage-corrected relative cerebral blood volume (rCBV) and relative cerebral blood flow (rCBF) maps were generated. In this study, rCBV and rCBF refer to the normalised to the rCBV values in the healthy brain (see below).

The acquired ASL raw data were pre-processed using MCFLIRT toolbox for motion correction[175], the ASL mean difference images between control and label conditions (ΔASL) were calculated[204], and the ASL-derived CBF-maps were generated using the simplified general kinetic model[100]. The MR scanner automatically generated ADC maps using the vendor's software.

a.



b.

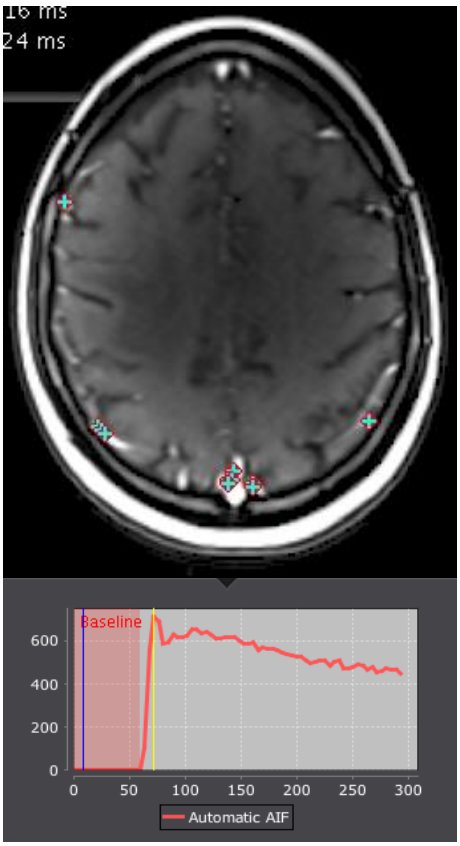
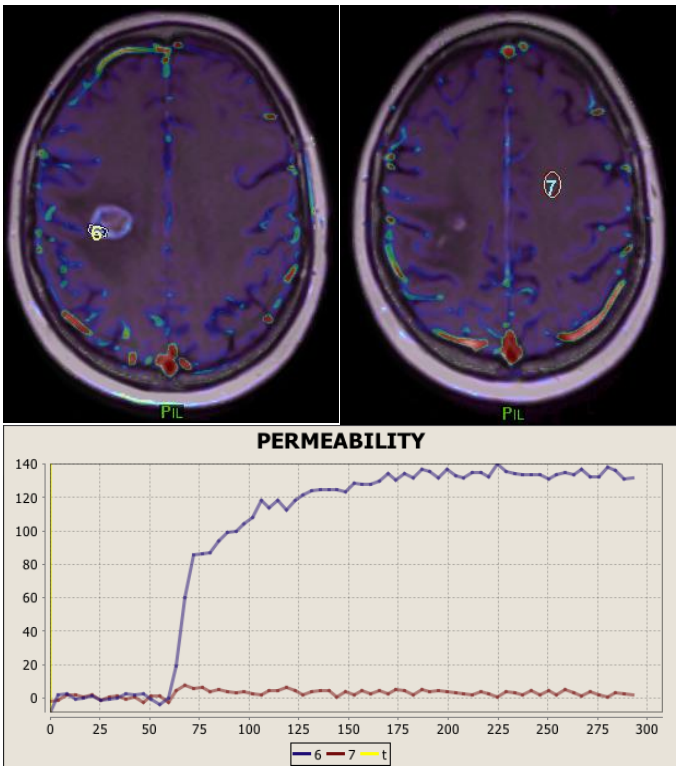


Figure 5.2 Examples of the arterial input function (AIF) a)from Dynamic susceptibility contrast (DSC), bb)from Dynamic contrast-enhanced (DCE),

vp map on T1-c

a.



rCBV map on

b.

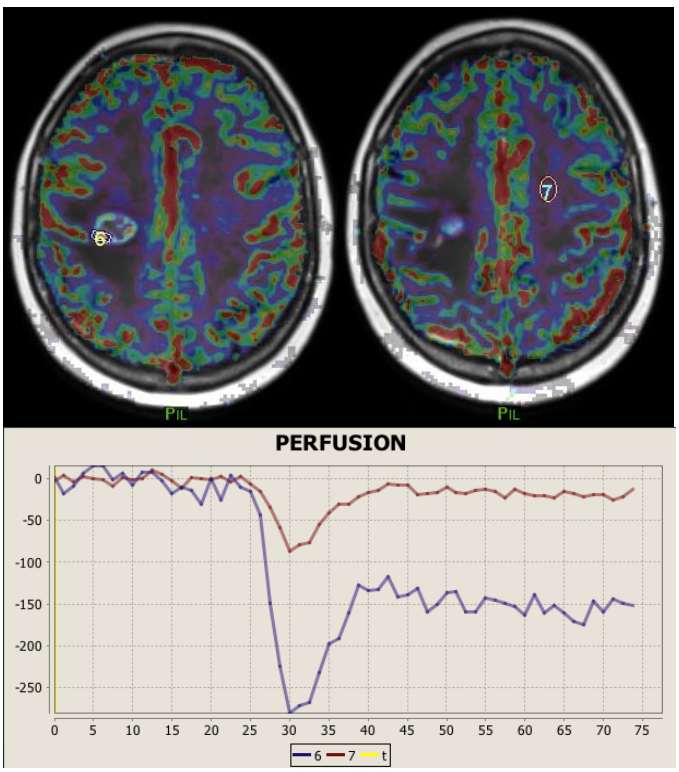


Figure 5.3 Example of the glioma uptake curve from ROI6 versus normal appeared white matter from ROI 7, a) from Dynamic susceptibility contrast (DSC), b) from Dynamic contrast-enhanced (DCE).

Table 5.1 Imaging parameters for DCE, DSC, DWI and ASL

	DCE	DSC	DWI	PCASL
sequence	Fast low angle shot (FLASH)-3D	Echo-planar imaging (EPI)-2D	Echo-planar imaging (EPI)-2D	3D GRASE
b-value (s/mm^2)	-	-	0, 2500	-
slices	20	20	54	28
Gap between slices (mm)	-	5.2	2.5	-
Dynamic measurements	70	58	-	-
TR/TE (ms)	3.16/1.24	1370/30	3600/79	4600/13.28
FOV (mm)	224 x 168	220 x 220	220 x 220	220 x 220
Matrix size	224 x 168	128 x 128	88 x 88	64 x 60
Voxel size (mm^3)	1 x 1 x 4	1.7 x 1.7 x 4	2.5 x 2.5 x 2.5	3.44 x 3.67 x 4
Slice thickness (mm)	4	4	2.5	4
Averages (NSA)	1	1	1	1
flip angle (°)	8	60	90	130
echo train length (ETL) or Turbo factor	1	63	38	Turbo / EPI factors: 28/15
CA	ProHance (Gadoteridol)	ProHance (Gadoteridol)	-	-
CA dose	2ml/10kg	2ml/10kg	-	-
CA injection rate	4ml/sec	4ml/sec	-	-
Delay before imaging	20 sec	20 sec	-	-
Scanning time	4min 57s	1min 26s	4min 35s	~1min
Number of repeats/averages	-	-	-	2
Echo spacing ms	-	-	-	0.51
Labelling plane thickness (mm)	-	-	-	10
Gap between labelling and imaging plane (mm)	-	-	-	45
Bolus duration (ms)	-	-	-	1800
Post-labelling delay time (PLD/TI) (ms)	-	-	-	1800

5.2.5 Image analysis and feature extraction

Figure 5.1 outlines the workflow of the image analysis. All MR images were converted to the Neuroimaging Informatics Technology Initiative (NIfTI) format. Conventional MR images (T1-weighted T1-W; T1 weighted with contrast T1-c, T2-weighted T2-W, and T2-FLAIR) were used as anatomical references. ITK-SNAP[177] (version 3.6.0; www.itksnap.org) was used to manually segment the entire tumour volume (VOI) on the FLAIR images avoiding the cystic, necrotic and haemorrhagic regions. The high resolution anatomical T1-W scan was used for automatic segmentation of the grey matter (GM) of the contralateral normal-appearing hemisphere (CNAH)[178]. In cases where the tumour crossed over to the contralateral hemisphere, tumour voxels were excluded from the segmented ROIs in the CNAH. NiftyReg, an open-source software[179], was used to resample all segmented ROIs to the generated multiparametric maps employing rigid registration followed by affine registration. Histogram features were extracted from the ASL, ADC, DSC and DCE maps using in-house MATLAB-based scripts (MATLAB 2017b, MathWorks Natick, MA, USA). For the perfusion modalities, the values over the 95-percentile and the values lower than 5-percentile were excluded, to avoid macrovascular and noise artefacts, respectively. All the values of the tumour pixels in ADC were entered into the analysis. The extracted histogram features were the mean; standard deviation (SD); 95th-percentile (95tile); kurtosis (kur); skewness (skew); median; interquartile range (iqr); mode; minimum (min); maximum (max); and variance (var). Also, histogram entropy (randomness) calculated based on probabilities of the histogram with 0.01 bin width; median z-score (z-score); and slope of the line between the 10% and the 95% percentiles based on the cumulative distribution function (CDF) curve were calculated, which accumulates the probabilities of the intensity of the observed pixels and reflects tumour heterogeneity. Additionally, additional percentiles from the ADC map, including 10tile to 90tile with an increment of 10 and 95tile were entered into the analysis.

5.2.6 Feature reduction and statistical analysis

The tumours were classified using both WHO 2007 and WHO 2016 classification schemes. Specifically, the tumours were grouped according to *grade (HGGs vs LGGs)* and *sub-grades (grade-II-IV)* following the 2007-WHO classification (Figure 5.2). According to the 2016-WHO classification, the tumours were broadly separated according to the *IDH-status (IDH-mut vs IDH-wt)*; *grouped* based on the molecular type *(GBM-wt, astro-mut, oligo-mut)*; and finally *sub-grouped* according to molecular type and histological grade *(GBM-wt, astro-mut-LG, astro-mut-HG, oligo-mut-LG, oligo-mut-HG)* (Figure 5.2). The extracted imaging features were tested for normality. Where the conditions of normality and variance homogeneity were not satisfied, the Kruskal-Wallis test was used to identify the features that demonstrated statistically significant differences ($P<0.05$) between the investigated gliomas classes. Since each of the comparisons was independent (i.e. there were no sub-group comparisons), besides the Kruskal-Wallis test was used for purpose of feature reduction, we felt that there was no need for Bonferroni correction.

The diagnostic performance was evaluated for these significant features for each of them separately as well as by grouping them in subsets. The features have been combined using backward elimination (BE) and pairwise-correlation (PC) methods. The diagnostic performance assessment was conducted for each MRI modality standalone. For the diagnostic performance evaluation, multinomial logistic regression analysis was undertaken to evaluate the model performance. Since the sample size is small, the leave-one-out-cross-validation (LOOCV) was used as a cross-validation method[205]. Cross-validation is a resampling procedure to evaluate the model against overfitting, which is the noise that negatively impacts the performance of the model and hence reduces its accuracy, for a limited sample size. This because the training data could fit the statistical model exactly (overfitting) and hence would not perform accurately against unseen data. LOOCV method uses all the samples, except one, to train the model. This process iterates many times equal to the sample size, to test one data point at each iteration. Thus, the output accuracy is the average of the resulted accuracy from the prediction models of each of the test points. The best-performed model and the optimal features' subsets were determined based on the results from the

LOOCV validation. LOOCV validation method overcomes the limitation of the small sample size that can lead to model overfitting and generate biased estimates for model performance. The model performance was reported using several indexes including sensitivity, specificity, F1 score, which is an accuracy index accounting for the sample size imbalance. In addition, accuracy (ACC) and Cohen's Kappa (K) were used as evaluation metrics for overall model performance. Accuracy is a ratio of the total right predicted class to the total prediction. Cohen's Kappa[206] measures how much better the model to predict classes compared to a model that simply predicts classes by chance (random guess). Thus, the validation model with the highest Kappa represents the best-performed model and hence the optimal features' subsets since high Kappa means the high agreement level between the tested models. Cohen's Kappa is always in the range of -1:+1, values less than or equal to zero indicates the model performance is useless, 0.01-0.20 as non to slight, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as substantial and 0.81-1 as perfect. For the statistical analyses, IBM SPSS version 24 (SPSS Inc, Chicago, IL) and R version 1.4.1717 (R Project for Statistical Computing; http://www.r-project.org) were used.

<table>
<tr><td colspan="10">32 adults with untreated gliomas</td></tr>
<tr><td colspan="4">WHO2007</td><td colspan="6">WHO2016</td></tr>
<tr><td colspan="4">1) Grading</td><td colspan="6">3) IDH-status</td></tr>
<tr><td colspan="2">LG</td><td colspan="2">HG</td><td colspan="5">IDH-mut</td><td>IDH-wt (HG)</td></tr>
<tr><td colspan="2">19</td><td colspan="2">12</td><td colspan="5">25</td><td>7</td></tr>
<tr><td colspan="4">2) Sub-grading</td><td colspan="6">4) Grouping</td></tr>
<tr><td>Grade-I</td><td>Grade-II</td><td>Grade-III</td><td>Grade-IV</td><td colspan="4">IDH-mut, ATRX loss (Astrocytoma)</td><td colspan="2">IDH-mut, 1p/19q codeleted (oligodendroglioma)</td><td>IDH-wt (HG)</td></tr>
<tr><td>-</td><td>19</td><td>6</td><td>6</td><td colspan="4">12</td><td colspan="2">13</td><td>7</td></tr>
<tr><td colspan="4"></td><td colspan="6">5) Sub-grouping</td></tr>
<tr><td colspan="4" rowspan="2">-</td><td>LG (II)</td><td>HG (III)</td><td>LG (II)</td><td>HG (III)</td><td>HG (IV)</td></tr>
<tr><td>10</td><td>2</td><td>9</td><td>4</td><td>7</td></tr>
<tr><td colspan="10">LG-low grade, HG-high grade gliomas</td></tr>
</table>

Figure 5.4 Histomolecular classification in the study population. Data from 32 adults with untreated gliomas were Prospectively recruited and underwent multimodal-MRI (PCASL, DSC, DCE and DWI). And classified according to 1) grading, 2) sub-grading, 3) IDH-status, 4) grouping and 5) sub-grouping.

5.3 Results

Fifty-eight adult patients with primary gliomas underwent multiparametric MRI, 26 of them were excluded from this study due to histopathological results showing gliosis (4 subjects), lack of histopathological sampling (11 patients), and incomplete MRI examination (11 patients). The histopathological results in the remaining 32 patients (average age of 43±13 (SD) years; M/F ratio 11/21) are shown in Figure 5.2.

5.3.1 Histogram features

The majority of the extracted histogram features across all advanced MRI modalities (ASL, DSC, DCE-mTK, DCE-L&L and DWI/ADC) revealed statistically significant differences among the defined classes of the gliomas. The numbers of the features (histogram-derived parameters) that showed statistically significant differences, based on the Kruskal-Wallis test, between grade, sub-grade, IDH-status, groups and sub-groups, respectively were: ASL features 10, 4, 8, 8, and 7; DSC features 10, 10, 14, 24, and 22; DCE-mTK features 3, 2, 1, 5, and 5 features; DCE-L&L features 6, 11, 15, and 12; and ADC features 12, 12, 26, 31, and 25, Illustrated in Figure 5.3 and Table 1-B in the Appendix B.

5.3.2 Diagnostic performance

ii. The significant features separately

Each of the significant features demonstrated various level of diagnostic accuracy. Table 5.2 and Figures 5.6 and 5.7 summarise the diagnostic performance of different MRI methods based on the best-performing models. Tables 2-B to 6-B in Appendix B show the diagnostic performance for each of the significant features. Figures 5.8 to 5.10 demonstrate two cases; one astrocytoma with mutant IDH status and ATRX loss in the right temporal lobe, and a wild type IDH glioblastoma localised in the left insula.

For gliomas *grading (HGGs vs. LGGs)* using the significant features separately, the diagnostic performance accuracy in order from the best to the least performing based on the LOOCV was DSC (rCBV-entropy, ACC=84%, K=67%), ASL (rTBF-max, ACC=75%, K=47%), ADC (rADC-80tile, ACC=73%, K=41%),

then both DCE-mTK (v_e-mode, ACC=72%, K=40%) and DCE-L&L (rv_p-mean, ACC=68%, K=29%).

For gliomas *sub-grading (II-IV)* using the significant features separately, the accuracy of the diagnostic performance was reduced and inadequate in predicting all the histological grades, see Table 5.2. DSC (rCBV-entropy, ACC=71%, K=42%) provided the highest accuracy, followed by the ADC (ADC-min, ACC=70%, K=38%), ASL (rTBF-mean, ACC=65%, K=25%), DCE-L&L (rv_p-max, ACC=66%, K=25%) and DCE-mTK (v_e-mode, ACC=56%, K=3%).

The result from the significant features standalone for the *IDH-status* prediction showed that both ASL (rTBF-mean) and DSC(rCBV-entropy) demonstrated similar performance (for both ACC=78% and K=24%), followed by the ADC (rADC-SD, ACC=80%, K=14%), DCE-L&L(rTC-min, ACC=78%, K=18%) and DCE-mTK (rk-mean, ACC=72%, K=0%).

For gliomas *grouping*, similar to gliomas sub-grading the diagnostic accuracy of the significant features separately was decreased and insufficient for prediction of every group. All of the used MRI methods expressed approximately similar performance, DCE-L&L (rFb-mean, ACC=61%, K=37%), DSC (rCBV-entropy, ACC=60%, K=37%), ADC (ADC-zscore, ACC=59%, K=37%), ASL (rTBF-skew, ACC=59%, K=34%) and DCE-mTK (rv_p-median, ACC=56%, K=28%).

Using the significant features standalone for *sub-grouping*, there was a further reduction in the models' accuracy to predict the gliomas sub-groups, as well as they were not able to predict all the included sub-groups. DSC exhibited the best performance (rCBV-entropy, ACC=56%, K=40%), followed by ASL (rTBF-median, ACC=50%, K=30%). All of ADC (ADC-SD), DCE-L&L (rFb-median) and DCE-mTK (rv_p-median) achieved performance of, respectively, ACC: 49%, 48%, 47% and K: 28%, 28%, 26%.

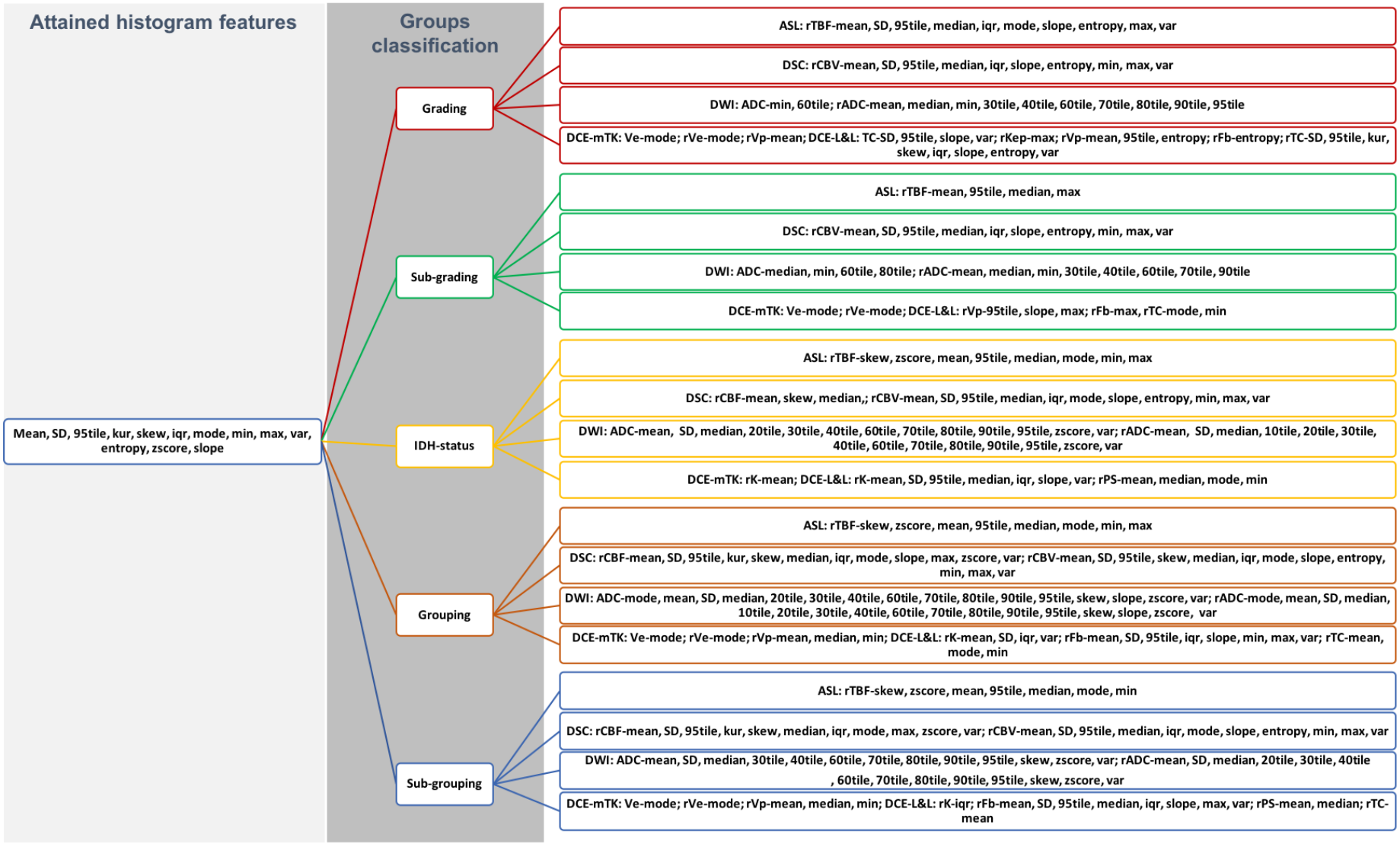
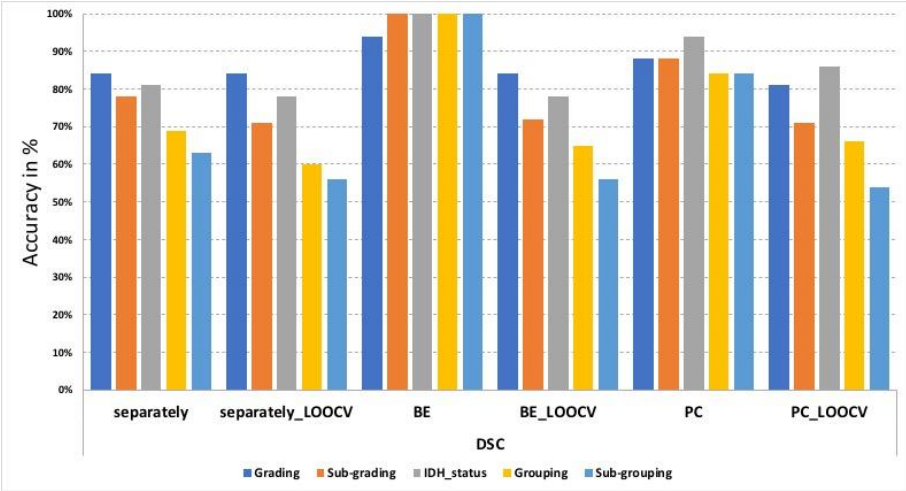
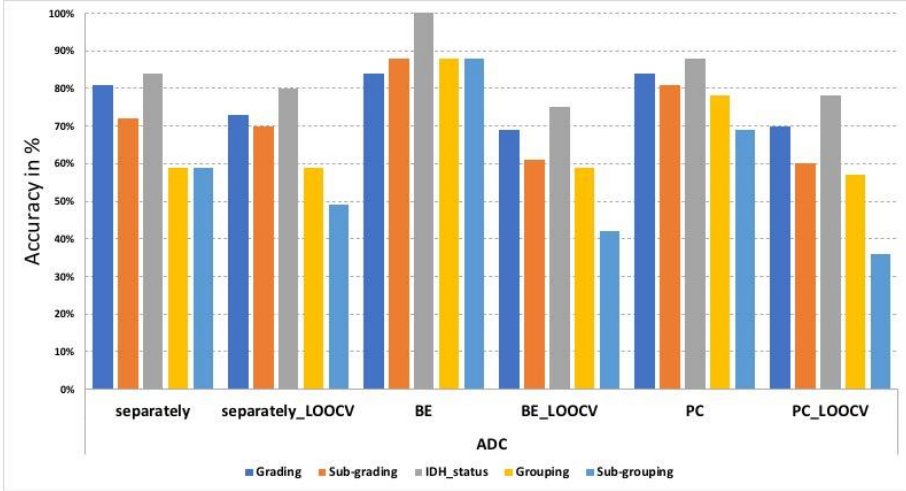


Figure 5.5 Hierarchy tree for the significant histogram features derived from Kruskal-Wallis test (P<0.05). Abbreviations: standard deviation (SD); 95th-percentile (95tile); kurtosis (kur); skewness (skew); inter-quartile range (iqr); minimum (min); maximum (max); variance (var); median z-score (zscore).

a.



b.



c.

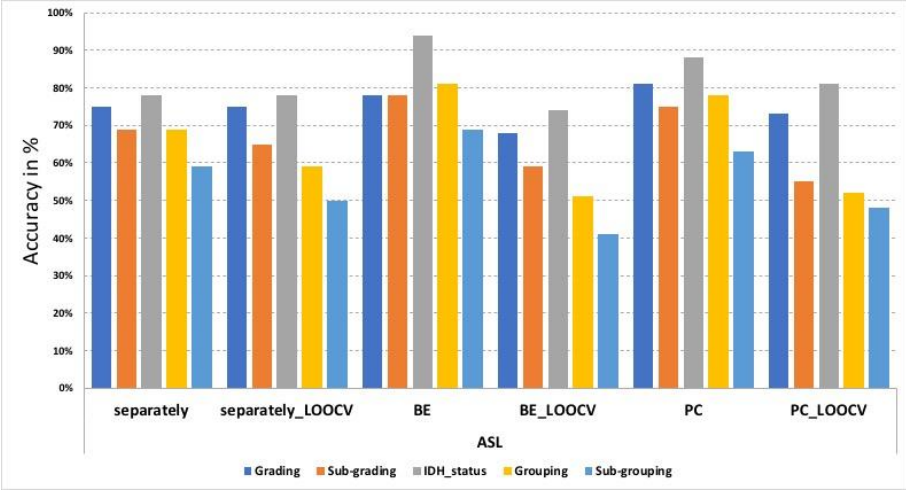
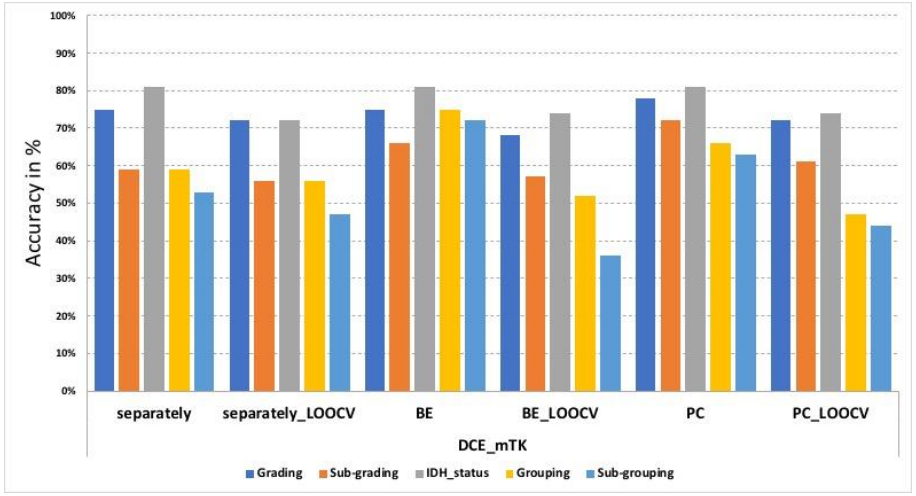


Figure 5.6 Bar charts of a) DSC, b) ADC and c) ASL, showing the diagnostic accuracy for different classification tasks using the significant histogram features from both the training models and the validation models using leave-one-out-cross-validation (LOOCV). The models based on the significant features separately and combined (backward elimination (BE) and pair-wise correlation (PC) combination methods).

a.



b.

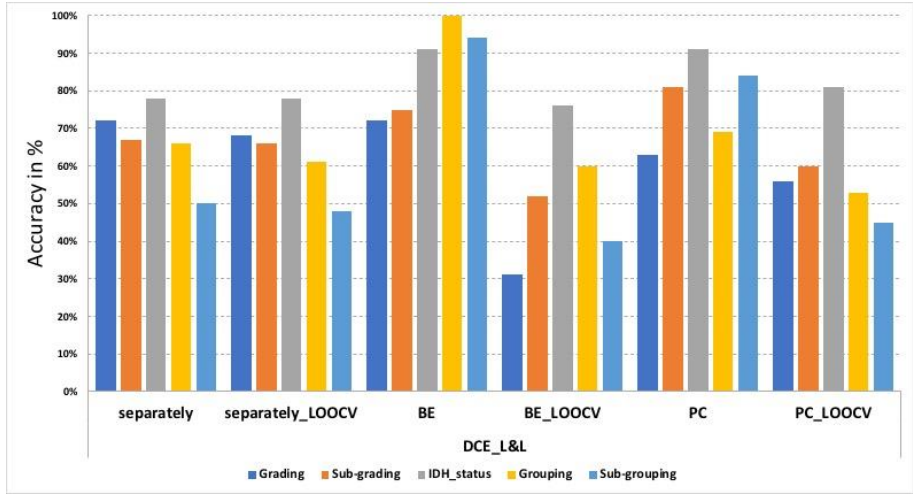


Figure 5.7 Bar charts of a) DCE_mTK and b) DCE_L&L, showing the diagnostic accuracy for different classification tasks using the significant histogram features from both the training models and the validation models using leave-one-out-cross-validation (LOOCV). The models based on the significant features separately and combined (backward elimination (BE) and pair-wise correlation (PC) combination methods).

iii. *The optimal features' subsets for different MRI methods*

The cross validation using LOOCV protects the model performance against the overfitting, consequently, we determined the optimal features sets as the ones that can diagnose the assigned gliomas' classes with high accuracy and generate similar accuracy among the resampled models from the LOOCV (high Kappa). Thus, the combined significant features in a validation model with the highest Kappa was determined the optimal features' subsets. Table 5.2 and Figures 5.6 and 5.7 shows the diagnostic performance of the optimal features' subsets from different modalities (the diagnostic performance of all the combined significant features are illustrated in Tables 7-B to 11-B in Appendix B). In contrast to using the significant features individually (univariate analysis), optimal features' subsets, for different MRI modalities, made it feasible to predict more classes. For example, as Table 5.2 illustrates for the sub-grading tasks from DSC, the individual significant feature was able to predict the grades-II, -III and -IV with sensitivity and specificity of 96% and 59%, 0% and 97%, 62% and 87%, respectively. In contrast, using the optimal features' subset achieved sensitivity and specificity of 88% and 69%, 28% and 91%, 67% and 89%, respectively.

Generally, the optimal features' subsets form the validated models demonstrated lower diagnostic performance compared to the training models. optimal features' subsets from the DSC demonstrated the best diagnostic performance among all of the involved tasks, followed by optimal features' subsets from ADC for grading, sub-grading. For grouping, the optimal features' subsets from DCE-L&L was the second superior while for IDH-status, whereas those from DCE-L&L and ASL showed the second-best diagnostic performance. For grading, the optimal features' subsets from DCE-mTK and ASL demonstrated the third-best diagnostic performance. From all of the MRI modalities, the models' overall diagnostic accuracy reduced as the involved classes increase particularly in sub-grouping (ACC <56%), where it involved 5 classes. The results from each of the MRI modalities are below.

The validated models of the optimal features' subsets from the DSC, using the PC and BE methods, showed moderate to substantial performance. Both PC and BE combination methods achieved similar diagnostic accuracy. The ACC from

PC and BE, respectively, for grading (81%, 84%), sub-grading (71%, 72%), IDH-status (86%, 78%), grouping (66%, 65%) and sub-grouping (54%, 56%).

The optimal features' subsets for ADC that derived from the validated models based on the PC and BE combination methods generated similar diagnostic accuracy and showed fair to moderate performance but with a bit lower diagnostic accuracy comparing to DSC. The ACC of the models using the optimal features from PC and BE, respectively, for grading (70%, 69%), sub-grading (60%, 61%), IDH-status (78%, 65%), grouping (57%, 59%) and sub-grouping (36%, 42%), respectively.

Similar to ADC, the validated models with optimal features from ASL showed fair to moderate performance. The ACC of the validation model from PC and BE respectively: for grading (73%, 68%), sub-grading (55%, 59%), IDH-status (81%, 74%), grouping (52%, 51%) and sub-grouping (48%, 41%).

In DCE-mTK, the validated models based on the optimal features' subsets using BE and PC methods provided approximately similar diagnostic accuracy to the individual features, still they facilitated the prediction of more classes (see table 5.2) as mentioned above. Similar to the individual features, the validated models based on the combined features were poorly performed for sub-grade and IDH-status form both PC (K: 6%, 19%) and BE (K: 2%, 2%), and fairly performed for grading, PC(K=29%) and BE(K=40%). The models diagnostic accuracy reduced for sub-grading and sub-grouping (ACC<55%).

In contrast to DCE-mTK, the optimal features' subsets from the DCE-L&L based on BE and PC combination methods, respectively, demonstrated fair (K=20%, ACC=76%) and moderate (K=45%, ACC=81%) diagnostic performance for IDH-status. Furthermore, for sub-grading the optimal features' subsets from the DCE-L&L based on PC combination method demonstrated fair performance with K=21% and ACC=60%. For grouping using the combined features from BE method the validation model showed fair performance with K=38% and ACC=60%. Similar to the DCE-mTK, the diagnostic accuracy for the validation models from DCE-L&L decreased for sub-grouping (ACC<50%). Also, the validation model from DCE-L&L poorly performed for grading.

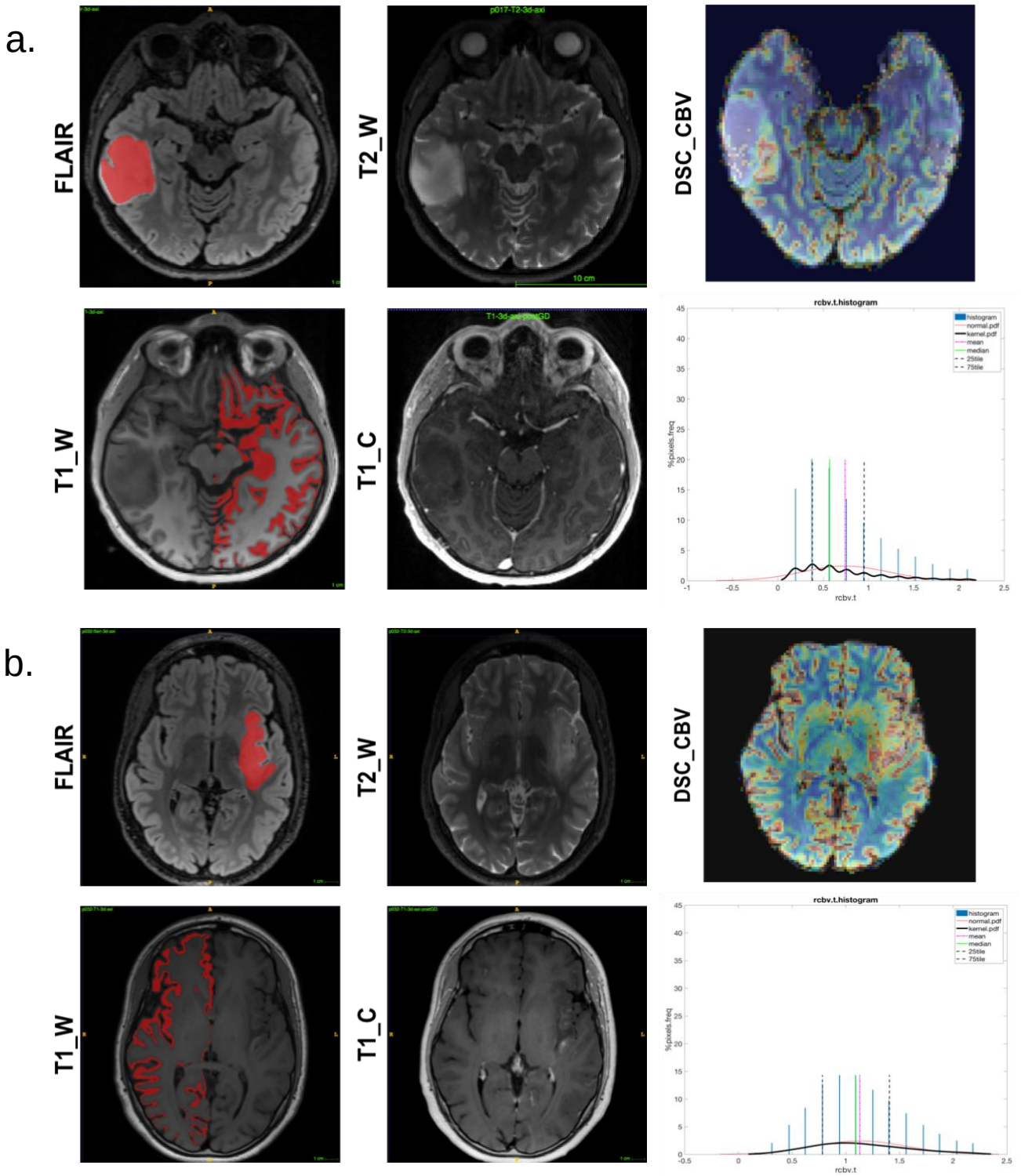


Figure 5.8 (a-b). Illustrative example of tumour segmentation and contralateral GM segmentation anatomical masks, rCBV maps and the corresponding normalised histogram from the whole tumour volume. **a**) 22-year-old male patient with low-grade astrocytoma with IDH mutation and ATRX loss in the right temporal lobe; and **b**) 30-year-old male patient with IDH-wt glioblastoma in the left insula. Compared to the astrocytoma, the glioblastoma demonstrated higher maximum rCBV value (as appears on the x-axis) with a wider range.

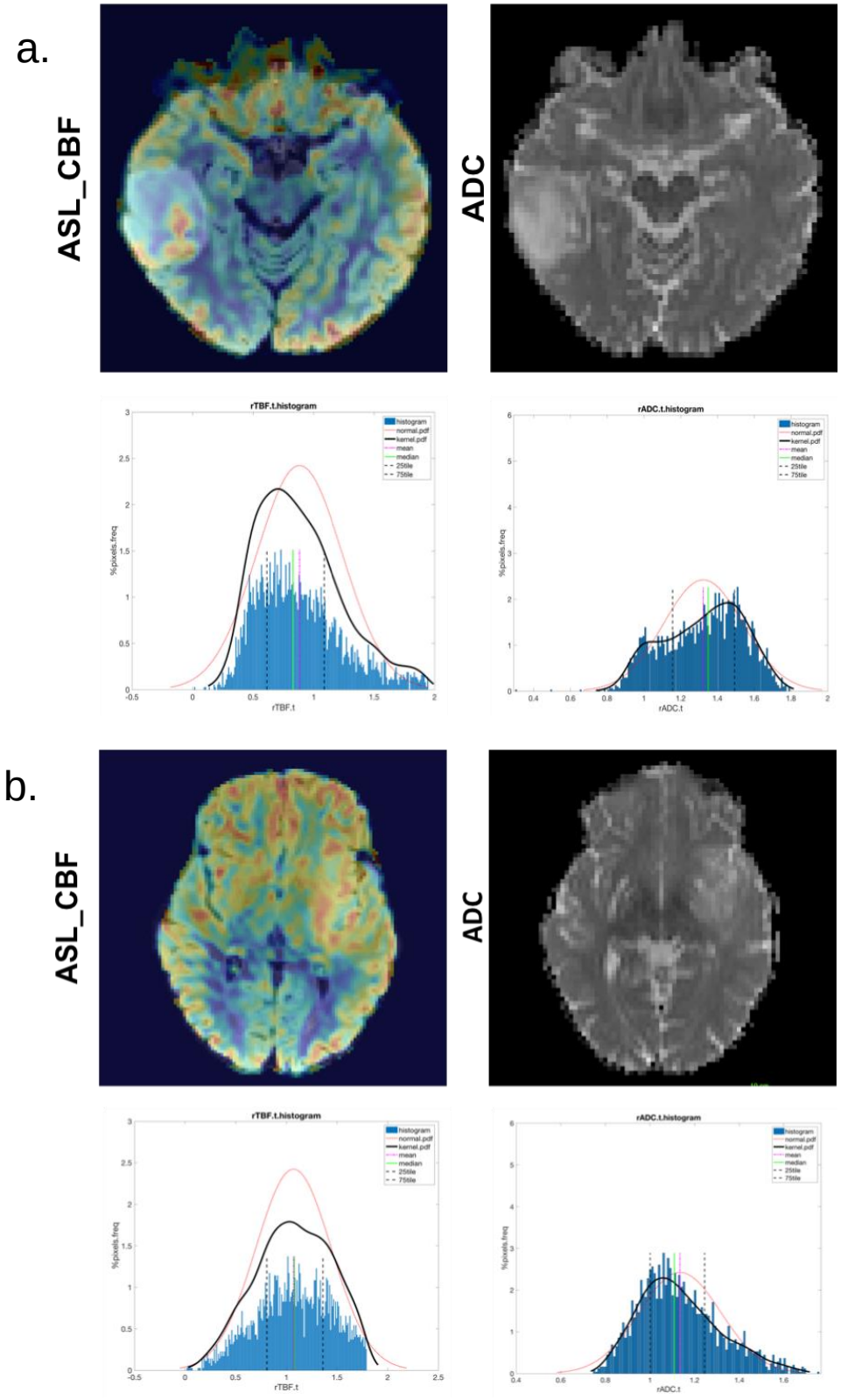


Figure 5.9 (a-b). ASL-CBF and ADC maps with corresponding normalised histograms from the whole tumour volume from the patients in figure 5. The ASL histogram in the low-grade astrocytoma (a) shows right-sided skewness whereas the ADC histogram demonstrates left-skewed peak, compared to the glioblastoma (b), where pattern is typically seen in gliomas transforming from lower to a higher grade.

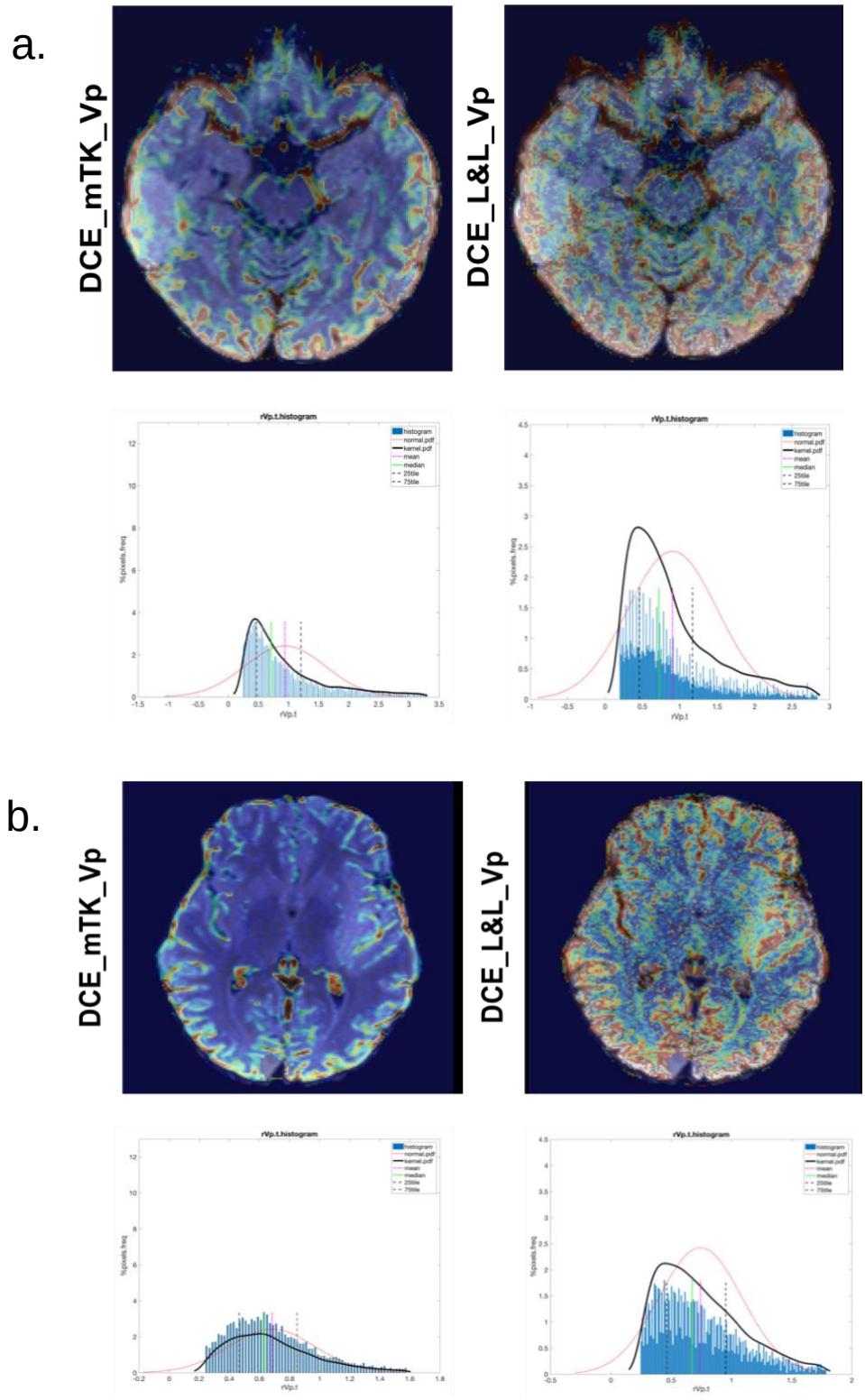


Figure 5.10 (a-b). DCE v_p maps generated by the modified Tofts and Kermode (mTK) model and the Lawrence and Lee (L&L) model with the corresponding normalised histograms from the whole tumour volume in the patients described in Figure 5. Compared to the low-grade astrocytoma (a) the intravascular volume values distribution from the glioblastoma (b) shows a broader and flatter peak. The skewness patterns seem to be similar across the two models in the LGG case while different in the HGG case, which is could be due to the hypervascularity of the HGG. This reflect the efficacy of the L&L model, unlike the simple mTK model, that account for the contrast agent gradient as a function of time and space hence reflect the underlying physiology more accurately in our case the hypervascularity in tumours, being a hallmark of diagnostic efficacy and grading in gliomas.

5.4 Discussion

This pilot study assessed the diagnostic role of advanced perfusion- and diffusion-MRI modalities (ASL, DSC, DCE, and DWI) in classifying gliomas by histological grade and molecular profile using radiomics extracted from the entire tumour volume. Also, it investigated the DCE-MRI diagnostic performance based on two different pharmacokinetic models (PKM), the mTK and the L&L models. We found that combining optimal features' subsets from standalone advanced MRI techniques enabled the model to predict more classes compared with using of single features (see Table 5.2). For the sub-grouping tasks, which included five classes with too small sample size in each, the evaluated models from all of the assessed MRI modalities displayed the lowest diagnostic accuracy. Notably among all the classification tasks the diagnostic performance of the models from DSC outweighed the rest of MRI modalities performance, followed by ADC except for IDH-status and grouping tasks. DCE-L&L was the second superior for grouping tasks while both ASL and DCE-L&L were the second superior for the IDH-status tasks. In addition, ASL and DCE-mTK were the third superiors for grading tasks.

Our work adds to the existing literature, which recently, and in the context of glioma staging, has been using different radiomics approaches[195,198–200,207]. Firstly, existing studies have reported the predictive accuracy of hand-crafted radiomics in a fragmented manner, either from standalone conventional MRI[197–199] or conventional MRI with DWI and ASL[195,207]. For example, Tian et al. reported better performance of GLCM than histogram features from multiparametric MRI (conventional, DWI and ASL) for glioma grading. They found that the diagnostic performance from the combined features proved more effective than each of the MRI methods alone (accuracy, sensitivity, specificity, area under the curve (AUC), 96.8%, 96.4%, 97.3%, 0.987, respectively)[195]; Su et al. [207] findings were in accordance with this, demonstrating that the use of multi-contrast MRI improved the AUC (0.911), sensitivity (85.2%) and specificity (85%) in comparison to using the CBF-ASL map alone or the contrast-enhanced T1-weighted images (AUC:0.750, 0.846, sensitivity: 71.4%, 75.9%, specificity: 63.9%, 77.9%, respectively). Lu et al. used multimodal conventional MRI to stratify gliomas according to the IDH status and 1p/19q status and reported the

model ability to the classification with estimated accuracy from the training dataset between 87.7% and 96.1% and comparable estimated accuracy from the validation dataset[208]. Although of promising results that have been reported there is no consensus on the throughput mining of quantitative image features (e.g., histogram-based textures, grey level co-occurrence matrices (GLCM) features, Gray-Level Run-Length Matrix (GLRLM) features), the mathematical formulas nor feature reduction and modelling, complicating the performance comparison prospectively. Secondly, the recently highlighted deep-learning (DL) approaches[63], can be utilised in a completely automated way to detect the lesion and extract high-level features which improve the prediction accuracy and even could be trained as an end-to-end process to the outcome prediction and decision making[209–211]. Two studies fulfilled using the same cohort to predict the IDH status in LGGs using T2 FLAIR[199,209], the first one carried out the classification based on machine learning and hand-crafted radiomics[199] while the other one based on DL trained as an end-to-end process[209]; The DL-radiomics model outperformed the machine learning-based radiomics (AUC: 92% vs 86%). However, DL based methods require large dataset and can incur high computational cost.

Our study adds to the existing literature by providing more comprehensive functional radiomics for gliomas classification. Radiomics was captured in the entire tumour for all currently available perfusion MR modalities standalone where ASL and ADC could be an alternative gadolinium-free approach, to gain a comprehensive overview of their performance. Pragmatically, the scanning time in clinical routine is limited, and it is imperative to design time- and diagnosis-efficient protocols with careful selection of the best performing sequences. Based on our results, DSC might be replaced by ASL if the latter is used beside ADC based on high b-values (>1000 s/mm^2). ASL has still not been fully adopted in routine MRI scanning bearing uncertainties for the optimal labelling technique and duration, acquisition method, post-processing, and results interpretation compared with the established DSC. However, the excellent diagnostic accuracy of ASL and ADC can be used in cases of gadolinium intolerance and contraindications, as well as in paediatric populations or for repeated scanning in the follow-up phase. This gadolinium-free approach poses a potential alternative to DSC diagnostic performance in the era of increasing concerns regarding

gadolinium deposition[163], with the caveats of longer scanning time required for ASL and limited dissemination of PCASL. Nonetheless, in this study, the short PCASL scanning time, similar to the DSC acquisition time (1min 26s), is an interesting approach that warrants more investigation and testing in future studies.

DCE has been widely explored in many clinical studies of pre-treatment characterisation of gliomas, being partially favoured over DSC due to its higher spatial resolution and its substantial immunity to susceptibility artefacts[212]. Also, it provides interpretable physiological parameters that reflect tumour tissue composition, such as EES volume and microvascular permeability, via fitting the acquired dynamic data into the PKM [80] by adapting the selected PKM to the temporal resolution used[22,83]. In our study, the DCE-MRI temporal resolution is indeed suboptimal for accurate AIF estimation, was acquired in ~ 5min with 4.2s temporal resolution, and we fitted the acquired data to the mTK[87] and the L&L[201] models. Usually, DCE data with low temporal resolution lead to biases in the parameters being fitted and would be the reason behind the low diagnostic performance (low kappa from the LOOCV), especially for DCE-mTK for sub-grading and IDH-status tasks and for DCE-L&L for grading tasks in our result. Temporal resolution is one of the important factors to maintain the accuracy and precession of the estimated parameters from the DCE model because depicting tissue perfusion requires rapid imaging[213]. The mTK model assumes negligible mean transit time (Tc) while L&L model implies its effect[80]. Thus, the former model could be analysed with a low temporal resolution whereas the latter requires the acquisition of the DCE data with a higher temporal resolution to capture the rapid changes in the arterial concentration and hence able to estimate Tc. The permeability from the mTK model was found to be not very sensitive to a reduction of the sampling frequency[214]. Accurate measurement of the AIF requires high temporal resolution to be able to describe the tracer kinetics during the vascular phase. A simulation study by Henderson et al.[215] showed that AIF sampling every second and tissue concentration sampled every 16s or less resulted in <10% error for parameters estimation except for vp, which has been reported as the most affected parameter by the insufficient temporal sampling of both AIF and tissue perfusion. For vp with an error of less than 10%, as reported

by Henderson et al., requires AIF sampling every second accompanied by a sampling of the tissue concentration at least every 4s. However, this sampling rate for the AIF is difficult to be achieved in real practice because it compromises the spatial resolution. Besides the influence of the low sampling rate of the AIF on the vp, from mTK model, vp estimation suffers from underestimation[88,216], which is mainly due to the assumption of negligible Tc[216]. So, it seems that the estimate of the vp, especially from the mTK model, will be severely biased. Also, in our study due to the low temporal resolution, the Tc from the L&L model would not be accurately measured, subsequently, the vp will not be accurate. Even though in our study, the vp estimates from both models were able to differentiate the gliomas grads, the vp-mean estimates from the L&L model demonstrated better diagnostic performance (ACC=68%, k=29%) than that from the mTK model (ACC=58%, k=9%), see Table 2-B in Appendix B. Also, a study by Kershaw and Cheng found that at similar temporal resolution (1.5s), apparently high, the adiabatic approximated tissue homogeneity (AATH) or L&L model (as we called in our study) was able to estimate the mTK parameters more accurate than mTK model itself [88]. In addition, their result showed that complex models with a large set of estimated parameters such as L&L require faster temporal resolution 1.5s to estimates the parameters with minimal error and can be relaxed to 6s if the accurate estimate of Tc is not needed (bias becomes >10%)[88] as it is the case in clinical tumour imaging. Thus, relaxing the temporal resolution increase the error and introduce further bias into the estimated parameters, and impact their diagnostic performance. In contrast, high temporal resolution is desirable for accurate parameters estimation, however, it is inevitability resulting in low spatial resolution and SNR which in turn lead to noisy and imprecise estimates of the parameters[217]. Another reason for the low diagnostic performance of the estimated parameters from DCE is the poor-quality AIF (Figure 5.2) which has a negative impact on their accuracy. Unfortunately, the automatic technique that has been used in our study generated a poor-quality AIF curve. An example of the typical AIF curve is shown in Figure 2 of the conducted study[218] by Rata et al. and among the literature[77,219]. The result of this study showed that the DCE-L&L optimal features' subsets performed better than those from DCE-mTK for sub-gradin, IDH-status and grouping tasks. Thus, using DCE-L&L is worth trialling if we consider its advantages for perfusion quantification, comprehensive capturing

of several pharmacokinetic parameters and no sensitivity to tumour or anatomy related susceptibility artifacts with the caveat of the time penalty.

In the clinical setup and in most of the previous studies, the discrimination ability of the measured values is gauged by the receiver operating curve (ROC) analysis of the tumour 'hot-spot' perfusion values. In light of the glioma heterogeneity, considering the distribution of the signal intensities in the entire tumour volume enables theoretically more accurate and effective diagnosis[60]. This is because segmentation of the entire tumour volume is objective and can be robustly performed even by less-experienced users[220]. ROC analysis, on the other hand, works as a binary classifier based on cut-off values. This makes it impractical, especially for MRI, due to its diversity in terms of imaging protocols, processing, and analyses, leading to threshold overlaps and inconsistencies across different reports in literature[42]. Instead of using one single cut-off value, joined multiple significant histogram features from the whole tumour (radiomics) may allow improved prediction rates, as confirmed by the current results. Similarly, previous studies showed that a combination of the most significant variables[193,194] or the determined optimal features' subsets[195,207] depending on the radiomics approach enhanced the diagnostic capacity.

A critical step in radiomics for diagnostic purposes is the concept of feature reduction[63]. The reduction was performed here using PC and BE[63,221]. Both resulted in enhancement of the model ability to predict more classes. BE method considers all the significant class-related features and ignores the feature-feature mutual information. In other words, the class prediction outcomes based on BE tend to be in favour of features with information that are related to specified classes (e.g. IDH-mut), having those classes gaining more information than other classes. In contrast, PC considers the relationship between both feature-class and feature-feature, therefore. Most of the recent studies[195,207] have used PC method (aka correlation redundancy) to reduce the attained features and avoid overfitting of the prediction model. It is worth mentioning that both PC and BE are supervised techniques, which are principally prone to overfitting.

Standardisation and validation are a big demand for the radiomics approaches to enable exploitation of their benefits and integration into clinical practice

(generalizability)[222,223]. This involves assessment of the Radiomics precision (repeatable and reproducible), which provide comparable data in any geographical location. In addition, model evaluation via internal and external validation using independent datasets with uniform quality and representative for the patient population. Model validation help reduces the danger of model overfitting that leading to overoptimistic model performance. Standardisation of the image acquisition, analysis and computational statistics of radiomics driven from quantitative images, like in our study, is a necessary step preceding validation[224]. Moreover, bias in the radiomics studies limits their generalizability[224], such as unbalanced training data, the manual volume of interest segmentation and locally developed method for feature extraction and reduction. Divorce and balanced sample would counteract falsely negative/positive outcomes. Automatic or semiautomatic segmentation that either conventional or deep learning (DL) based, would reduce the uncertainty of tumour delineation and hence the inter-observer variability[225–227].

Several limitations exist, the main limitation in the current study is the small and unbalanced sample size. This due to the Patient data is acquired prospectively besides the comprehensive study design limit the recruitment rate of this vulnerable patient category. However, our study wants to mine information in a rich but limited dataset in order to extract useful conclusions that will signpost a couple of meaningful combinations of advanced MRI modalities that can be validated in future larger studies with less time-consuming exam protocols. In addition, the images were acquired in a single centre using the same scanner, thereby limiting wider generalisation of the results. As such, our findings need to be validated and confirmed in multicentre large-scale MRI studies that apply post-processing techniques akin to these in this work. In addition, MR acquisitions with similar parameters or appropriate phantoms to enable harmonisation of quantitative MRI across sites and times would be a requisite to confirm the findings of our work and will allow broader dissemination and validation in the clinical routine environment [228]. We did our best to minimise bias related to post-processing and statistics, whereas the study included patients in a consecutive way to avoid selection bias.

5.5 Conclusion

Our results suggest that advanced MRI methods can provide higher diagnostic value by using the radiomics approach and combining multiple features than by relying on a single diagnostic feature. ASL and ADC represent a gadolinium-free approach poses a potential alternative to DSC. The radiomics from DCE-L&L yielded better diagnostic performance than those from DCE-mTK for the majority of the classification tasks, prompting for the use of more detailed physiological models for tumour perfusion imaging.

Table 5.2 Diagnostic performance summary of the significant tumour histogram features (from DSC, DCE, ADC and ASL) among different histomolecular groups

MRI method		Prediction performance	Grading		Sub-grading			IDH_status		Grouping			Sub-grouping				
			LGGs	HGGs	II	III	IV	wt	mut	GBM-wt	Astro-mut	Oligo-mut	GBM-wt	Astro-mut-LG	Astro-mut-HG	Oligo-mut-LG	Oligo-mut-HG
ASL		sensitivity	84%	62%	95%	0%	57%	43%	88%	43%	83%	69%	71%	90%	0%	56%	0%
Separately		specificity	62%	84%	46%	100%	88%	88%	43%	96%	80%	74%	84%	77%	100%	83%	100%
		F	80%	67%	81%	0%	57%	46%	86%	55%	77%	67%	63%	75%	0%	56%	0%
		Accuracy	75%		69%			78%		69%			59%				
		Kappa	47%		36%			33%		50%			44%				
	Internal validation LOOCV	sensitivity	84%	62%	93%	0%	43%	29%	92%	10%	83%	64%	38%	90%	0%	48%	0%
		specificity	62%	84%	33%	100%	89%	92%	29%	95%	78%	61%	87%	74%	100%	70%	100%
		F	80%	67%	78%	0%	47%	36%	87%	15%	76%	58%	41%	73%	0%	43%	0%
		Accuracy	75%		65%			78%		59%			50%				
		Kappa	47%		25%			24%		34%			30%				
Backwards-elimination		sensitivity	89%	62%	84%	67%	71%	71%	100%	71%	92%	77%	71%	90%	50%	56%	50%
		specificity	62%	89%	69%	96%	92%	100%	71%	96%	90%	84%	92%	82%	100%	87%	96%
		F	83%	70%	82%	73%	72%	83%	96%	77%	88%	77%	71%	78%	67%	59%	57%
		Accuracy	78%		78%			94%		81%			69%				

		Kappa	53%		60%			80%		71%			58%				
	Internal validation LOOCV	sensitivity	74%	59%	84%	11%	33%	48%	81%	48%	75%	31%	38%	83%	0%	22%	0%
		specificity	59%	74%	36%	95%	87%	81%	48%	79%	78%	68%	73%	80%	96%	80%	93%
		F	73%	60%	74%	17%	37%	44%	83%	43%	71%	35%	33%	74%	0%	26%	0%
		Accuracy	68%		59%			74%		51%			41%				
		Kappa	33%		18%			28%		26%			20%				
Pairwise-correlation		sensitivity	89%	69%	89%	50%	57%	71%	92%	57%	92%	77%	43%	90%	0%	67%	50%
		specificity	69%	89%	62%	96%	92%	92%	71%	96%	90%	79%	92%	81%	100%	83%	93%
		F	85%	75%	83%	60%	62%	71%	92%	67%	88%	74%	50%	78%	0%	63%	50%
		Accuracy	81%		75%			88%		78%			63%				
		Kappa	60%		53%			63%		66%			49%				
	Internal validation LOOCV	sensitivity	84%	56%	72%	22%	38%	48%	91%	0%	89%	46%	29%	90%	0%	48%	0%
		specificity	56%	84%	41%	92%	81%	91%	48%	87%	80%	58%	85%	74%	98%	75%	96%
		F	79%	63%	68%	29%	37%	53%	88%	0%	80%	44%	32%	73%	0%	46%	0%
		Accuracy	73%		55%			81%		52%			48%				
		Kappa	42%		16%			41%		24%			28%				
DSC		sensitivity	89%	77%	100%	0%	86%	43%	92%	57%	75%	69%	86%	70%	0%	78%	0%
Separately		specificity	77%	89%	69%	100%	88%	92%	43%	92%	90%	68%	80%	91%	100%	78%	100%

		F	87%	80%	90%	0%	75%	50%	88%	62%	78%	64%	67%	74%	0%	67%	0%
		Accuracy	84%		78%			81%		69%			63%				
		Kappa	67%		57%			39%		51%			49%				
	Internal validation LOOCV	sensitivity	89%	77%	96%	0%	62%	29%	92%	29%	75%	64%	81%	63%	0%	67%	0%
		specificity	77%	89%	59%	97%	87%	92%	29%	93%	85%	58%	79%	86%	100%	75%	100%
		F	87%	80%	86%	0%	59%	36%	87%	38%	75%	57%	63%	66%	0%	58%	0%
		Accuracy	84%		71%			78%		60%			56%				
		Kappa	67%		42%			24%		37%			40%				
Backwards-elimination		sensitivity	95%%	92%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
		specificity	92%%	95%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
		F	95%	92%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
		Accuracy	94%		100%			100%		100%			100%				
		Kappa	87%		100%			100%		100%			100%				
	Internal validation LOOCV	sensitivity	84%	85%	88%	28%	67%	62%	83%	81%	69%	51%	71%	63%	0%	63%	25%
		specificity	85%	84%	69%	91%	89%	83%	62%	92%	78%	74%	92%	89%	93%	73%	95%
		F	86%	81%	84%	33%	65%	55%	86%	77%	68%	54%	71%	68%	0%	54%	32%
		Accuracy	84%		72%			78%		65%			56%				
		Kappa	68%		48%			41%		46%			42%				

Pairwise-correlation		sensitivity	89%	85%	95%	67%	86%	86%	96%	86%	83%	85%	100%	80%	50%	89%	75%
		specificity	85%	89%	85%	92%	100%	96%	86%	96%	95%	84%	100%	95%	97%	87%	100%
		F	89%	85%	92%	67%	92%	86%	96%	86%	87%	81%	100%	84%	50%	80%	86%
		Accuracy	88%		88%			94%		84%			84%				
		Kappa	74%		77%			82%		76%			79%				
	Internal validation LOOCV	sensitivity	86%	74%	93%	22%	52%	67%	92%	57%	81%	56%	62%	73%	0%	56%	17%
		specificity	74%	86%	54%	95%	92%	92%	67%	92%	78%	75%	89%	88%	56%	78%	89%
		F	84%	76%	83%	31%	58%	68%	91%	62%	74%	59%	62%	73%	0%	53%	17%
		Accuracy	81%		71%			86%		66%			54%				
		Kappa	61%		42%			60%		47%			39%				
DCE_mTK		sensitivity	84%	62%	95%	0%	14%	14%	100%	0%	83%	69%	0%	90%	0%	67%	50%
Separately		specificity	62%	84%	38%	100%	80%	100%	14%	100%	65%	68%	100%	68%	100%	70%	96%
		F	80%	67%	80%	0%	15%	25%	89%	0%	69%	64%	0%	69%	0%	55%	57%
		Accuracy	75%		59%			81%		59%			53%				
		Kappa	47%		15%			21%		33%			35%				
	Internal validation LOOCV	sensitivity	84%	54%	95%	0%	0%	0%	92%	0%	83%	62%	0%	90%	0%	52%	33%
		specificity	54%	84%	21%	96%	89%	92%	0%	100%	60%	68%	100%	62%	100%	67%	96%
		F	78%	61%	76%	0%	0%	0%	84%	0%	67%	59%	0%	66%	0%	44%	42%

		Accuracy	72%		56%			72%		56%			47%					
		Kappa	40%		3%			0%		28%			26%					
Backwards-elimination		sensitivity	84%	62%	95%	33%	14%	43%	92%	43%	92%	77%	57%	80%	0%	89%	75%
		specificity	62%	84%	38%	96%	92%	92%	43%	96%	80%	84%	88%	82%	100%	91%	100%
		F	80%	67%	80%	44%	20%	50%	88%	55%	81%	77%	57%	73%	0%	84%	86%
		Accuracy	75%		66%			81%		75%			72%					
		Kappa	47%		28%			39%		60%			62%					
	Internal validation LOOCV	sensitivity	84%	44%	91%	17%	0%	10%	92%	0%	83%	51%	10%	77%	0%	19%	42%
		specificity	44%	84%	18%	91%	97%	92%	10%	92%	63%	68%	84%	70%	100%	71%	89%
		F	76%	52%	74%	21%	0%	14%	85%	0%	68%	52%	11%	63%	0%	19%	38%
		Accuracy	68%		57%			74%		52%			36%					
		Kappa	29%		6%			2%		23%			14%					
Pairwise-correlation		sensitivity	89%	62%	89%	33%	57%	43%	92%	43%	83%	62%	29%	90%	0%	68%	75%
		specificity	62%	89%	54%	100%	88%	92%	43%	88%	75%	84%	92%	77%	100%	78%	100%
		F	83%	70%	81%	50%	57%	50%	88%	46%	74%	67%	36%	75%	0%	60%	86%
		Accuracy	78%		72%			81%		66%			63%					
		Kappa	53%		24%			39%		47%			48%					
	Internal	sensitivity	84%	54%	91%	17%	19%	10%	92%	19%	64%	46%	10%	83%	0%	48%	17%

		specificity	54%	84%	33%	96%	89%	92%	10%	89%	63%	63%	88%	73%	100%	65%	96%
		F	78%	61%	77%	25%	24%	14%	85%	24%	57%	46%	13%	68%	0%	41%	24%
		Accuracy	72%		61%			74%		47%			44%				
		Kappa	40%		19%			2%		16%			22%				
DCE_L&L		sensitivity	89%	46%	95%	67%	0%	29%	92%	0%	92%	77%	0%	80%	0%	56%	75%
Separately		specificity	46%	89%	31%	96%	100%	92%	29%	100%	60%	84%	100%	59%	100%	74%	96%
		F	79%	57%	78%	73%	0%	36%	87%	0%	71%	77%	0%	59%	0%	50%	75%
		Accuracy	72%		67%			78%		66%			50%				
		Kappa	38%		33%			24%		44%			31%				
	Internal validation LOOCV	sensitivity	86%	41%	95%	50%	0%	19%	95%	0%	81%	78%	0%	80%	0%	56%	58%
		specificity	41%	86%	23%	96%	100%	95%	19%	100%	60%	77%	100%	59%	100%	71%	96%
		F	76%	51%	77%	60%	0%	28%	87%	0%	65%	73%	0%	59%	0%	48%	64%
		Accuracy	68%		66%			78%		61%			48%				
		Kappa	29%		25%			18%		37%			28%				
Backwards-elimination		sensitivity	79%	62%	84%	67%	57%	71%	96%	100%	100%	100%	86%	100%	100%	89%	100%
		specificity	62%	79%	62%	96%	92%	96%	71%	100%	100%	100%	100%	91%	100%	100%	100%
		F	77%	64%	80%	72%	62%	77%	94%	100%	100%	100%	92%	91%	100%	94%	100%
		Accuracy	72%%		75%			91%		100%			94%				

		Kappa	41%		54%			71%		100%			92%					
	Internal validation LOOCV	sensitivity	46%	10%	72%	39%	10%	29%	89%	33%	64%	72%	19%	53%	0%	41%	58%	
		specificity	10%	46%	33%	91%	83%	89%	29%	87%	77%	75%	75%	74%	100%	72%	96%	
		F	44	11%	66%	44%	11%	34%	85%	37%	63%	69%	18%	51%	0%	39%	64%	
		Accuracy	31%		52%			76%		60%			40%					
		Kappa	45%		9%			20%		38%			18%					
Pairwise-correlation		sensitivity	79%	38%	95%	67%	57%	57%	100%	57%	75%	69%	86%	70%	100%	89%	100%	
		specificity	38%	79%	77%	96%	92%	100%	57%	92%	70%	89%	96%	91%	97%	96%	100%	
		F	71%	45%	90%	72%	62%	73%	94%	62%	67%	75%	86%	74%	80%	89%	100%	
		Accuracy	63%		81%			91%		69%			84%					
		Kappa	18%		65%			68%		51%			80%					
	Internal validation LOOCV	sensitivity	70%	36%	88%	28%	14%	57%	88%	33%	53%	64%	48%	43%	0%	52%	50%	
		specificity	36%	70%	41%	94%	87%	88%	57%	83%	67%	79%	93%	65%	91%	75%	100%	
		F	66%	40%	77%	36%	18%	57%	88%	34%	51%	66%	56%	39%	0%	48%	67%	
		Accuracy	56%		60%			81%		53%			45%					
		Kappa	6%		21%			45%		27%			26%					
ADC		sensitivity	84%	77%	100%	67%	0%	29%	100%	0%	83%	69%	71%	80%	0%	67%	0%	
Separately		specificity	77%	84%	46%	92%	100%	100%	29%	96%	80%	58%	88%	73%	100%	83%	100%	

		F	84%	77%	84%	67%	0%	44%	91%	0%	77%	60%	67%	67%	0%	63%	0%
		Accuracy	81%		72%			84%		59%			59%				
		Kappa	61%		42%			38%		34%			44%				
	Internal validation LOOCV	sensitivity	88%	51%	96%	67%	0%	10%	100%	0%	83%	69%	29%	83%	0%	59%	0%
		specificity	51%	88%	44%	91%	100%	100%	10%	96%	80%	58%	89%	68%	100%	71%	100%
		F	79%	61%	82%	65%	0%	17%	89%	0%	77%	60%	34%	66%	0%	51%	0%
		Accuracy	73%		70%			80%		59%			49%				
		Kappa	41%		38%			14%		34%			28%				
Backwards-elimination		sensitivity	89%	77%	100%	50%	86%	100%	100%	86%	92%	85%	86%	100%	50%	89%	75%
		specificity	77%	89%	85%	96%	96%	100%	100%	96%	95%	89%	96%	95%	100%	96%	96%
		F	87%	80%	95%	60%	86%	100%	100%	86%	92%	85%	86%	95%	67%	89%	75%
		Accuracy	84%		88%			100%		88%			88%				
		Kappa	67%		77%			100%		81%			83%				
	Internal validation LOOCV	sensitivity	77%	56%	82%	39%	19%	48%	83%	38%	78%	54%	29%	73%	0%	44%	0%
		specificity	56%	77%	54%	88%	85%	83%	48%	87%	80%	70%	84%	77%	96%	80%	87%
		F	75%	59%	77%	41%	22%	45%	84%	41%	74%	55%	31%	66%	0%	45%	0%
		Accuracy	69%		61%			75%		59%			42%				
		Kappa	34%		26%			29%		37%			23%				

Pairwise-correlation		sensitivity	89%	77%	89%	67%	71%	57%	96%	57%	92%	77%	71%	70%	100%	56%	75%
		specificity	77%	89%	69%	96%	96%	96%	57%	96%	90%	79%	92%	91%	100%	78%	96%
		F	87%	80%	85%	72%	77%	67%	92%	67%	88%	74%	71%	74%	100%	53%	75%
		Accuracy	84%		81%			88%		78%			69%				
		Kappa	67%		65%			59%		66%			59%				
	Internal validation LOOCV	sensitivity	79%	56%	81%	33%	29%	29%	92%	24%	69%	64%	26%	57%	0%	33%	25
		specificity	56%	79%	51%	88%	87%	92%	29%	88%	83%	61%	80%	76%	96%	77%	88%
		F	76%	60%	75%	36%	32%	36%	87%	29%	70%	58%	29%	54%	0%	35%	24%
		Accuracy	70%		60%			78%		57%			36%				
		Kappa	46%		26%			24%		32%			15%				

Chapter 6

6 Conclusion and further work 185

6.1 Summary 185

6.2 Recommendation for further work 186

6.2.1 Image acquisition and processing 187

6.2.2 Further quantitative estimates from ASL 187

6.2.3 Sample size and validation 189

6.3 Role of ASL in other brain tumours and surveillance 189

6.3.1 ASL in other brain tumours 189

6.3.2 ASL in brain tumours surveillance 190

6.4 ASL in paediatric brain tumours 191

6 Conclusion and further work

6.1 Summary

This thesis aims to assess the diagnostic value of the arterial spin labelling (ASL)-based perfusion estimates in untreated adult gliomas. The first chapter provided an introduction into the magnetic resonance imaging (MRI) basic principles as well as the advanced MRI for primary gliomas diagnosis. The second chapter provided an overview of the methodological background of the three most common MRI perfusion methods (DSC, DCE and ASL), and the shortcomings that could bias the measurements. In addition, it focused on the ASL technique and explored the available suggestions to mitigate these harmful effects.

The third chapter provided high-level evidence through systematic review and meta-analysis for the diagnostic value of the relative tumour blood flow (rTBF) from ASL. Especially the maximum-rTBF helps in both differentiation between the HGGs and the LGGs as well as in gliomas sub-grading. The estimated effect size for the rTBF was approximately similar between HGGs and LGGs and grade II and grade III gliomas, (-1.46 (-2.00, -0.91), p-value<0.001), (-1.39 (-1.89, -0.89), p-value<0.001), respectively. The effect size between grade III and grade IV tumours was smaller (-1.05 (-1.82, -0.27)), p<0.05).

The fourth chapter elucidated the role of the normalised perfusion estimates (rTBF), compared to the absolute TBF (aTBF). The rTBF expressed lower -WCoV and CR, which quantified the variation (error) in the same unit as the measured value (ml/100 g/min in this case). This result supports the better performing diagnostic role of rTBF compared to aTBF. Specifically, the rTBF that was normalised to the segmented whole GM from the CNAH (CR with 95% CI was: 0.07 (0.023, 0.096), 0.156 (0.1025, 0.195) for the PASL and PCASL, respectively). This means that if the effect size is lower than the CR, it is deemed to be due to physiological, scanner-related variation. However, when estimating rTBF, large ROIs need to be used to filter out systemic misleading within-subject variation and thus provide more robust perfusion measurements. Finally, no large difference between repeatabilities were observed between PASL and PCASL, which might have been attributed primarily to the difference in acquisition parameters, and both methods can therefore be recommended to assess brain

perfusion, with the caveat that PCASL allows for longer PLD than PASL, which therefore minimises the macro-vascular artefacts, and provide better estimates of patient perfusion.

The fifth chapter demonstrated the remarkable diagnostic performance of the radiomics approach in comparison to the simple univariate summary results. This was based on different advanced MRI methods, including DSC, DCE, ASL, and diffusion-weighted imaging (DWI) to classify primary gliomas. Even though, the diagnostic performances of the ASL, DSC and DCE for the attained features separately were comparable to combining optimal features' subsets, where the latter made it feasible to predict more classes. The optimal features' subsets from DSC showed superior diagnostic performance, with ACC for grading 84%, sub-grading 72%, IDH-status 86%, grouping 66% and sub-grouping 56%. Followed by ADC for grading (ACC=70%) and sub-grading (ACC=61%), DCE-L&L for grouping (ACC=60%). Both DCE-L&L and ASL showed the second-best diagnostic performance for IDH-status (ACC=81%). DCE-mTK surpass DCE-L&L for grading (ACC: 56% vs 72%). ASL and ADC represent a gadolinium-free approach poses a potential alternative to DSC. in selected cases and in the era of increasing awareness for gadolinium deposition[163]. The practicality of the gadolinium-free approach is obvious in cases of gadolinium intolerance and contraindication as well as in the pediatric population or repeated scanning. The radiomics from the DCE based on the St Lawrence and Lee model revealed better diagnostic performance than those based on the modified Tofts-Kermode model, which highlights the importance of exploring more advanced pharmacokinetic models for DCE, which likely more accurately reflects the underlying physiology of glioma .

6.2 Recommendation for further work

The subsidiary role of the ASL in adult gliomas classification is highlighted in this thesis. Nevertheless, the practical and technical challenges delay its adoption into daily clinical practice. Besides that, these challenges have an impact on the estimations of numerically meaningful data in radiomics That is worth considering in future studies.

6.2.1 Image acquisition and processing

Variations in image acquisition and processing introduce undesirable changes that obscure the actual biological changes. This is most significant in cases where the assessment is quantitative . The stability of these quantitative estimations is generally evaluated using either test-retest (repeatability and reproducibility) and inter-observer reliability measurements, both of which can determine the source of the variance in the measurements and visualise it in error bars besides the reported number. In this aspect, the professional societies (e.g. QIBA) are making efforts to improve the effectiveness of quantitative imaging, by providing consensus on the measurement accuracy and technical guidelines[100,229].

Delineation of tumour volume is a step in the image analysis that could introduce variation and should be taken care of it. The precision of tumour segmentation is crucial for the reproducibility of the computed quantitative values and radiomics features. It has been shown that semiautomatic segmentation reduces the uncertainty of tumour delineation and hence the inter-observer variability[225,226]. Even if the impact of the manual segmentation method on the radiomics quantification is deemed to be small, (semi)automatic segmentation would be a useful step towards process robustness.

6.2.2 Further quantitative estimates from ASL

Arterial arrival time (AAT) and arterial blood volume (aBV) can be estimated from ASL if using multiple delay times (TIs) and the QUASAR (quantitative STAR labelling of arterial regions) sequence, respectively. Westen et al. compared the ASL-aBV to the DSC-CBV in 10 patients with brain tumours[160] and found that CBV and aBV increased from low grade to higher-grade tumours, which suggests that aBV could be a promising tool for tumour grading. A study in paediatric brain tumours estimated the AAT with and without a crusher gradient and argued that these estimates provided capillary and pre-capillary arrival times, respectively. This is because the usage of crusher gradients in ASL removes the signal of the large vessels. Also, blood arrival time from the large vessels to the capillary level could be estimated and utilised as an indirect index for the permeability. It can be postulated that the delay at the arrival time between the proximal location to the tumour's large vessels, and the capillary level is related to the dispersion of the

magnetically labelled blood due to the increased permeability in tumours with neo-angiogenesis. This non-invasive, Gd-free biomarker of prolonged arrival time might be relevant to characterising the grade of glioma and may correlate with permeability estimates derived from DCE-MRI. Thus, the suggestion is to obtain ASL images in tumours over a range of delay times without using crusher gradients. The acquired image signal will reflect the perfusion from the large and small vessels as well as the arrival time of the labelled blood mainly at the large vessels. Subsequently, another identical set of ASL images using crusher gradients, as earlier published in the QUASAR study[126,170], is expected that will show perfusion and ATT only at the capillary level. The difference between the arrival time at the levels of large vessels and capillaries (without and with crusher gradient, respectively) will reflect the degree of the blood dispersion due to the increased vessel permeability; hence, the coined biomarker name blood arrival time might be indirectly related to the tumour permeability. If the proposed estimate of blood arrival time is highly correlated to the established permeability index derived from the DCE-MRI or alternatively from other existing ASL-related permeability indices to water[230], this would offer a new surrogate biomarker for staging and monitoring gliomas that utilises the comparative advantages of ASL over the Gd-based perfusion MRI.

Unlike ASL image acquisitions at multiple delay times, ASL acquisition at single TI does not provide enough information regarding the AAT. Nevertheless, other methods might be useful to overcome this limitation; in example, Mutsaerts et al. calculated the spatial COV as the standard deviation (SD) between voxel signals inside the ROI divided by their mean, and interpreted it as a predictor for the ATT at single PLD[231]. It would be interesting to explore this parameter in various brain diseases, particularly, gliomas, which are the focus of this thesis.

ASL estimates are attractive biomarkers due to their Gd-free acquisition nature. Besides its capability to quantify absolute CBF may provide a more accurate estimation of the tumour microvascular bed. Each ASL-derived parameter might highlight different tumour pathophysiology and using them all together could enhance the diagnostic accuracy. For instance, using the rTBF and bolus arrival time (BAT) raised the diagnostic accuracy to 72% [129]. Moreover, the combination

of mTIs-ASL (nCBF & nBAT) and conventional MRI provided better diagnostic accuracy than using the conventional MRI alone, 81% and 56%, respectively[129].

6.2.3 Sample size and validation

A common problem among the most neuroimaging studies is that the sample sizes are usually limited and from a single or a few centres. The small sample size leads to noisy measurements; hence the results reported by the various studies differ widely and lack consensus. Even though quantitative imaging biomarkers are constantly validated against the clinical outcome (i.e., disease course or molecular/genetic biomarkers), they are rarely validated across different platforms/institutions. Multicentre validation is important, as it provides evidence of reproducibility of the results across different centres and their generalisability.

6.3 Role of ASL in other brain tumours and surveillance

6.3.1 ASL in other brain tumours

Perfusion estimates from ASL besides their benefits in gliomas diagnosis have also been used in brain tumour differentiation, treatment surveillance and prognosis. Weber et al. discriminated between GBM, metastasis and primary CNS-lymphoma (PCNSL) using ASL with better discrimination ability and predictive value than MRS[154]. The rTBF in the tumour hot spot was higher for the GBM than that for the PCNSL and higher in the peritumoral region for the GBM than for the metastases[154]. This result is plausible since the peritumoral regions in metastases are consisted of vasogenic oedema, in contrast to gliomas, which infiltrate the peritumoral areas[232,233].

Yamashita et al. assessed the role of ASL in differentiating hemangioblastoma from metastatic brain tumours using both rTBF and aTBF[234]. Both estimated metrics expressed higher values in hemangioblastomas than in metastases as hemangioblastomas arise from blood vessels and are highly vascularised[235]. However, the metastatic renal cell carcinoma (RCC) expressed high perfusion values, which appeared to mimic the hemangioblastomas[234]. This phenomenon may be explained by the fact that in metastatic brain tumours perfusion values depend on the primary tumour's vascularisation. Primary RCC are known for their

increased vascularity[236]. Weber et al. also reported metastases from RCCs displaying high perfusion values[166]. A recent study that investigated the ability of ASL to differentiate between the hemangioblastomas and metastases that only were in the posterior fossa[237]concluded that if both ASL and T1W-contrasted showed increased values, the size ratio (defined as ROI of the tumour on ASL in comparison to that on the enhanced-T1W) larger than 0.98 is more likely to be hemangioblastoma. Yoo et al. used crushed PCASL not only to discriminate HGGs from PCNSL but also to investigate the connection between HGGs' perfusion values and genetic biomarkers[165]. This study found both aTBF and rTBF to be significantly higher in HGGs than PCNSL, which agrees with previous studies[154,238]. Among the six investigated genetic biomarkers , only the epidermal growth factor receptor (EGFR) expressed significant positive correlation with rTBF and aTBF. EGFR was also positively correlated with poor prognosis in anaplastic oligodendroglioma[239].

6.3.2 ASL in brain tumours surveillance

ASL for brain tumour follow-up is still under evaluation. Some studies have been investigating radiation impact on normal brain tissue[240,241] while other works have evaluated the treatment response [166,167]. Weber et al. evaluated the perfusion of the contralateral normal brain tissue in 62 metastatic brain tumour patients prior to single dose of stereotactic radiosurgery (median of 18 Gy, 12 to 20 Gy) and up to 5 follow-up sessions after treatment. Perfusion values obtained from PASL and DSC remain unchanged on surveillance [241], which may be due to the low administered dose less than 0.5 Gy. A recent study evaluated perfusion in healthy brain tissue using PCASL in 24 GBM patients after radio-chemotherapy (RCT). This study was searching for any relationship between the received dose and tissue perfusion alteration[242]. It reported a significant reduction in rCBF between the pre-treatment session and the first follow-up . This reduction was more pronounced in the regions that received doses > 50 Gy than in regions receiving doses <10Gy.

Weber et al. investigated the value of PASL in the prediction of treatment outcomes by estimating rTBF in 25 metastatic brain tumour patients before stereotactic radiosurgery and during early follow-up[166]. The reduction of rTBF in

the enhancing tumour tissue between baseline and six weeks after treatment showed higher predictive value than tumour volume changes . In addition, the rCBF in the contralateral normal brain tissue, which received less than 0.5 Gy, was unchanged, a finding consistent with the previous studies[241,243]. A recent study explored the usefulness of PCASL in terms of the evaluation of bevacizumab anti-angiogenic effect in nude rats with GBM[167], and this found a significant reduction in rTBF after the treatment. These findings suggest that ASL is a suitable method for evaluation of brain tumour treatment.

Other recent studies have also assessed ASL prognostic value. Rau et al. used crushed PASL as compared to DSC in 69 patients with high-grade treatment-naïve gliomas . Cases with rTBF from the ASL and rCBF and rCBV from DSC below the ROC curve analysis-derived cut-off thresholds demonstrated longer progression-free-survival . However, this result reached statistical significance only for the rCBV[244]. Another study included 24 pre-treated astrocytoma with different WHO grades and similarly investigated the prognostic value of the crushed PASL [245]. Regardless of tumour grade, the authors subdivided the lesions in high and low perfusion tumours using aTBF. An optimal cut-off value of 182 ml/100 g/ min was set and patients with high perfusion values had shorter periods of event-free survival than patients with lower a TBF values . However, further studies in larger populations are needed to establish the prognostic role of ASL.

6.4 ASL in pediatric brain tumours

Although several studies have examined tumour perfusion in adults using ASL, scarce data exists for children. It may not be straightforward to apply adult estimated perfusion parameters to the paediatric population because CBF accuracy is subject to scaling imaging parameters, which include subject-specific parameters[246]. For example, T1 relaxation time is longer in children than in adults[246]. Particularly in terms of tumour evaluation, paediatric tumours have several unique features, such as their predominance in the posterior fossa and diverse histological presentations. The use of ASL in children has the advantage of reduced distortion artefacts in the frontal and inferior brain regions due to the

immature skull, no use of gadolinium, and increased repeat rate in case of unsuccessful sedation or motion artefacts.

Yeom et al. evaluated the role of ASL in paediatric brain tumours in a similar manner to the results reported in adults; the rTBFmax found to be significantly higher in HGGs than in LGGs[247]. Among the posterior fossa tumours, rTBF was significantly higher in the medulloblastoma than in the pilocytic astrocytoma. A study demonstrated that ASL could discriminate pilomyxoid astrocytoma from pilocytic counterparts [156]. The role of ASL in paediatric brain tumours was also demonstrated in a case study that emphasised its value in distinguishing hemangioblastoma from other medullary low-grade tumours that displayed similar conventional MRI findings[248]. Finally, a recent meta-analysis synthesised the evidence for the high diagnostic accuracy of ASL perfusion estimates in the differentiation of the HGGs from LGGs in the paediatric population[249].

In conclusion, when considered against other available MRI perfusion methods, the ASL main strengths consist of its capability to estimate absolute CBF, and its minimally invasive nature with no need for exogenous contrast medium. Although these advantages are quite remarkable, the modality has yet to be validated before being incorporated into routine clinical practice. Owing to the continuous development over the last decade, ASL has reached a stage that encourages its trialling and adoption in clinical institutions. With these issues in mind, ASL appears a useful modality for glioma patients, especially the vulnerable and paediatric subjects, and post-treatment tumour patients with a low tolerance for high-rate contrast injections and complex venous access.

7 References

1. Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol*. 2007;114(2):97-109. doi:10.1007/s00401-007-0243-4.

2. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol*. 2016;131(6):803-820. doi:10.1007/s00401-016-1545-1.

3. Brandner S, Deimling A Von. Invited review : Diagnostic , prognostic and predictive relevance of molecular markers in gliomas. 2015;41(6):694-720. doi:10.1111/nan.12246.

4. Yan H, Parsons DW, Jin G, et al. IDH1 and IDH2 Mutations in Gliomas. *N Engl J Med*. 2009;360(8):765-773. doi:10.1056/NEJMoa0808710.

5. Wick W, Hartmann C, Engel C, et al. NOA-04 randomized phase III trial of sequential radiochemotherapy of anaplastic glioma with Procarbazine, Lomustine, and Vincristine or Temozolomide. *J Clin Oncol*. 2009;27(35):5874-5880. doi:10.1200/JCO.2009.23.6497.

6. Sanson M, Marie Y, Paris S, et al. Isocitrate dehydrogenase 1 codon 132 mutation is an important prognostic biomarker in gliomas. *J Clin Oncol*. 2009;27(25):4150-4154. doi:10.1200/JCO.2009.21.9832.

7. Van Den Bent MJ, Brandes AA, Taphoorn MJB, et al. Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: Long-term follow-up of EORTC brain tumor group study 26951. *J Clin Oncol*. 2013;31(3):344-350. doi:10.1200/JCO.2012.43.2229.

8. Cairncross G, Wang M, Shaw E, et al. Phase III trial of chemoradiotherapy for anaplastic oligodendroglioma: Long-term results of RTOG 9402. *J Clin Oncol*. 2013;31(3):337-343. doi:10.1200/JCO.2012.43.2674.

9. Buckner JC, Shaw EG, Pugh SL, et al. Radiation plus procarbazine, CCNU, and vincristine in low-grade glioma. *N Engl J Med*. 2016;374(14):1344-1355. doi:10.1056/NEJMoa1500925.

10. Cairns RA, Harris I, Mccracken S, Mak TW. Cancer cell metabolism. *Cold Spring Harb Symp Quant Biol*. 2011;76(December 2011):299-311. doi:10.1101/sqb.2011.76.012856.

11. Marie SKN, Shinjo SMO. Metabolism and brain cancer. *Clinics (Sao Paulo)*. 2011;66 Suppl 1:33-43. doi:10.1590/S1807-59322011001300005.

12. McMahon KL, Cowin G, Galloway G. Magnetic resonance imaging: The underlying principles. *J Orthop Sports Phys Ther*. 2011;41(11):806-819. doi:10.2519/jospt.2011.3576.

13. Reifenberger G, Wirsching HG, Knobbe-Thomsen CB, Weller M. Advances in the

molecular genetics of gliomas-implications for classification and therapy. *Nat Rev Clin Oncol*. 2017;14(7):434-452. doi:10.1038/nrclinonc.2016.204.

14. van den Bent MJ. Anaplastic Oligodendroglioma and Oligoastrocytoma. *Neurol Clin*. 2007;25(4):1089-1109. doi:10.1016/j.ncl.2007.07.013.

15. Assi H, Candolfi M, Baker G, Mineharu Y, Lowenstein PR, Castro MG. Gene therapy for brain tumors: Basic developments and clinical implementation. *Neurosci Lett*. 2012;527(2):71-77. doi:10.1016/j.neulet.2012.08.003.

16. Thust SC, Heiland S, Falini A, et al. Glioma imaging in Europe: A survey of 220 centres and recommendations for best clinical practice. *Eur Radiol*. 2018;28(8):3306-3317. doi:10.1007/s00330-018-5314-5.

17. Keevil SF. Magnetic resonance imaging in medicine. *Phys Educ*. 2001;36(6):476-485. doi:10.1088/0031-9120/36/6/305.

18. Grover VPB, Tognarelli JM, Crossey MME, Cox IJ, Taylor-Robinson SD, McPhail MJW. Magnetic Resonance Imaging: Principles and Techniques: Lessons for Clinicians. *J Clin Exp Hepatol*. 2015;5(3):246-255. doi:10.1016/j.jceh.2015.08.001.

19. Worthoff WA, Yun SD, Shah NJ. CHAPTER 1: Introduction to Magnetic Resonance Imaging. In: *New Developments in NMR*. Vol 2019-Janua. ; 2019:3-44. doi:10.1039/9781788013062-00001.

20. Westbrook C, Roth CK, Talbot J. *MRI in Practice. Westbrook.Pdf*. 4th ed. Oxford : Wiley-Blackwell; 2011.

21. Currie S, Hoggard N, Craven IJ, Hadjivassiliou M, Wilkinson ID. Understanding MRI: Basic MR physics for physicians. *Postgrad Med J*. 2013;89(1050):209-223. doi:10.1136/postgradmedj-2012-131342.

22. Cuenod CA, Balvay D. Perfusion and vascular permeability: Basic concepts and measurement in DCE-CT and DCE-MRI. *Diagn Interv Imaging*. 2013;94(12):1187-1204. doi:10.1016/j.diii.2013.10.010.

23. Thomsen HS, Morcos SK, Almén T, et al. Nephrogenic systemic fibrosis and gadolinium-based contrast media: Updated ESUR Contrast Medium Safety Committee guidelines. *Eur Radiol*. 2013;23(2):307-318. doi:10.1007/s00330-012-2597-9.

24. Huisman TAGM. Diffusion-weighted imaging: Basic concepts and application in cerebral stroke and head trauma. *Eur Radiol*. 2003;13(10):2283-2297. doi:10.1007/s00330-003-1843-6.

25. Hagmann P, Jonasson L, Maeder P, Thiran J-P, Wedeen VJ, Meuli R. Understanding Diffusion MR Imaging Techniques: From Scalar Diffusion-weighted Imaging to Diffusion Tensor Imaging and Beyond. *RadioGraphics*. 2006;26(suppl_1):S205-S223. doi:10.1148/rg.26si065510.

26. Le Bihan D, Breton E, Lallemand D, Aubin ML, Vignaud J, Laval-Jeantet M. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology*. 1988;168(2):497-505. doi:10.1148/radiology.168.2.3393671.

27. Le Bihan D. Intravoxel Incoherent Motion Perfusion MR Imaging: A Wake-Up Call. *Radiology*. 2008;249(3):748-752. doi:10.1148/radiol.2493081301.

28. Soares DP, Law M. Magnetic resonance spectroscopy of the brain: review of metabolites and clinical applications. *Clin Radiol*. 2009;64(1):12-21. doi:10.1016/j.crad.2008.07.002.

29. Buonocore MH, Maddock RJ. Magnetic resonance spectroscopy of the brain: A review of physical principles and technical methods. *Rev Neurosci*. 2015;26(6):609-632. doi:10.1515/revneuro-2015-0010.

30. Copen WA, Lev MH, Rapalino O. Brain perfusion. In: *Neuroimaging Part I*. Vol 135. 1st ed. Elsevier B.V.; 2016:117-135. doi:10.1016/B978-0-444-53485-9.00006-4.

31. Folkman J. Role of angiogenesis in tumor growth and metastasis. *Semin Oncol*. 2002;29(6 Suppl 16):15-18. doi:10.1053/sonc.2002.37263.

32. Dean BL, Drayer BP, Bird CR, et al. Gliomas: Classification with MR imaging. *Radiology*. 1990;174(2):411-415. doi:10.1148/radiology.174.2.2153310.

33. Felix R, Schorner W, Laniado M, et al. Brain tumors: MR imaging with gadolinium-DTPA. *Radiology*. 1985;156(3):681-688. doi:10.1148/radiology.156.3.4040643.

34. Runge VM, Clanton JA, Price AC, et al. The use of GD DTPA as a perfusion agent and marker of blood-brain barrier disruption. *Magn Reson Imaging*. 1985;3(1):43-55. doi:10.1016/0730-725X(85)90008-6.

35. Maier SE, Sun Y, Mulkern R V. Diffusion imaging of brain tumors. *NMR Biomed*. 2010;23(7):849-864. doi:10.1002/nbm.1544.

36. Bre S, Cotran R, Folkman J. Tumor angiogenesis: A quantitative method for histologic grading. *J Natl Cancer Inst*. 1972;48(2):347-356. doi:10.1093/jnci/48.2.347.

37. Wesseling P, Ruiter DJ, Burger PC. Angiogenesis in brain tumors; pathobiological and clinical aspects. *J Neurooncol*. 1997;32(3):253-265. doi:10.1023/A:1005746320099.

38. Folkman J. What is the evidence that tumors are angiogenesis dependent? *J Natl Cancer Inst*. 1990;82(1):4-7. doi:10.1093/jnci/82.1.4.

39. Noguchi T, Yoshiura T, Hiwatashi A, et al. Perfusion imaging of brain tumors using arterial spin-labeling: Correlation with histopathologic vascular density. *Am J Neuroradiol*. 2008;29(4):688-693. doi:10.3174/ajnr.A0903.

40. Ningning D, Haopeng P, Xuefei D, et al. Perfusion imaging of brain gliomas using arterial spin labeling: correlation with histopathological vascular density in MRI-guided biopsies. *Neuroradiology*. 2017;59(1):51-59. doi:10.1007/s00234-016-1756-0.

41. Tietze A, Mouridsen K, Lassen-Ramshad Y, Østergaard L. Perfusion MRI derived indices of microvascular shunting and flow control correlate with tumor grade and outcome in patients with cerebral glioma. *PLoS One*. 2015;10(4):1-16. doi:10.1371/journal.pone.0123044.

42. Abrigo JM, Fountain DM, Provenzale JM, et al. Magnetic resonance perfusion for differentiating low-grade from high-grade gliomas at first presentation. *Cochrane Database Syst Rev*. 2018;2018(1). doi:10.1002/14651858.CD011551.pub2.

43. Cebeci H, Aydin O, Ozturk-Isik E, et al. Assesment of perfusion in glial tumors with arterial spin labeling; Comparison with dynamic susceptibility contrast method. *Eur J Radiol*. 2014;83(10):1914-1919. doi:10.1016/j.ejrad.2014.07.002.

44. Liang J, Liu D, Gao P, et al. Diagnostic Values of DCE-MRI and DSC-MRI for Differentiation Between High-grade and Low-grade Gliomas: A Comprehensive Meta-analysis. *Acad Radiol*. 2018;25(3):338-348. doi:10.1016/j.acra.2017.10.001.

45. Alsaedi A, Doniselli F, Jäger HR, et al. The value of arterial spin labelling in adults glioma grading: systematic review and meta-analysis. *Oncotarget*. 2019;10(16):1589-1601. doi:10.18632/oncotarget.26674.

46. Hourani R, Brant LJ, Rizk T, Weingart JD, Barker PB, Horská A. Can proton MR spectroscopic and perfusion imaging differentiate between neoplastic and nonneoplastic brain lesions in adults? *Am J Neuroradiol*. 2008;29(2):366-372. doi:10.3174/ajnr.A0810.

47. Nakamura H, Doi M, Suzuki T, et al. The significance of lactate and lipid peaks for predicting primary neuroepithelial tumor grade with proton MR spectroscopy. *Magn Reson Med Sci*. 2018;17(3):238-243. doi:10.2463/mrms.mp.2017-0042.

48. Fellah S, Caudal D, De Paula AM, et al. Multimodal MR imaging (diffusion, perfusion, and spectroscopy): Is it possible to distinguish oligodendroglial tumor grade and 1p/19q codeletion in the pretherapeutic diagnosis? *Am J Neuroradiol*. 2013;34(7):1326-1333. doi:10.3174/ajnr.A3352.

49. Spampinato MV, Smith JK, Kwock L, et al. Cerebral blood volume measurements and proton MR spectroscopy in grading of oligodendroglial tumors. *Am J Roentgenol*. 2007;188(1):204-212. doi:10.2214/AJR.05.1177.

50. Unconfirmed 732269.crdownload.

51. Smits, Marion MD P, van den Bent, Martin J M. Imaging Correlates of Adult Glioma Genotypes. *Radiology*. 2017;284(2):316-331. https://pubs.rsna.org/doi/pdf/10.1148/radiol.2017151930.

52. Suh CH, Kim HS, Jung SC, Choi CG, Kim SJ. Imaging prediction of isocitrate dehydrogenase (IDH) mutation in patients with glioma: a systemic review and meta-analysis. *Eur Radiol*. 2019;29(2):745-758. doi:10.1007/s00330-018-5608-7.

53. Lasocki A, Anjari M, Örs Kokurcan S, Thust SC. Conventional MRI features of adult diffuse glioma molecular subtypes: a systematic review. *Neuroradiology*. 2020;(2016). doi:10.1007/s00234-020-02532-7.

54. Yamauchi T, Ohno M, Matsushita Y, et al. Radiological characteristics based on isocitrate dehydrogenase mutations and 1p/19q codeletion in grade II and III gliomas. *Brain Tumor Pathol*. 2018;35(3):148-158. doi:10.1007/s10014-018-0321-4.

55. Liu Y, Zhou Y, Zhang XS, Shen BZ. Expression of VEGF and MMP-9 and MRI imaging changes in cerebral glioma. *Oncol Lett*. 2011;2(6):1171-1175. doi:10.3892/ol.2011.384.

56. Polívka J, Pešta M, Pitule P, et al. IDH1 mutation is associated with lower expression of VEGF but not microvessel formation in glioblastoma multiforme. *Oncotarget*. 2018;9(23):16462-16476. doi:10.18632/oncotarget.24536.

57. Tan W, Xiong J, Huang W, Wu J, Zhan S, Geng D. Noninvasively detecting Isocitrate dehydrogenase 1 gene status in astrocytoma by dynamic susceptibility contrast MRI. *J Magn Reson Imaging*. 2017;45(2):492-499. doi:10.1002/jmri.25358.

58. Suh CH, Kim HS, Jung SC, Choi CG, Kim SJ. 2-Hydroxyglutarate MR spectroscopy for prediction of isocitrate dehydrogenase mutant glioma: A systemic review and meta-analysis using individual patient data. *Neuro Oncol*. 2018;20(12):1573-1583. doi:10.1093/neuonc/noy113.

59. Broen MPG, Smits M, Wijnenga MMJ, et al. The T2-FLAIR mismatch sign as an imaging marker for non-enhancing IDH-mutant, 1p/19q-intact lower-grade glioma: a validation study. *Neuro Oncol*. 2018;20(10):1393-1399. doi:10.1093/neuonc/noy048.

60. Just N. Improving tumour heterogeneity MRI assessment with histograms. *Br J Cancer*. 2014;111(12):2205-2213. doi:10.1038/bjc.2014.512.

61. Cvancarova M, Bjornerud A, Borota OC, et al. Glioma Grading by Using Histogram Analysis of Blood Volume Heterogeneity from MR-derived Cerebral Blood Volume Maps. *Radiology*. 2008;247(3):808-817. doi:10.1148/radiol.2473070571.

62. Materka A, Strzelecki M. Texture Analysis Methods – A Review. http://www.eletel.p.lodz.pl/programy/cost/pdf_1.pdf.

63. Afshar P, Mohammadi A, Plataniotis KN, Oikonomou A, Benali H. From handcrafted to deep-learning-based cancer radiomics: Challenges and opportunities. *IEEE Signal Process Mag*. 2019;36(4):132-160. doi:10.1109/MSP.2019.2900993.

64. Parekh V, Jacobs MA. Radiomics: a new application from established techniques. *Expert Rev Precis Med Drug Dev*. 2016;1(2):207-226. doi:10.1080/23808993.2016.1164013.

65. Alsaedi A, Thomas D, Bisdas S, Golay X. Overview and Critical Appraisal of Arterial Spin Labelling Technique in Brain Perfusion Imaging. *Contrast Media Mol Imaging*. 2018;2018:1-15. doi:10.1155/2018/5360375.

66. Zierler KL. Theoretical Basis of Indicator-Dilution Methods For Measuring Flow and Volume. *Circ Res*. 1962;10(3):393-407. doi:10.1161/01.RES.10.3.393.

67. MEIER P, ZIERLER KL. On the theory of the indicator-dilution method for measurement of blood flow and volume. *J Appl Physiol*. 1954;6(12):731-744. doi:10.1152/jappl.1954.6.12.731.

68. Buxton RB, Buxton RB. Contrast agent techniques. In: *Introduction to Functional Magnetic Resonance Imaging*. Cambridge: Cambridge University Press; 2010:281-306. doi:10.1017/cbo9780511605505.016.

69. Bleeker EJW, Van Osch MJP, Connelly A, Van Buchem MA, Webb AG, Calamante F. New criterion to aid manual and automatic selection of the arterial input function in dynamic susceptibility contrast MRI. *Magn Reson Med*. 2011;65(2):448-456. doi:10.1002/mrm.22599.

70. Bleeker EJW, Van Buchem MA, Van Osch MJP. Optimal location for arterial input function measurements near the middle cerebral artery in first-pass perfusion MRI. *J Cereb Blood Flow Metab*. 2009;29(4):840-852. doi:10.1038/jcbfm.2008.155.

71. Hadizadeh DR, Mädler B, Gieseke J, Fimmers R, Hattingen E, Schild HH. Effects of arterial input function selection on kinetic parameters in brain dynamic contrast-enhanced MRI. *Magn Reson Imaging*. 2017;40:83-90. doi:10.1016/j.mri.2017.04.006.

72. Calamante F, Thomas DL, Pell GS, Wiersma J, Turner R. Measuring cerebral blood flow using magnetic resonance imaging techniques. *J Cereb Blood Flow Metab*. 1999;19(7):701-735. doi:10.1097/00004647-199907000-00001.

73. Petrella JR, Provenzale JM. MR perfusion imaging of the brain: Techniques and applications. *Am J Roentgenol*. 2000;175(1):207-219. doi:10.2214/ajr.175.1.1750207.

74. Vonken EJPA, Van Osch MJP, Bakker CJG, Viergever MA. Measurement of cerebral perfusion with dual-echo multi-slice quantitative dynamic susceptibility contrast MRI. *J Magn Reson Imaging*. 1999;10(2):109-117. doi:10.1002/(sici)1522-2586(199908)10:2<109::aid-jmri1>3.0.co;2-%23.

75. Jochimsen TH, Newbould RD, Skare ST, et al. Identifying systematic errors in quantitative dynamic-susceptibility contrast perfusion imaging by high-resolution multi-echo parallel EPI. *NMR Biomed*. 2007;20(4):429-438. doi:10.1002/nbm.1107.

76. Welker K, Boxerman J, Kalnin A, Kaufmann T, Shiroishi M, Wintermark M. ASFNR recommendations for clinical performance of MR dynamic susceptibility contrast perfusion imaging of the brain. *Am J Neuroradiol*. 2015;36(6):E41-E51. doi:10.3174/ajnr.A4341.

77. Calamante F. Arterial input function in perfusion MRI: A comprehensive review. *Prog Nucl Magn Reson Spectrosc*. 2013;74:1-32. doi:10.1016/j.pnmrs.2013.04.002.

78. Benner T, Heiland S, Erb G, Forsting M, Sartor K. Accuracy of gamma-variate fits to concentration-time curves from dynamic susceptibility-contrast enhanced MRI: Influence of time resolution, maximal signal drop and signal-to-noise. *Magn Reson Imaging*. 1997;15(3):307-317. doi:10.1016/S0730-725X(96)00392-X.

79. Mouridsen K, Friston K, Hjort N, Gyldensted L, Østergaard L, Kiebel S. Bayesian estimation of cerebral perfusion using a physiological model of microvasculature. *Neuroimage*. 2006;33(2):570-579. doi:10.1016/j.neuroimage.2006.06.015.

80. Khalifa F, Soliman A, El-Baz A, et al. Models and methods for analyzing DCE-MRI: A review. *Med Phys*. 2014;41(12). doi:10.1118/1.4898202.

81. Cheng HLM, Wright GA. Rapid high-rsolution T1 mapping by variable flip angles: Accurate and precise measurements in the presence of radiofrequency field inhomogeneity. *Magn Reson Med*. 2006;55(3):566-574. doi:10.1002/mrm.20791.

82. Guo JY, Reddick WE, Rosen MA, Song HK. Dynamic contrast-enhanced magnetic resonance imaging parameters independent of baseline T10 values. *Magn Reson Imaging*. 2009;27(9):1208-1215. doi:10.1016/j.mri.2009.05.015.

83. Balvay D, Frouin F, Calmon G, et al. New criteria for assessing fit quality in dynamic contrast-enhanced T 1-weighted MRI for perfusion and permeability imaging. *Magn Reson Med*. 2005;54(4):868-877. doi:10.1002/mrm.20650.

84. Brendle C, Hempel JM, Schittenhelm J, et al. Glioma Grading and Determination of IDH Mutation Status and ATRX loss by DCE and ASL Perfusion. *Clin Neuroradiol*. 2018;28(3):421-428. doi:10.1007/s00062-017-0590-z.

85. Arevalo-Perez J, Peck KK, Young RJ, Holodny AI, Karimi S, Lyo JK. Dynamic Contrast-Enhanced Perfusion MRI and Diffusion-Weighted Imaging in Grading of Gliomas. *J Neuroimaging*. 2015;25(5):792-798. doi:10.1111/jon.12239.

86. Jung SC, Yeom JA, Kim JH, et al. Glioma: Application of histogram analysis of pharmacokinetic parameters from T1-weighted dynamic contrast-enhanced MR imaging to tumor grading. *Am J Neuroradiol*. 2014;35(6):1103-1110. doi:10.3174/ajnr.A3825.

87. Tofts PS. Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *J Magn Reson Imaging*. 1997;7(1):91-101. doi:10.1002/jmri.1880070113.

88. Kershaw LE, Cheng HLM. Temporal resolution and SNR requirements for accurate DCE-MRI data analysis using the AATH model. *Magn Reson Med*. 2010;64(6):1772-1780. doi:10.1002/mrm.22573.

89. Detre JA, Leigh JS, Williams DS, Koretsky AP. Perfusion imaging. *Magn Reson Med*. 1992;23(1):37-45. doi:10.1002/mrm.1910230106.

90. Frank LR, Wong EC, Buxton RB. Slice profile effects in adiabatic inversion: Application to multislice perfusion imaging. *Magn Reson Med*. 1997;38(4):558-564. doi:10.1002/mrm.1910380409.

91. Zhang W, Williams DS, Detre JA, Koretsky AP. Measurement of brain perfusion by volume-localized NMR spectroscopy using inversion of arterial water spins: Accounting for transit time and cross-relaxation. *Magn Reson Med*. 1992;25(2):362-371. doi:10.1002/mrm.1910250216.

92. Pekar J, Jezzard P, Roberts DA, Leigh JS, Frank JA, McLaughlin AG. Perfusion imaging with compensation for asymmetric magnetization transfer effects. *Magn Reson Med*. 1996;35(1):70-79. doi:10.1002/mrm.1910350110.

93. McLaughlin AC, Ye FQ, Pekar JJ, Santha AKS, Frank JA. Effect of magnetization transfer on the measurement of cerebral blood flow using steady-state arterial spin tagging approaches: A theoretical investigation. *Magn Reson Med*. 1997;37(4):501-510. doi:10.1002/mrm.1910370406.

94. Williams DS, Detre JA, Leigh JS, Koretsky AP. Magnetic resonance imaging of perfusion using spin inversion of arterial water. *Proc Natl Acad Sci*. 1992;89(9):4220-4220. doi:10.1073/pnas.89.9.4220e.

95. Dai W, Garcia D, De Bazelaire C, Alsop DC. Continuous flow-driven inversion for arterial spin labeling using pulsed radio frequency and gradient fields. *Magn Reson Med*. 2008;60(6):1488-1497. doi:10.1002/mrm.21790.

96. Wang SJ, Nishimura DG, Macovski A. Fast angiography using selective inversion recovery. *Magn Reson Med*. 1992;23(1):109-121. doi:10.1002/mrm.1910230112.

97. Edelman RR, Siewert B, Adamis M, Gaa J, Laub G, Wielopolski P. Signal targeting with alternating radiofrequency (STAR) sequences: Application to MR angiography. *Magn Reson Med*. 1994;31(2):233-238. doi:10.1002/mrm.1910310219.

98. Kim S -G. Quantification of relative cerebral blood flow change by flow-sensitive alternating inversion recovery (FAIR) technique: Application to functional mapping. *Magn Reson Med*. 1995;34(3):293-301. doi:10.1002/mrm.1910340303.

99. Wong EC, Buxton RB, Frank LR. Implementation of quantitative perfusion imaging techniques for functional brain mapping using pulsed arterial spin labeling. *NMR Biomed*. 1997;10(4-5):237-249. doi:10.1002/(SICI)1099-1492(199706/08)10:4/5<237::AID-NBM475>3.0.CO;2-X.

100. Alsop DC, Detre JA, Golay X, et al. Recommended implementation of arterial spin-labeled Perfusion mri for clinical applications: A consensus of the ISMRM Perfusion Study group and the European consortium for ASL in dementia. *Magn Reson Med*. 2015;73(1):102-116. doi:10.1002/mrm.25197.

101. Silva AC, Zhang W, Williams DS, Koretsky AP. Multi-Slice MRI of Rat Brain Perfusion During Amphetamine Stimulation Using Arterial Spin Labeling. *Magn Reson Med*. 1995;33(2):209-214. doi:10.1002/mrm.1910330210.

102. Zaharchuk G, Ledden PJ, Kwong KK, Reese TG, Rosen BR, Wald LL. Multislice perfusion and perfusion territory imaging in humans with separate label and image coils. *Magn Reson Med*. 1999;41(6):1093-1098. doi:10.1002/(SICI)1522-2594(199906)41:6<1093::AID-MRM4>3.0.CO;2-0.

103. Alsop DC, Detre JA. Multisection cerebral blood flow MRI imaging with continuous arterial spin labeling. *Radiology*. 1998;208(2):410-416. doi:10.1148/radiology.208.2.9680569.

104. Buxton RB. *Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques*. Vol 9780521899. Cambridge: Cambridge University Press; 2009. doi:10.1017/CBO9780511605505.

105. Wong EC. An introduction to ASL labeling techniques. *J Magn Reson Imaging*. 2014;40(1):1-10. doi:10.1002/jmri.24565.

106. Alsop DC, Detre JA. Reduced transit-time sensitivity in noninvasive magnetic resonance imaging of human cerebral blood flow. *J Cereb Blood Flow Metab*. 1996;16(6):1236-1249. doi:10.1097/00004647-199611000-00019.

107. Wong EC, Buxton RB, Frank LR. Quantitative imaging of perfusion using a single subtraction (QUIPSS and QUIPSS II). *Magn Reson Med*. 1998;39(5):702-708. doi:10.1002/mrm.1910390506.

108. Luh WM, Wong EC, Bandettini PA, Hyde JS. QUIPSS II with thin-slice TI1 periodic saturation: A method for improving accuracy of quantitative perfusion imaging using pulsed arterial spin labeling. *Magn Reson Med*. 1999;41(6):1246-1254. doi:10.1002/(SICI)1522-2594(199906)41:6<1246::AID-MRM22>3.0.CO;2-N.

109. Trampel R, Jochimsen TH, Mildner T, Norris DG, Möller HE. Efficiency of flow-driven adiabatic spin inversion under realistic experimental conditions: A computer simulation. *Magn Reson Med*. 2004;51(6):1187-1193. doi:10.1002/mrm.20080.

110. Maccotta L, Detre JA, Alsop DC. The efficiency of adiabatic inversion for perfusion imaging by arterial spin labeling. *NMR Biomed*. 1997;10(4-5):216-221. doi:10.1002/(SICI)1099-1492(199706/08)10:4/5<216::AID-NBM468>3.0.CO;2-U.

111. Buxton RB, Frank LR, Wong EC, Siewert B, Warach S, Edelman RR. A general kinetic model for quantitative perfusion imaging with arterial spin labeling. *Magn Reson Med*. 1998;40(3):383-396. doi:10.1002/mrm.1910400308.

112. Wansapura JP, Holland SK, Dunn RS, Ball WS. NMR relaxation times in the human brain at 3.0 Tesla. *J Magn Reson Imaging*. 1999;9(4):531-538. doi:10.1002/(SICI)1522-2586(199904)9:4<531::AID-JMRI4>3.0.CO;2-L.

113. Lu H, Clingman C, Golay X, Van Zijl PCM. Determining the longitudinal relaxation time (T1) of blood at 3.0 tesla. *Magn Reson Med*. 2004;52(3):679-682. doi:10.1002/mrm.20178.

114. Lu H, Nagae-Poetscher LM, Golay X, Lin D, Pomper M, Van Zijl PCM. Routine clinical brain MRI sequences for use at 3.0 tesla. *J Magn Reson Imaging*. 2005;22(1):13-22. doi:10.1002/jmri.20356.

115. Franke C, Van Dorsten FA, Olah L, Schwindt W, Hoehn M. Arterial spin tagging perfusion imaging of rat brain dependency on magnetic field strength. *Magn Reson Imaging*. 2000;18(9):1109-1113. doi:10.1016/S0730-725X(00)00211-3.

116. Wang J, Alsop DC, Li L, et al. Comparison of quantitative perfusion imaging using arterial spin labeling at 1.5 and 4.0 Tesla. *Magn Reson Med*. 2002;48(2):242-254. doi:10.1002/mrm.10211.

117. Van Gelderen P, De Zwart JA, Duyn JH. Pittfalls of MRI measurement of white matter perfusion based on arterial spin labeling. *Magn Reson Med*. 2008;59(4):788-795. doi:10.1002/mrm.21515.

118. Helenius J, Perkiö J, Soinne L, et al. Cerebral hemodynamics in a healthy population measured by dynamic susceptibility contrast MR imaging. *Acta radiol*. 2003;44(5):538-546. doi:10.1080/j.1600-0455.2003.00104.x.

119. Ye FQ, Frank JA, Weinberger DR, McLaughlin AC. Noise reduction in 3D perfusion imaging by attenuating the static signal in arterial spin tagging (ASSIST). *Magn Reson Med*. 2000;44(1):92-100. doi:10.1002/1522-2594(200007)44:1<92::AID-MRM14>3.0.CO;2-M.

120. Garcia DM, Duhamel G, Alsop DC. Efficiency of inversion pulses for background suppressed arterial spin labeling. *Magn Reson Med*. 2005;54(2):366-372. doi:10.1002/mrm.20556.

121. Gunther M, Bock M, Schad LR. Arterial spin labeling in combination with a look-locker sampling strategy: Inflow turbo-sampling EPI-FAIR (ITS-FAIR). *Magn Reson Med*. 2001;46(5):974-984. doi:10.1002/mrm.1284.

122. Francis ST, Bowtell R, Gowland PA. Modeling and optimization of look-locker spin labeling for measuring perfusion and transit time changes in activation studies taking into account arterial blood volume. *Magn Reson Med*. 2008;59(2):316-325. doi:10.1002/mrm.21442.

123. Wells JA, Lythgoe MF, Gadian DG, Ordidge RJ, Thomas DL. In vivo Hadamard encoded Continuous arterial spin labeling (H-CASL). *Magn Reson Med*. 2010;63(4):1111-1118. doi:10.1002/mrm.22266.

124. Dai W, Shankaranarayanan A, Alsop DC. Volumetric measurement of perfusion

and arterial transit delay using hadamard encoded continuous arterial spin labeling. *Magn Reson Med*. 2013;69(4):1014-1022. doi:10.1002/mrm.24335.

125. Dai W, Robson PM, Shankaranarayanan A, Alsop DC. Reduced resolution transit delay prescan for quantitative continuous arterial spin labeling perfusion imaging. *Magn Reson Med*. 2012;67(5):1252-1265. doi:10.1002/mrm.23103.

126. Petersen ET, Lim T, Golay X. Model-free arterial spin labeling quantification approach for perfusion MRI. *Magn Reson Med*. 2006;55(2):219-232. doi:10.1002/mrm.20784.

127. Shen N, Zhao L, Jiang J, et al. Intravoxel incoherent motion diffusion-weighted imaging analysis of diffusion and microperfusion in grading gliomas and comparison with arterial spin labeling for evaluation of tumor perfusion. *J Magn Reson Imaging*. 2016;44(3):620-632. doi:10.1002/jmri.25191.

128. Lin Y, Li J, Zhang Z, et al. Comparison of intravoxel incoherent motion diffusion-weighted MR imaging and arterial spin labeling MR imaging in gliomas. *Biomed Res Int*. 2015;2015:1-10. doi:10.1155/2015/234245.

129. Yang XS, Zhao B, Wang G, et al. Improving the grading accuracy of astrocytic neoplasms noninvasively by combining timing information with cerebral blood flow: A multi-Ti arterial spin-labeling MR imaging study. *Am J Neuroradiol*. 2016;37(12):2209-2216. doi:10.3174/ajnr.A4907.

130. Roy B, Awasthi R, Bindal A, et al. Comparative evaluation of 3-dimensional pseudocontinuous arterial spin labeling with dynamic contrast-enhanced perfusion magnetic resonance imaging in grading of human glioma. *J Comput Assist Tomogr*. 2013;37(3):321-326. doi:10.1097/RCT.0b013e318282d7e2.

131. Bai Y, Lin Y, Zhang W, et al. Noninvasive amide proton transfer magnetic resonance imaging in evaluating the grading and cellularity of gliomas. *Oncotarget*. 2017;8(4):5834-5842. doi:10.18632/oncotarget.13970.

132. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*. 2009;339(7716):332-336. doi:10.1136/bmj.b2535.

133. Sun Y, Schmidt NO, Schmidt K, et al. Perfusion MRI of U87 Brain Tumors in a Mouse Model. *Magn Reson Med*. 2004;51(5):893-899. doi:10.1002/mrm.20029.

134. Wong ET, Timmons J, Callahan A, O'Loughlin L, Giarusso B, Alsop DC. Phase I study of low-dose metronomic temozolomide for recurrent malignant gliomas. *BMC Cancer*. 2016;16(1):914. doi:10.1186/s12885-016-2945-2.

135. Lyu Y, Liu S, You H, et al. Evaluation of recurrent high-grade gliomas treated with bevacizumab: A preliminary report of 3D pseudocontinuous artery spin labeling. *J Magn Reson Imaging*. 2017;46(2):565-573. doi:10.1002/jmri.25558.

136. Morana G, Piccardo A, Tortora D, et al. Grading and outcome prediction of pediatric diffuse astrocytic tumors with diffusion and arterial spin labeling

perfusion MRI in comparison with 18F–DOPA PET. *Eur J Nucl Med Mol Imaging*. 2017;44(12):2084-2093. doi:10.1007/s00259-017-3777-2.

137. Whiting PF, Rutjes AWS, Westwood ME, et al. Quadas-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536. doi:10.7326/0003-4819-155-8-201110180-00009.

138. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Br Med J*. 2003;327(7414):557-560. doi:10.1136/bmj.327.7414.557.

139. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539-1558. doi:10.1002/sim.1186.

140. Li PMMC, Linos E, Gurgel RK, Fischbein NJ, Blevins NH. Evaluating the utility of non-echo-planar diffusion-weighted imaging in the preoperative evaluation of cholesteatoma: A meta-analysis. *Laryngoscope*. 2013;123(5):1247-1250. doi:10.1002/lary.23759.

141. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med*. 2001;20(19):2865-2884. doi:10.1002/sim.942.

142. Fudaba H, Shimomura T, Abe T, et al. Comparison of multiple parameters obtained on 3T pulsed arterial spin-labeling, diffusion tensor imaging, and MRS and the Ki-67 labeling index in evaluating glioma grading. *Am J Neuroradiol*. 2014;35(11):2091-2098. doi:10.3174/ajnr.A4018.

143. Kong L, Chen H, Yang Y, Chen L. A meta-analysis of arterial spin labelling perfusion values for the prediction of glioma grade. *Clin Radiol*. 2017;72(3):255-261. doi:10.1016/j.crad.2016.10.016.

144. Wolf RL, Wang J, Wang S, et al. Grading of CNS neoplasms using continuous arterial spin labeled perfusion MR imaging at 3 Tesla. *J Magn Reson Imaging*. 2005;22(4):475-482. doi:10.1002/jmri.20415.

145. Chawla S, Wang S, Wolf RL, et al. Arterial spin-labeling and MR spectroscopy in the differentiation of gliomas. *Am J Neuroradiol*. 2007;28(9):1683-1689. doi:10.3174/ajnr.A0673.

146. Furtner J, Schöpf V, Schewzow K, et al. Arterial spin-labeling assessment of normalized vascular intratumoral signal intensity as a predictor of histologic grade of astrocytic neoplasms. *Am J Neuroradiol*. 2014;35(3):482-489. doi:10.3174/ajnr.A3705.

147. Kim MJ, Kim HS, Kim JH, Cho KG, Kim SY. Diagnostic accuracy and interobserver variability of pulsed arterial spin labeling for glioma grading. *Acta radiol*. 2008;49(4):450-457. doi:10.1080/02841850701881820.

148. Warmuth C, Günther M, Zimmer C. Quantification of blood flow in brain tumors: Comparison of arterial spin labeling and dynamic susceptibility-weighted

contrast-enhanced MR imaging. *Radiology*. 2003;228(2):523-532. doi:10.1148/radiol.2282020409.

149. Yamashita K, Hiwatashi A, Togao O, et al. MR imaging-based analysis of glioblastoma multiforme: Estimation of IDH1 mutation status. *Am J Neuroradiol*. 2016;37(1):58-65. doi:10.3174/ajnr.A4491.

150. Wong EC, Buxton RB, Frank LR. A theoretical and experimental comparison of continuous and pulsed arterial spin labeling techniques for quantitative perfusion imaging. *Magn Reson Med*. 1998;40(3):348-355. doi:10.1002/mrm.1910400303.

151. Chen Y, Wang DJJ, Detre JA. Test-retest reliability of arterial spin labeling with common labeling strategies. *J Magn Reson Imaging*. 2011;33(4):940-949. doi:10.1002/jmri.22345.

152. Canale S, Rodrigo S, Tourdias T, et al. Évaluation du degré de malignité des tumeurs gliales primitives de l'adulte en IRM de perfusion par marquage des spins. *J Neuroradiol*. 2011;38(4):207-213. doi:10.1016/j.neurad.2010.12.003.

153. Kim HS, Kim SY. A prospective study on the added value of pulsed arterial spin-labeling and apparent diffusion coefficients in the grading of gliomas. *Am J Neuroradiol*. 2007;28(9):1693-1699. doi:10.3174/ajnr.A0674.

154. Weber MA, Zoubaa S, Schlieter M, et al. Diagnostic performance of spectroscopic and perfusion MRI for distinction of brain tumors. *Neurology*. 2006;66(12):1899-1906. doi:10.1212/01.wnl.0000219767.49705.9c.

155. Wetzel SG, Cha S, Johnson G, et al. Relative cerebral blood volume measurements in intracranial mass lesions: Interobserver and intraobserver reproducibility study. *Radiology*. 2002;224(3):797-803. doi:10.1148/radiol.2243011014.

156. Nabavizadeh SA, Assadsangabi R, Hajmomenian M, Santi M, Vossough A. High accuracy of arterial spin labeling perfusion imaging in differentiation of pilomyxoid from pilocytic astrocytoma. *Neuroradiology*. 2015;57(5):527-533. doi:10.1007/s00234-015-1497-5.

157. Jiang J, Zhao L, Zhang Y, et al. Comparative analysis of arterial spin labeling and dynamic susceptibility contrast perfusion imaging for quantitative perfusion measurements of brain tumors. *Int J Clin Exp Pathol*. 2014;7(6):2790-2799.

158. Gao F, Guo R, Hu XJ, Li CJ, Li M. Noninvasive tumor grading of glioblastomas before surgery using arterial spin labeling: A cohort study. *Anal Quant Cytol Histol*. 2015;37(6):339-346. http://www.ncbi.nlm.nih.gov/pubmed/26860009.

159. White CM, Pope WB, Zaw T, et al. Regional and Voxel-Wise Comparisons of Blood Flow Measurements Between Dynamic Susceptibility Contrast Magnetic Resonance Imaging (DSC-MRI) and Arterial Spin Labeling (ASL) in Brain Tumors. *J Neuroimaging*. 2014;24(1):23-30. doi:10.1111/j.1552-6569.2012.00703.x.

160. Van Westen D, Petersen ET, Wirestam R, et al. Correlation between arterial blood volume obtained by arterial spin labelling and cerebral blood volume in intracranial tumours. *Magn Reson Mater Physics, Biol Med*. 2011;24(4):211-223. doi:10.1007/s10334-011-0255-x.

161. Hirai T, Kitajima M, Nakamura H, et al. Quantitative blood flow measurements in gliomas using arterial spin-labeling at 3T: Intermodality agreement and inter- and intraobserver reproducibility study. *Am J Neuroradiol*. 2011;32(11):2073-2079. doi:10.3174/ajnr.A2725.

162. Lee S, Choi SH, Ryoo I, et al. Evaluation of the microenvironmental heterogeneity in high-grade gliomas with IDH1/2 gene mutation using histogram analysis of diffusion-weighted imaging and dynamic-susceptibility contrast perfusion imaging. *J Neurooncol*. 2015;121(1):141-150. doi:10.1007/s11060-014-1614-z.

163. Guo BJ, Yang ZL, Zhang LJ. Gadolinium Deposition in Brain: Current Scientific Evidence and Future Perspectives. *Front Mol Neurosci*. 2018;11(September):1-12. doi:10.3389/fnmol.2018.00335.

164. Paulson ES, Schmainda KM. Comparison of dynamic susceptibility-weighted contrast-enhanced MR methods: Recommendations for measuring relative cerebral blood volume in brain tumors. *Radiology*. 2008;249(2):601-613. doi:10.1148/radiol.2492071659.

165. Yoo RE, Choi SH, Cho HR, et al. Tumor blood flow from arterial spin labeling perfusion MRI: A key parameter in distinguishing high-grade gliomas from primary cerebral lymphomas, and in predicting genetic biomarkers in high-grade gliomas. *J Magn Reson Imaging*. 2013;38(4):852-860. doi:10.1002/jmri.24026.

166. Weber MA, Thilmann C, Lichy MP, et al. Assessment of Irradiated Brain Metastases by Means of Arterial Spin-Labeling and Dynamic Susceptibility-Weighted Contrast-Enhanced Perfusion MRI: Initial Results. *Invest Radiol*. 2004;39(5):277-287. doi:10.1097/01.rli.0000119195.50515.04.

167. Yun TJ, Cho HR, Choi SH, et al. Antiangiogenic effect of bevacizumab: Application of arterial spin-labeling perfusion MR imaging in a rat glioblastoma model. *Am J Neuroradiol*. 2016;37(9):1650-1656. doi:10.3174/ajnr.A4800.

168. Wang L, Wei L, Wang J, et al. Evaluation of perfusion MRI value for tumor progression assessment after glioma radiotherapy: A systematic review and meta-analysis. *Medicine (Baltimore)*. 2020;99(52):e23766. doi:10.1097/MD.0000000000023766.

169. O'Connor JPB, Aboagye EO, Adams JE, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol*. 2017;14(3):169-186. doi:10.1038/nrclinonc.2016.162.

170. Petersen ET, Mouridsen K, Golay X. The QUASAR reproducibility study, Part II: Results from a multi-center Arterial Spin Labeling test-retest study. *Neuroimage*.

2010;49(1):104-113. doi:10.1016/j.neuroimage.2009.07.068.

171. Zhou L, Wang Y, Pinho MC, et al. Intrasession Reliability of Arterial Spin-Labeled MRI–Measured Noncontrast Perfusion in Glioblastoma at 3 T. *Tomography*. 2020;6(2):139-147. doi:10.18383/j.tom.2020.00010.

172. Figueiredo PM, Clare S, Jezzard P. Quantitative perfusion measurements using pulsed arterial spin labeling: Effects of large region-of-interest analysis. *J Magn Reson Imaging*. 2005;21(6):676-682. doi:10.1002/jmri.20329.

173. Fallatah S. Perfusion Imaging in Brain Tumours. 2016.

174. Melbourne A, Toussaint N, Owen D, et al. NiftyFit: a Software Package for Multi-parametric Model-Fitting of 4D Magnetic Resonance Imaging Data. *Neuroinformatics*. 2016;14(3):319-337. doi:10.1007/s12021-016-9297-6.

175. Jenkinson M, Bannister P, Brady M, Smith S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *Neuroimage*. 2002;17(2):825-841. doi:10.1006/nimg.2002.1132.

176. Mignani S, Rosa R. The moving block bootstrap to assess the accuracy of statistical estimates in Ising model simulations. *Comput Phys Commun*. 1995;92(2-3):203-213. doi:10.1016/0010-4655(95)00114-7.

177. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*. 2006;31(3):1116-1128. doi:10.1016/j.neuroimage.2006.01.015.

178. Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage*. 2011;56(3):907-922. doi:10.1016/j.neuroimage.2011.02.046.

179. Modat M, Ridgway GR, Taylor ZA, et al. Fast free-form deformation using graphics processing units. *Comput Methods Programs Biomed*. 2010;98(3):278-284. doi:10.1016/j.cmpb.2009.09.002.

180. NEMA. NEMA Standards Publication MS 1-2001: Determination of Signal-to-Noise Ratio (SNR) in Diagnostic Magnetic Resonance Imaging. *Natl Electr Manuf Assoc*. 2014;2008:21. https://www.nema.org/Standards/Pages/Determination-of-Signal-to-Noise-Ratio-in-Diagnostic-Magnetic-Resonance-Imaging.aspx.

181. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012.

182. Bland M. How do I calculate a within-subject coefficient of variation. https://www-users.york.ac.uk/~mb55/meas/cv.htm. Published 2006.

183. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The Case for Using the Repeatability Coefficient When Calculating Test-Retest Reliability. *PLoS One*.

2013;8(9):1-7. doi:10.1371/journal.pone.0073990.

184. Schuck P, Zwingmann C. The 'smallest real difference' as a measure of sensitivity to change: a critical analysis. *Int J Rehabil Res*. 2003;26(2):85-91. doi:10.1097/01.mrr.0000070759.63544.65.

185. Wang Y, Saykin AJ, Pfeuffer J, et al. Regional reproducibility of pulsed arterial spin labeling perfusion imaging at 3T. *Neuroimage*. 2011;54(2):1188-1195. doi:10.1016/j.neuroimage.2010.08.043.

186. Mutsaerts HJMM, Steketee RME, Heijtel DFR, et al. Inter-vendor reproducibility of pseudo-continuous arterial spin labeling at 3 Tesla. *PLoS One*. 2014;9(8). doi:10.1371/journal.pone.0104108.

187. Aslan S, Lu H. On the sensitivity of ASL MRI in detecting regional differences in cerebral blood flow. *Magn Reson Imaging*. 2010;28(7):928-935. doi:10.1016/j.mri.2010.03.037.

188. Mezue M, Segerdahl AR, Okell TW, Chappell MA, Kelly ME, Tracey I. Optimization and reliability of multiple postlabeling delay pseudo-continuous arterial spin labeling during rest and stimulus-induced functional task activation. *J Cereb Blood Flow Metab*. 2014;34(12):1919-1927. doi:10.1038/jcbfm.2014.163.

189. Raunig DL, McShane LM, Pennello G, et al. Quantitative imaging biomarkers: A review of statistical methods for technical performance assessment. *Stat Methods Med Res*. 2015;24(1):27-67. doi:10.1177/0962280214537344.

190. Sokolska M, Bainbridge A, Rojas-Villabona A, Golay X, Thomas DL. Effect of labelling plane angulation and position on labelling efficiency and cerebral blood flow quantification in pseudo-continuous arterial spin labelling. *Magn Reson Imaging*. 2019;59(November 2018):61-67. doi:10.1016/j.mri.2019.02.007.

191. Robertson AD, Matta G, Basile VS, et al. Temporal and spatial variances in arterial spin-labeling are inversely related to large-artery blood velocity. *Am J Neuroradiol*. 2017;38(8):1555-1561. doi:10.3174/ajnr.A5257.

192. Lin Y, Xing Z, She D, et al. IDH mutant and 1p/19q co-deleted oligodendrogliomas: tumor grade stratification using diffusion-, susceptibility-, and perfusion-weighted MRI. *Neuroradiology*. 2017;59(6):555-562. doi:10.1007/s00234-017-1839-6.

193. Xing Z, Yang X, She D, Lin Y, Zhang Y, Cao D. Noninvasive assessment of IDH mutational status in World Health Organization grade II and III astrocytomas using DWI and DSC-PWI combined with conventional MR imaging. *Am J Neuroradiol*. 2017;38(6):1134-1144. doi:10.3174/ajnr.A5171.

194. Latysheva A, Emblem KE, Brandal P, et al. Dynamic susceptibility contrast and diffusion MR imaging identify oligodendroglioma as defined by the 2016 WHO classification for brain tumors: histogram analysis approach. *Neuroradiology*. 2019;61(5):545-555. doi:10.1007/s00234-019-02173-5.

195. Tian Q, Yan LF, Zhang X, et al. Radiomics strategy for glioma grading using texture features from multiparametric MRI. *J Magn Reson Imaging*. 2018;48(6):1518-1528. doi:10.1002/jmri.26010.

196. Ranjith G, Parvathy R, Vikas V, Chandrasekharan K, Nair S. Machine learning methods for the classification of gliomas: initial results using features extracted from MR spectroscopy. *Neuroradiol J*. 2015;28(2):106-111. doi:10.1177/1971400915576637.

197. Zhang X, Tian Q, Wang L, et al. Radiomics Strategy for Molecular Subtype Stratification of Lower-Grade Glioma: Detecting IDH and TP53 Mutations Based on Multimodal MRI. *J Magn Reson Imaging*. 2018;48(4):916-926. doi:10.1002/jmri.25960.

198. Jakola AS, Zhang YH, Skjulsvik AJ, et al. Quantitative texture analysis in the prediction of IDH status in low-grade gliomas. *Clin Neurol Neurosurg*. 2018;164(December 2017):114-120. doi:10.1016/j.clineuro.2017.12.007.

199. Yu J, Shi Z, Lian Y, et al. Noninvasive IDH1 mutation estimation based on a quantitative radiomics approach for grade II glioma. *Eur Radiol*. 2017;27(8):3509-3522. doi:10.1007/s00330-016-4653-3.

200. Hsieh KLC, Chen CY, Lo CM. Radiomic model for predicting mutations in the isocitrate dehydrogenase gene in glioblastomas. *Oncotarget*. 2017;8(28):45888-45897. doi:10.18632/oncotarget.17585.

201. St. Lawrence KS, Lee TY. An adiabatic approximation to the tissue homogeneity model for water exchange in the brain: II. Experimental validation. *J Cereb Blood Flow Metab*. 1998;18(12):1378-1385. doi:10.1097/00004647-199812000-00012.

202. Kang Y, Choi SH, Kim YJ, et al. Gliomas: Histogram analysis of apparent diffusion coefficient maps with standard- or high-b-value diffusion-weighted MR imaging -correlation with tumor grade. *Radiology*. 2011;261(3):882-890. doi:10.1148/radiol.11110686.

203. Mouridsen K, Christensen S, Gyldensted L, Østergaard L. Automatic selection of arterial input function using cluster analysis. *Magn Reson Med*. 2006;55(3):524-531. doi:10.1002/mrm.20759.

204. Chappell MA, Groves AR, Whitcher B, Woolrich MW. Variational Bayesian inference for a nonlinear forward model. *IEEE Trans Signal Process*. 2009;57(1):223-236. doi:10.1109/TSP.2008.2005752.

205. Rao RB, Fung G, Rosales R. On the dangers of cross-validation. An experimental evaluation. *Soc Ind Appl Math - 8th SIAM Int Conf Data Min 2008, Proc Appl Math 130*. 2008;2:588-596. doi:10.1137/1.9781611972788.54.

206. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics*. 1977;33(1):159. doi:10.2307/2529310.

207. Su C, Jiang J, Zhang S, et al. Radiomics based on multicontrast MRI can precisely

differentiate among glioma subtypes and predict tumour-proliferative behaviour. *Eur Radiol*. 2019;29(4):1986-1996. doi:10.1007/s00330-018-5704-8.

208. Lu CF, Hsu FT, Hsieh KLC, et al. Machine learning–based radiomics for molecular subtyping of gliomas. *Clin Cancer Res*. 2018;24(18):4429-4436. doi:10.1158/1078-0432.CCR-17-3445.

209. Li Z, Wang Y, Yu J, Guo Y, Cao W. Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. *Sci Rep*. 2017;7(1):5467. doi:10.1038/s41598-017-05848-2.

210. Pei L, Vidyaratne L, Rahman MM, Iftekharuddin KM. Context aware deep learning for brain tumor segmentation, subtype classification, and survival prediction using radiology images. *Sci Rep*. 2020;10(1):19726. doi:10.1038/s41598-020-74419-9.

211. Ge C, Gu IY-H, Jakola AS, Yang J. Deep semi-supervised learning for brain tumor classification. *BMC Med Imaging*. 2020;20(1):87. doi:10.1186/s12880-020-00485-0.

212. Okuchi S, Rojas-Garcia A, Ulyte A, et al. Diagnostic accuracy of dynamic contrast-enhanced perfusion MRI in stratifying gliomas: A systematic review and meta-analysis. *Cancer Med*. 2019;8(12):5564-5573. doi:10.1002/cam4.2369.

213. Ewing JR, Bagher-Ebadian H. Model selection in measures of vascular parameters using dynamic contrast-enhanced MRI: Experimental and clinical applications. *NMR Biomed*. 2013;26(8):1028-1041. doi:10.1002/nbm.2996.

214. Cramer SP, Larsson HBW. Accurate determination of blood-brain barrier permeability using dynamic contrast-enhanced T1-weighted MRI: A simulation and in vivo study on healthy subjects and multiple sclerosis patients. *J Cereb Blood Flow Metab*. 2014;34(10):1655-1665. doi:10.1038/jcbfm.2014.126.

215. Henderson E, Rutt BK, Lee TY. Temporal sampling requirements for the tracer kinetics modeling of breast disease. *Magn Reson Imaging*. 1998;16(9):1057-1073. doi:10.1016/S0730-725X(98)00130-1.

216. Cheng HLM. Investigation and optimization of parameter accuracy in dynamic contrast-enhanced MRI. *J Magn Reson Imaging*. 2008;28(3):736-743. doi:10.1002/jmri.21489.

217. Parker GJM, Padhani AR. T1-W DCE-MRI:T1-Weighted Dynamic Contrast-Enhanced MRI. In: *Quantitative MRI of the Brain*. Vol 13. Chichester, UK: John Wiley & Sons, Ltd; 2012:341-364. doi:10.1002/0470869526.ch10.

218. Rata M, Collins DJ, Darcy J, et al. Assessment of repeatability and treatment response in early phase clinical trials using DCE-MRI: comparison of parametric analysis using MR- and CT-derived arterial input functions. *Eur Radiol*. 2016;26(7):1991-1998. doi:10.1007/s00330-015-4012-9.

219. Ziayee F, Müller-Lutz A, Gross J, et al. Influence of arterial input function (AIF) on

quantitative prostate dynamic contrast-enhanced (DCE) MRI and zonal prostate anatomy. *Magn Reson Imaging*. 2018;53(March):28-33. doi:10.1016/j.mri.2018.06.004.

220. Zeng Q, Jiang B, Shi F, Ling C, Dong F, Zhang J. 3D pseudocontinuous arterial spin-labeling MR imaging in the preoperative evaluation of gliomas. *Am J Neuroradiol*. 2017;38(10):1876-1883. doi:10.3174/ajnr.A5299.

221. Hall MA, Smith LA. Correlation-based Filter Approach vs. Wrapper. *FLAIRS Conf*. 1995:235-239.

222. Bhandari AP, Liong R, Koppen J, Murthy SV, Lasocki A. Noninvasive Determination of IDH and 1p19q Status of Lower-grade Gliomas Using MRI Radiomics: A Systematic Review. *Am J Neuroradiol*. 2021;42(1):94-101. doi:10.3174/ajnr.A6875.

223. Park JE, Kim HS, Kim D, et al. A systematic review reporting quality of radiomics research in neuro-oncology: Toward clinical utility and quality improvement using high-dimensional imaging features. *BMC Cancer*. 2020;20(1):1-11. doi:10.1186/s12885-019-6504-5.

224. Fournier L, Costaridou L, Bidaut L, et al. Incorporating radiomics into clinical trials: expert consensus on considerations for data-driven compared to biologically driven quantitative biomarkers. *Eur Radiol*. 2021. doi:10.1007/s00330-020-07598-8.

225. Parmar C, Velazquez ER, Leijenaar R, et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS One*. 2014;9(7):1-8. doi:10.1371/journal.pone.0102107.

226. Yip SSF, Aerts HJWL. Applications and limitations of radiomics. *Phys Med Biol*. 2016;61(13):R150-R166. doi:10.1088/0031-9155/61/13/R150.

227. Pei L, Vidyaratne L, Rahman MM, Iftekharuddin KM. Context aware deep learning for brain tumor segmentation, subtype classification, and survival prediction using radiology images. *Sci Rep*. 2020;10(1):19726. doi:10.1038/s41598-020-74419-9.

228. Keenan KE, Ainslie M, Barker AJ, et al. Quantitative magnetic resonance imaging phantoms: A review and the need for a system phantom. *Magn Reson Med*. 2018;79(1):48-61. doi:10.1002/mrm.26982.

229. Buckler AJ, Bresolin L, Dunnick NR, Sullivan DC. A collaborative enterprise for multi-stakeholder participation in the advancement of quantitative imaging. *Radiology*. 2011;258(3):906-914. doi:10.1148/radiol.10100799.

230. Dickie BR, Parker GJM, Parkes LM. Measuring water exchange across the blood-brain barrier using MRI. *Prog Nucl Magn Reson Spectrosc*. 2020;116:19-39. doi:10.1016/j.pnmrs.2019.09.002.

231. Mutsaerts HJMM, Petr J, Václavů L, et al. The spatial coefficient of variation in

arterial spin labeling cerebral blood flow images. *J Cereb Blood Flow Metab*. 2017;37(9):3184-3192. doi:10.1177/0271678X16683690.

232. Law M, Cha S, Knopp EA, Johnson G, Arnett J, Litt AW. High-grade gliomas and solitary metastases: Differentiation by using perfusion and proton spectroscopic MR imaging. *Radiology*. 2002;222(3):715-721. doi:10.1148/radiol.2223010558.

233. Burtscher IM, Skagerberg G, Geijer B, Englund E, Ståhlberg F, Holtås S. Proton MR spectroscopy and preoperative diagnostic accuracy: An evaluation of intracranial mass lesions characterized by stereotactic biopsy findings. *Am J Neuroradiol*. 2000;21(1):84-93.

234. Yamashita K, Yoshiura T, Hiwatashi A, et al. Arterial spin labeling of hemangioblastoma: Differentiation from metastatic brain tumors based on quantitative blood flow measurement. *Neuroradiology*. 2012;54(8):809-813. doi:10.1007/s00234-011-0977-5.

235. Chaudhry AP, Montes M, Cohn GA. Ultrastructure of cerebellar hemangioblastoma. *Cancer*. 1978;42(4):1834-1850. doi:10.1002/1097-0142(197810)42:4<1834::AID-CNCR2820420423>3.0.CO;2-Z.

236. Gruys ME, Subleski J, Lee JK, et al. Induction of transplantable mouse renal cell cancers by streptozotocin: In vivo growth, metastases, and angiogenic phenotype. *Cancer Res*. 2001;61(16):6255-6263.

237. Kang KM, Sohn CH, You SH, et al. Added value of arterial spin-labeling mr imaging for the differentiation of cerebellar hemangioblastoma from metastasis. *Am J Neuroradiol*. 2017;38(11):2052-2058. doi:10.3174/ajnr.A5363.

238. Yamashita K, Yoshiura T, Hiwatashi A, et al. Differentiating primary CNS lymphoma from glioblastoma multiforme: Assessment using arterial spin labeling, diffusion-weighted imaging, and 18F-fluorodeoxyglucose positron emission tomography. *Neuroradiology*. 2013;55(2):135-143. doi:10.1007/s00234-012-1089-6.

239. Idbaih A, Criniére E, Marie Y, et al. Gene amplification is a poor prognostic factor in anaplastic oligodendrogliomas. *Neuro Oncol*. 2008;10(4):540-547. doi:10.1215/15228517-2008-022.

240. Ata ES, Turgut M, Eraslan C, Dayanir YÖ. Comparison between dynamic susceptibility contrast magnetic resonance imaging and arterial spin labeling techniques in distinguishing malignant from benign brain tumors. *Eur J Radiol*. 2016;85(9):1545-1553. doi:10.1016/j.ejrad.2016.05.015.

241. Weber MA, Günther M, Lichy MP, et al. Comparison of Arterial Spin-Labeling Techniques and Dynamic Susceptibility-Weighted Contrast-Enhanced MRI in Perfusion Imaging of Normal Brain Tissue. *Invest Radiol*. 2003;38(11):712-718. doi:10.1097/01.rli.0000084890.57197.54.

242. Petr J, Platzek I, Seidlitz A, et al. Early and late effects of radiochemotherapy on cerebral blood flow in glioblastoma patients measured with non-invasive

perfusion MRI. *Radiother Oncol*. 2016;118(1):24-28. doi:10.1016/j.radonc.2015.12.017.

243. Taki S, Higashi K, Oguchi M, et al. Changes in regional cerebral blood flow in irradiated regions and normal brain after stereotactic radiosurgery. *Ann Nucl Med*. 2002;16(4):273-277. doi:10.1007/BF03000106.

244. Rau MK, Braun C, Skardelly M, et al. Prognostic value of blood flow estimated by arterial spin labeling and dynamic susceptibility contrast-enhanced MR imaging in high-grade gliomas. *J Neurooncol*. 2014;120(3):557-566. doi:10.1007/s11060-014-1586-z.

245. Furtner J, Bender B, Braun C, et al. Prognostic value of blood flow measurements using arterial spin labeling in gliomas. *PLoS One*. 2014;9(6):1-6. doi:10.1371/journal.pone.0099616.

246. Varela M, Petersen ET, Golay X, Hajnal J V. Cerebral blood flow measurements in infants using look-locker arterial spin labeling. *J Magn Reson Imaging*. 2015;41(6):1591-1600. doi:10.1002/jmri.24716.

247. Yeom KW, Mitchell LA, Lober RM, et al. Arterial spin-labeled perfusion of pediatric brain tumors. *Am J Neuroradiol*. 2014;35(2):395-401. doi:10.3174/ajnr.A3670.

248. Goo HW, Ra YS. Medullary hemangioblastoma in a child with von Hippel-Lindau disease: Vascular tumor perfusion depicted by arterial spin labeling and dynamic contrast-enhanced imaging. *J Neurosurg Pediatr*. 2015;16(1):50-53. doi:10.3171/2014.12.PEDS14609.

249. Delgado AF, Delgado AF, De Luca F, Hanagandi P, Van Westen D. Arterial spin-labeling in children with brain tumor: A meta-analysis. *Am J Neuroradiol*. 2018;39(8):1543-1549. doi:10.3174/ajnr.A5727.

Appendix A: Figures

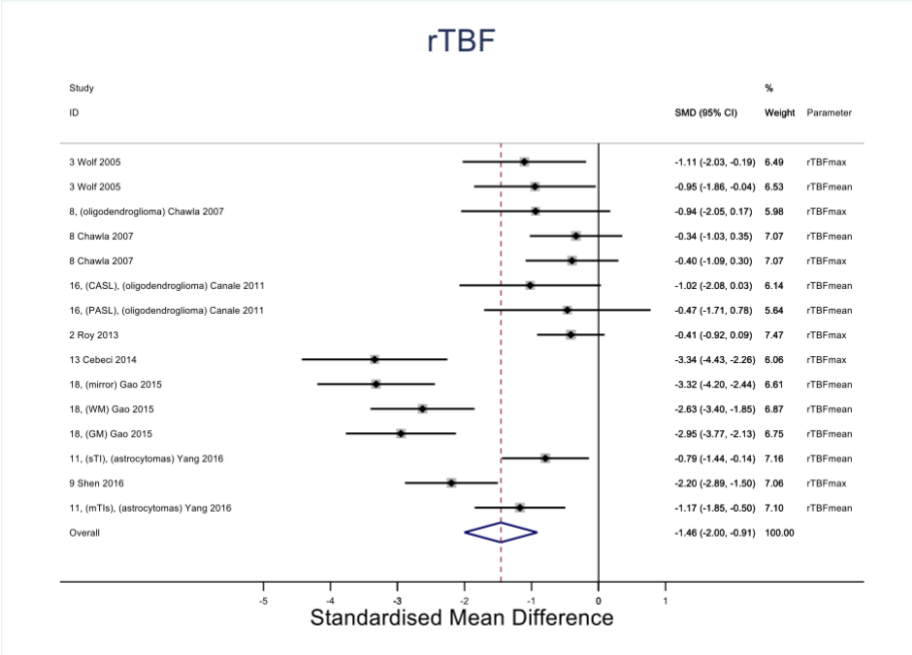
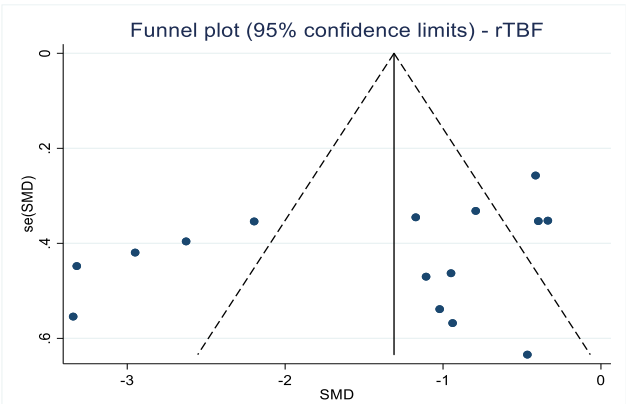
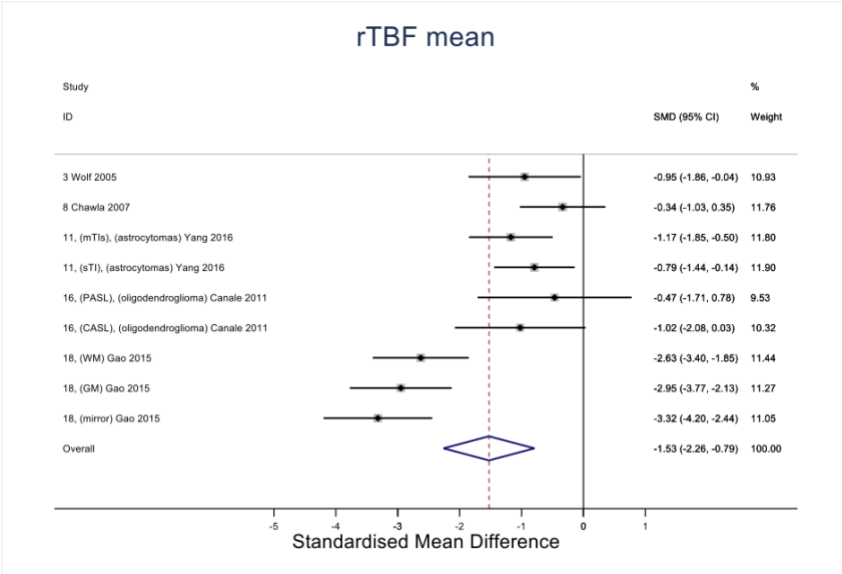


Figure 1A: rTBF for LGG patients' relative to the value for HGG patients'. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies have very different distribution (heterogeneity) around this line. The pooled effect and their 95% CI (the diamond at the bottom) express that the LGG have significantly lower rTBF than the HGG (-1.46, (-2.00, -0.91)). The funnel plot is symmetric and does not show publication bias.





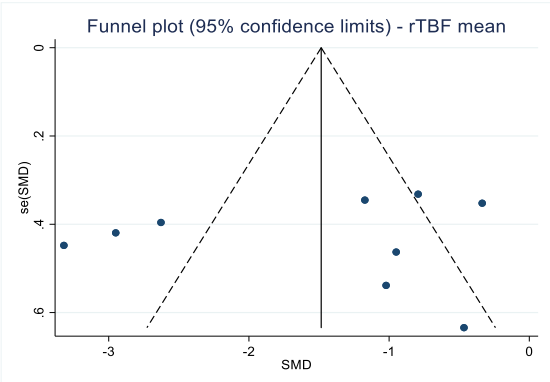
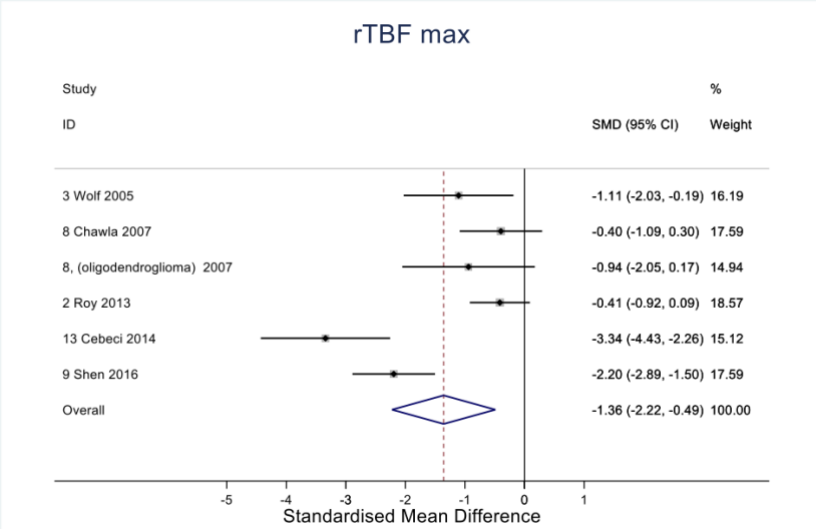


Figure 2A: rTBF-mean for LGG patients relative to the value for HGG patients. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies have very different distribution (heterogeneity) around this line. The pooled effect and their 95% CI (the diamond at the bottom) express that the LGG have significantly lower rTBFmean than the HGG (-1.53, (-2.26, -0.79)). The funnel plot is symmetric and does not show publication bias.



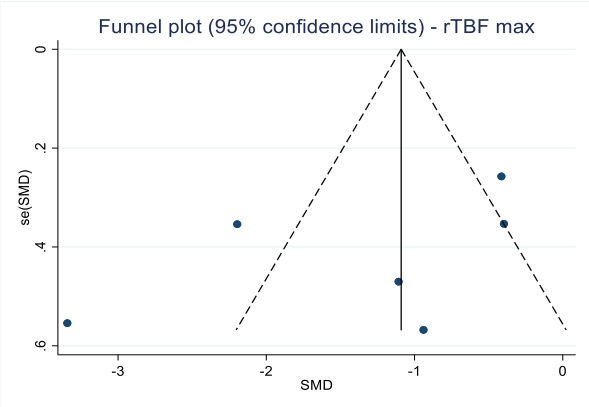


Figure 3A: rTBF-max for LGG patients' relative to the value for HGG patients'. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies have very different distribution (heterogeneity) around this line. The pooled effect and their 95% CI (the diamond at the bottom) express that the LGG have significantly lower rTBFmax than the HGG (-1.36, (-2.23, -0.49)). The funnel plot is symmetric and does not show publication bias.

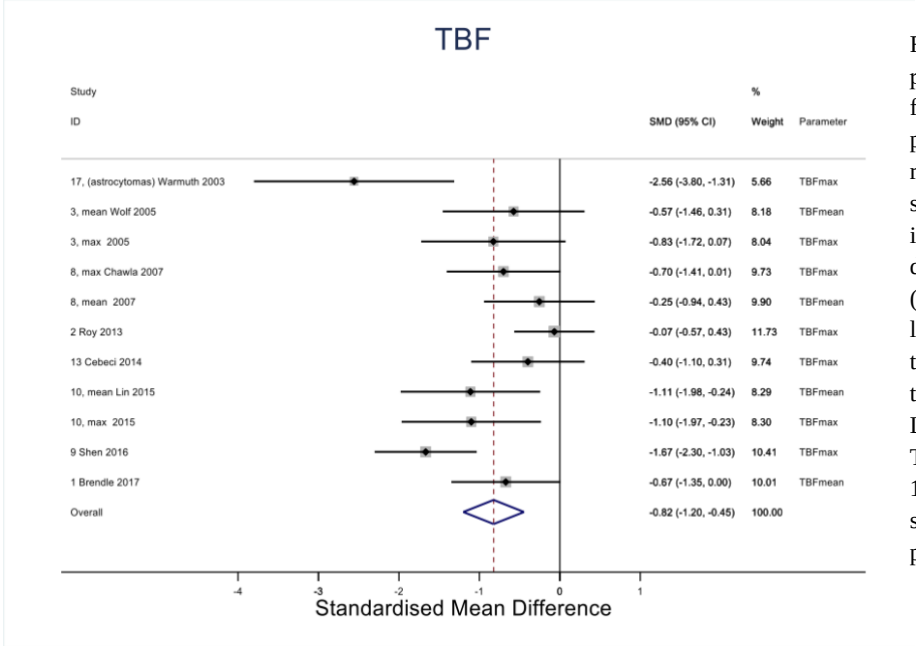
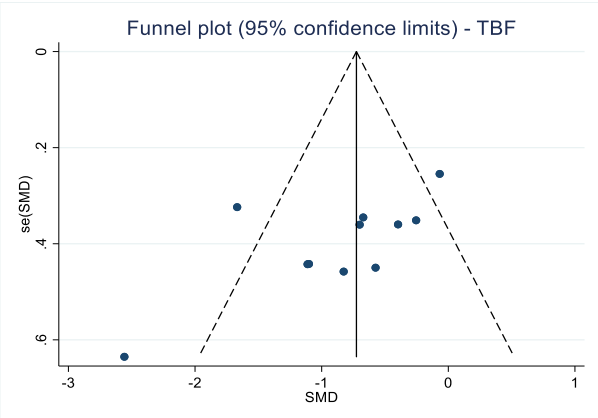


Figure 4A: TBF for LGG patients' relative to the value for HGG patients'. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies have very different distribution (heterogeneity) around this line. The pooled effect and their 95% CI (the diamond at the bottom) express that the LGG have significantly lower TBF than the HGG (-0.82, (-1.20, -0.45)). The funnel plot is symmetric and does not show publication bias.



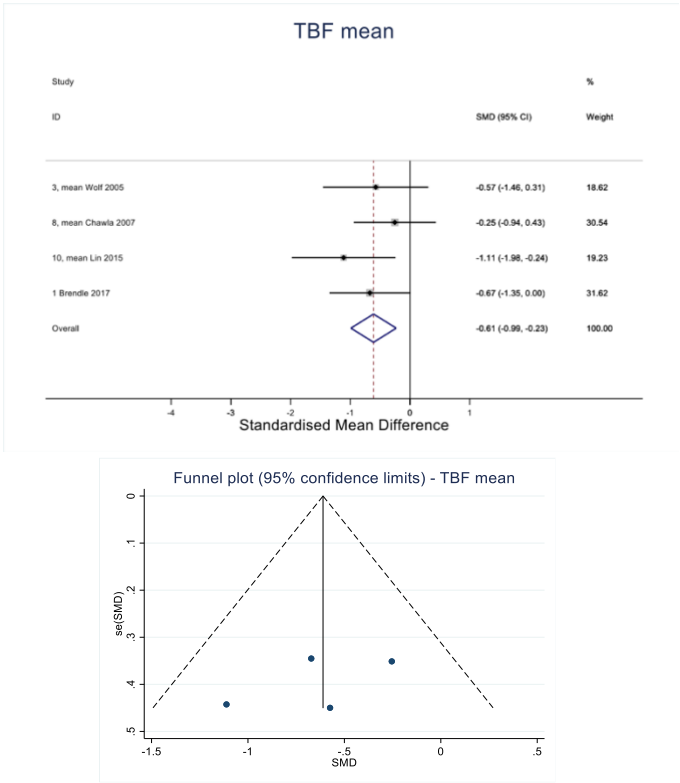


Figure 5A: TBFmean for LGG patients' relative to the value for HGG patients'. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies have small distribution around this line with small degree of heterogeneity. The pooled effect and their 95% CI (the diamond at the bottom) express that the LGG have significantly lower TBFmean than the HGG (-0.61, (-0.99, -0.23)). The funnel plot is symmetric and does not show publication bias.

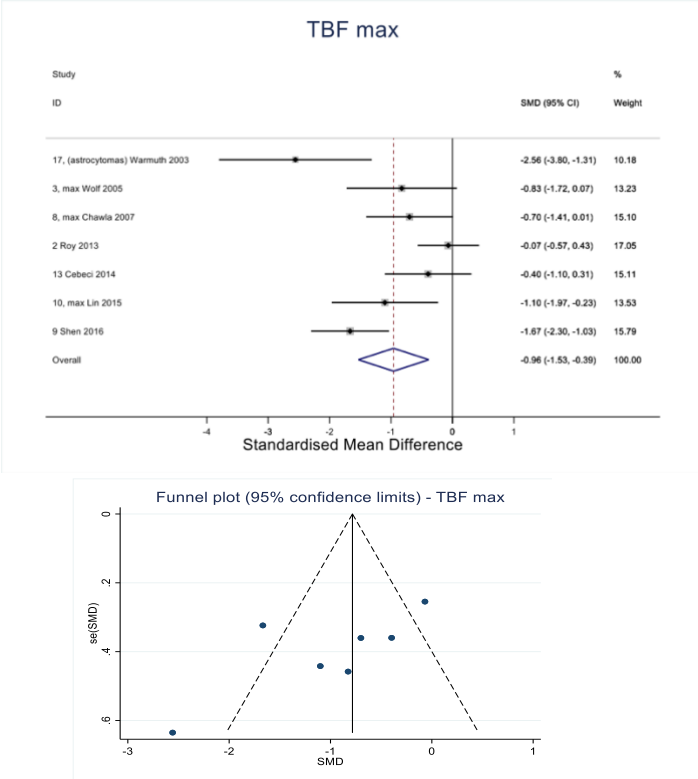


Figure 6A: TBFmax for LGG patients' relative to the value for HGG patients'. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies have very different distribution (heterogeneity) around this line. The pooled effect and their 95% CI (the diamond at the bottom) express that the LGG have significantly lower TBFmax than the HGG (-0.96, (-1.53, -0.39)). The funnel plot is symmetric and does not show publication bias.

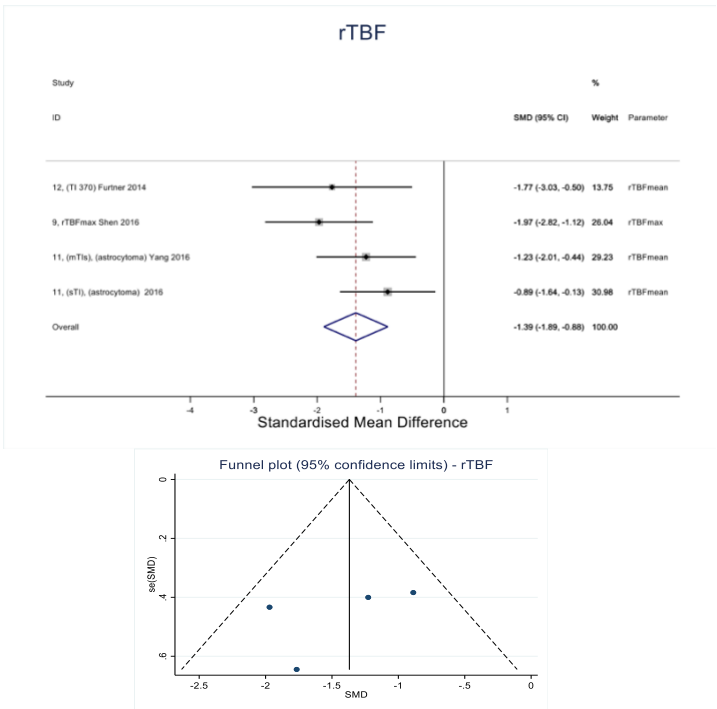


Figure 7A: rTBF for grade-II patients' relative to the value for grade-III patients'. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies have low distribution (small heterogeneity degree) around this line. The pooled effect and their 95% CI (the diamond at the bottom) express that the grade-II rTBF value about significantly lower than the that of the grade-III (-1.39, (-1.89, -0.89)). The funnel plot is symmetric and does not show publication bias.

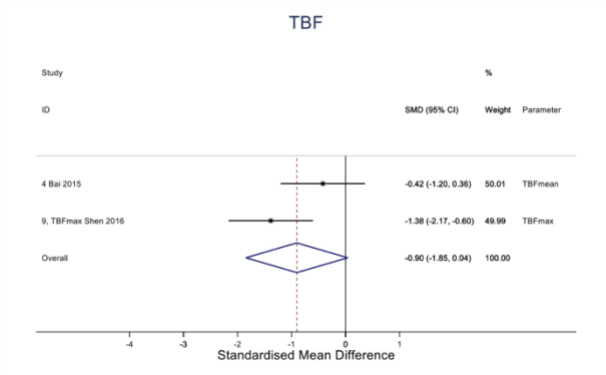
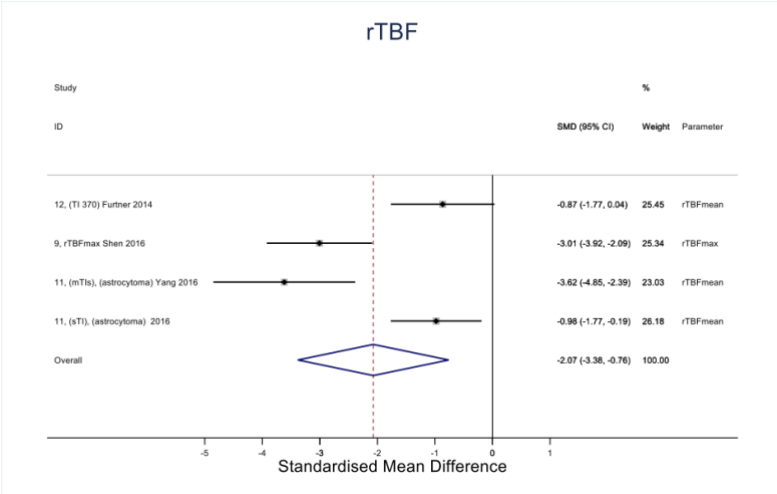


Figure 8A: TBF for grade-II patients' relative to the value for grade-III patients'. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies have moderate distribution (moderate heterogeneity degree) around this line. The pooled effect and their 95% CI (the diamond at the bottom) express that the grade-II has approximately significant lower TBF value than the grade-III (-0.90, (-1.85, 0.04)). The funnel plot cannot be produced due to the small study number.



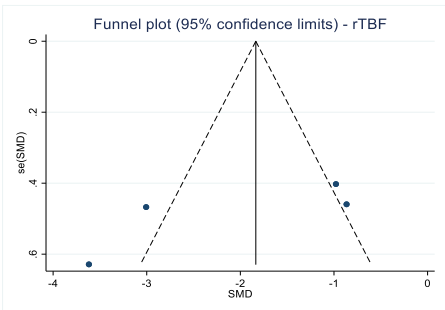


Figure 9A: rTBF for grade-II patients' relative to the value for grade-IV patients'. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies have very large distribution (heterogeneity) around this line. The pooled effect and their 95% CI (the diamond at the bottom) express that the grade-II rTBF value was significantly lower than the that of the grade-IV (-2.07, (-3.38, -0.76)). The funnel plot is symmetric and does not show publication bias.

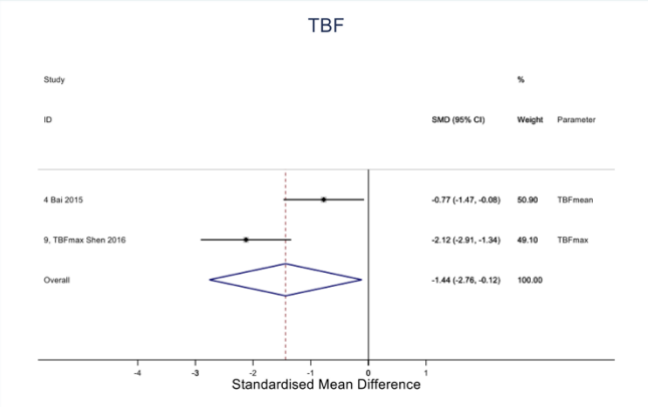


Figure 10A: TBF for grade-II patients' relative to the value for grade-IV patients'. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies represent very large distribution (heterogeneity) around this line. The pooled effect and their 95% CI (the diamond at the bottom) express that the grade-II significantly lower TBF value than the grade-IV (-1.44, (-2.76, -0.12)). The funnel plot cannot be produced due to the small study number.

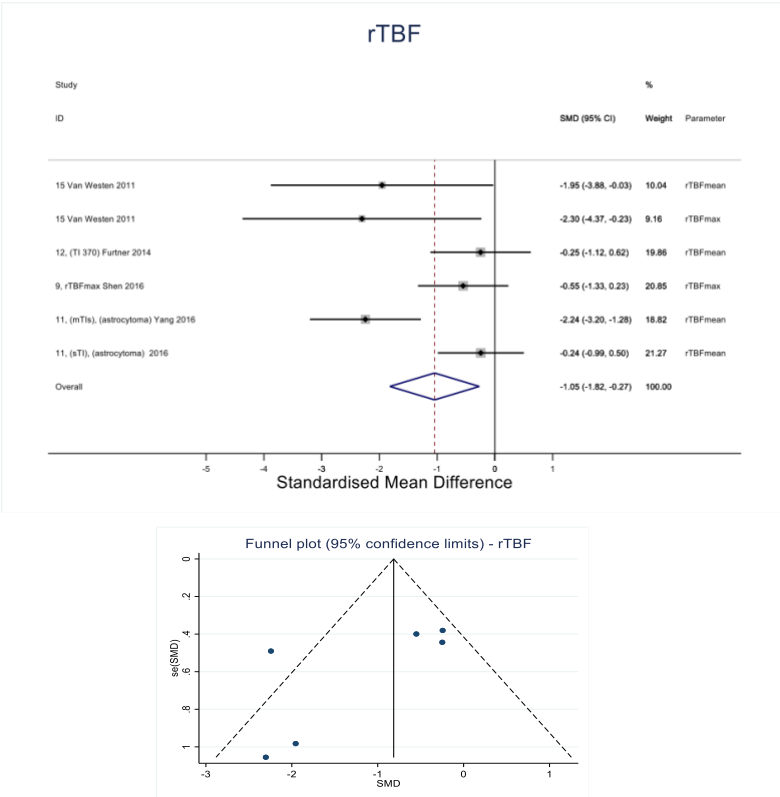
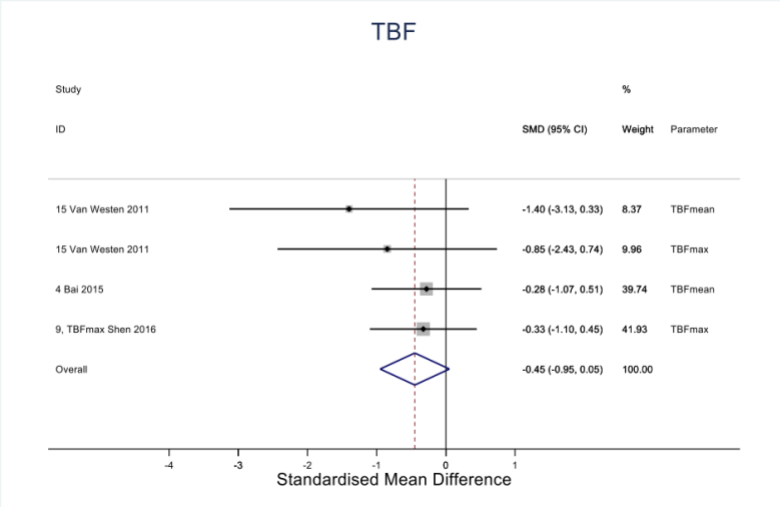


Figure 11A: rTBF for grade-III patients' relative to the value for grade-IV patients'. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies have large distribution (heterogeneity) around this line. The pooled effect and their 95% CI (the diamond at the bottom) express that the grade-III significantly has lower rTBF value than the grade-IV (-1.05, (-1.82, -0.27)). The funnel plot is symmetric and does not show publication bias.



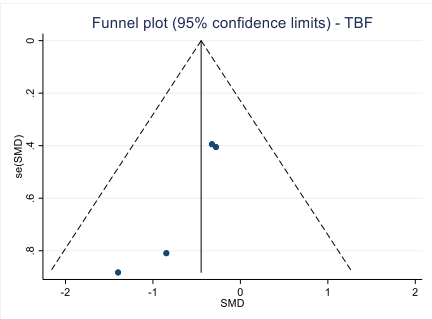


Figure 12A: TBF for grade-III patients' relative to the value for grade-IV patients'. In the forest plot, the dotted vertical line represents the pooled effect size point where the effect size in individual studies have low distribution (small heterogeneity degree) around this line. The pooled effect and their 95% CI (the diamond at the bottom) express a trend of lower rTBF value in grade-III than in grade-IV (-0.45, (-0.95, 0.05)). The funnel plot is asymmetric and does show publication bias.

Tables

Table 1A Studies performed using PCASL

Study no.	Authors	Publication year	Country of origin	Gliomas type Oligodendrogliomas/astrocytomas/ mixed	LGGs		HGGs		Histologic analysis obtained with	Study design	MRI field strength	2D/3D	Bolus width (ms)	TI/PLD (ms)	Examined perfusion metrics	Significance for differentiation between HGGs and LGGs
					Grade I	Grade II	Grade III	Grade IV								
2[130]	Roy, B. et al.	2013	India	indistinct	3	23	9 astrocytoma	29 GBM	NA	prospective	3T	3D	1450	1525	TBFmax	P = 0.78
															rTBFmax	P = 0.12
4[131]	Bai, Y. et al.	2015	USA	mixed	NA	18(13 astrocytoma, 2 oligodendroglioma)	10(5 anaplastic astrocytoma, 2 anaplastic oligodendroglioma, 3 anaplastic oligo-astrocytoma)	16 GBM	Surgical resection	prospective	3T	3D	2025	1525	TBFmean	Grade-II / III, p = 0.874
																Grade-II / IV, p = 0.023
																Grade-III/ IV, p = 0.213
9[127]	Shen, N. et al.	2016	China	mixed	25 (19 astrocytoma, 6 oligodendroglioma)		27 (10 anaplastic astrocytoma, 1 anaplastic oligodendroglioma, 16 GBM)		NA	prospective	3T	3D	1500	1525	TBFmax	P < 0.001, including sub-grading (P < 0.001)
															rTBFmax	P < 0.001, including sub-grading (P < 0.001)
10[128]	Lin, Y. et al.	2015	China	mixed	11 (7 diffuse astrocytoma, 3 oligodendroglioma, 1 capillary astrocytoma)		8 anaplastic astrocytoma	5 GBM		prospective	3T	3D	Not mentioned	1500	TBFmean	P = 0.011
															TBFmax	P = 0.002

(*)18[158]	Gao, F. et al.	2015	China	indistinct	28	21	NA	prospective	3T	2D	Not mentioned	1400	rTBFmean (WM)	P < 0.001
													rTBFmean (GM)	P < 0.001
													rTBFmean (mirror)	P < 0.001

(*) study 18 did not mentioned the used ASL labelling method

Not available (NA);

Table 2A Studies performed using CASL

Study no.	Author	Publication year	Country of origin	Gliomas type Oligodendrogliomas/astrocytomas/mixed	LGGs		HGGs		Histologic analysis obtained with	Study design	MRI field strength	2D/3D	Bolus width (ms)	TI/PLD (ms)	Examined perfusion metrics	Significance for differentiation between HGGs and LGGs
					Grade I	Grade II	Grade III	Grade IV								
3[144]	Wolf, R. et al.	2005	USA	mixed	2 (ganglioma)	5 (1 oligodendroglioma, 1 astrocytoma, 3 oligoastrocytoma)	8 (1 anaplastic oligodendroglioma, 4 anaplastic astrocytoma, 3 anaplastic oligoastrocytoma)	11 GBM	NA	NA	3T	2D	2000	1200	TBFmean	P = 0.39
															TBFmax	P = 0.04
															rTBFmean	P = 0.06
															rTBFmax	P = 0.01
8[145]	Chawla, S. et al.	2007	USA	mixed	1 gabglioma	12 (1 astrocytoma, 11 oligodendroglioma)	9 (4 astrocytoma, 5 oligodendroglioma)	13 GBM	NA	retrospective	3T	3D	2000	1200	TBFmax	P < 0.05
															TBFmean	P > 0.05
															rTBFmean	P > 0.05
															rTBFmax	P > 0.05
				Oligodendrogliomas											rTBFmax	P > 0.05
16[152]	Canale, S. et al.	2011	France	indistinct	NA	5 oligodendroglioma	11 (9 oligodendroglioma, 2 TGNM)	5 GBM	NA	retrospective	1.5T	3D	Not mentioned	1200	rTBFmean	yes

Not available (NA)

Table 3A Studies performed using PASL

Study no.	Author	Publication year	Country of origin	Gliomas type Oligodendrogliomas/astrocytomas/mixed	LGGs		HGGs		Histologic analysis obtained with	Study design	PASL approach	MRI field strength	2D/3D	Bolus width (ms)	TI/PLD (ms)	Examined perfusion metrics	Significance for differentiation between HGGs and LGGs
					Grade I	Grade II	Grade III	Grade IV									
1[84]	Brendle et al.	2017	Germany	mixed	NA	20(7 oligodendroglioma II, 14 astrocytoma II)	11(2 oligodendroglioma III, 9 astrocytoma III)	5 GBM	NA	retrospective	PICORE	3T	2D	700	1800	TBFmean	P = 0.1030
5[153]	Kim, H.S. et al.	2007	Korea	indistinct	NA	11	7	15	Surgery or stereotactic biopsy	prospective	FAIR	1.5T	2D	indistinct	1200	ROC-analysis	indistinct
6[154]	Weber, M. et al.	2006	Germany	indistinct	NA	9	11 anaplastic gliomas	35 GBM	stereotactic biopsy	prospective	Not mentioned	1.5	2D	1000	1200	ROC-analysis	indistinct
7[142]	Fudaba, H. et al.	2014	Japan	Both Mix and astrocytomas	NA	9 (3 diffuse astrocytoma, 3 oligodendroglioma and 3 oligoastrocytoma)	8 (3 anaplastic astrocytoma, 4 anaplastic oligodendroglioma, 1 oligoastrocytoma)	15 (14 GBM, 1 GBM with oligodendroglioma component)	Surgery or stereotactic biopsy	retrospective	Not mentioned	3T	2D	Not mentioned	1800	ROC-analysis	indistinct
11[129]	Yang, X. et al.	2016	China	astrocytomas	NA	15 diffuse astrocytoma	15 anaplastic astrocytoma	13 GBM	NA	prospective	FAIR	3T	3D	700	1920	rTBFmean (sTI)	HGG vs LGG, P = 0.003 II vs III, P = 0.098 II vs IV, P = 0.006

																	III vs IV, P = 0.0905
																rTBFmean (mTIs)	HGG vs LGG, P < 0.001
																	II vs III, P = 0.021
																	II vs IV, P < 0.001
																	III vs IV, P = 0.023
12[146]	Furtner, J. et al.	2014	Austria	astrocytomas	NA	7 (diffuse astrocytoma)	7 (anaplastic astrocytoma)	19 GBM	Surgery or stereotactic biopsy	prospective	PICORE	3T	2D	Not mentioned	370	rTBFmean	P = 0.003
13[43]	Cebeci, H. et al.	2014	Turkey	mixed	13 (11 oligodendroglioma, 1 disembryoblastic neuroepithelial tumour (DNET), 1 pilocytic astrocytoma)		20 (18 GBM, 1 astrocytoma, 1 gliosarcoma)		NA	retrospective	EPISTAR	3T	2D	Not mentioned	Not mentioned	TBFmax	P < 0.001
																rTBFmax	P < 0.001
14[147]	Kim, M J. et al.	2008	Korea	astrocytomas	NA	26	12 anaplastic astrocytomas	23 GBM	Surgery or stereotactic biopsy	prospective	FAIR	1.5T	2D	Not mentioned	1200	rTBFmax	P < 0.05
15[160]	Van, W. et al.	2011	Sweden	indistinct	NA	NA	3	4	Biopsy proven	NA	QUASAR	3T	2D	Not mentioned	Not mentioned	TBFmax	Just mentioned the trend of increasing the TBF from grade-III
																TBFmean	
																rTBFmax	

																rTBFmean	towards grade-IV
16[152]	Canale, S. et al.	2016	France	indistinct	NA	3 oligodendroglioma	12 (9 oligodendroglioma III, 3 TGNM)	4 GBM	NA	retrospective	Not mentioned	1.5T	2D	Not mentioned	1200	rTBFmean	Using PASL, P > 0.05 Using CASL, P < 0.05
17[148]	Warmuth, C. et al.	2003	Germany	mixed	3 (2 ganglioma, 1 pleomorphic xanthoastrocytoma)	6 (5 astrocytoma, 1 optic astrocytoma)	3 (1 anaplastic oligodendroglioma, 1 anaplastic astrocytoma, 1 astrocytoma)	7 GBM	NA	prospective	FAIR	1.5T	2D	1200	1300	TBFmax	P < 0.001
																rTBFmax	P < 0.001

Not available (NA); Proximal Inversion with Control of Off-Resonance Effects (PICORE); flow alternating inversion recovery (FAIR); quantitative STAR labeling of arterial regions (QUASAR)

Table 4A Sensitivity, specificity, negative predictive values (NPV) and positive predictive values (PPV) of published ASL-derived biomarkers cut-off values for glioma grading.

II vs III gliomas								
Author/year	Study No.	ASL parameters	Cut-off	Sensitivity	Specificity	Prevalence of grade III	PPV	NPV
Weber, M A. et al.; 2006	6	rTBFmean	1	0.92	0.33	0.55	62.66	77.14
Shen, N. et al.; 2016	9	TBFmax	43.62	1	0.69	0.31	59.12	100
Yang, X. et al.; 2016	11, (multiple TIs), (astrocytoma)	rTBFmean	2.43	1	0.51	0.5	67.11	100
	11, (single TI), (astrocytoma)	rTBFmean	1.88	0.78	0.73	0.5	74.29	76.84
II vs IV gliomas								
Author/year	Study No.	ASL parameters	Cut-off	Sensitivity	Specificity	Prevalence of grade-IV	PPV	NPV
Weber, M A. et al.; 2006	6	rTBFmean	1.6	0.94	0.78	0.80	94.32	76.97
Yang, X. et al.; 2016	11, (multiple TIs), (astrocytoma)	rTBFmean	4	1	0.87	0.46	86.96	100
	11, (single TI), (astrocytoma)	rTBFmean	3.01	0.67	0.87	0.46	81.71	75.26
III vs IV gliomas								
Author/year	Study No.	ASL parameters	Cut-off	Sensitivity	Specificity	Prevalence of grade-IV	PPV	NPV
Weber, M A. et al.; 2006	6	rTBFmean	1.4	0.97	0.5	0.76	86.06	83.97
Fudaba, H. et al.; 2014	7	rTBFmean	2.562	0.87	0.77	0.65	87.37	75.42
		rTBFmax	2.845	0.87	0.82	0.65	90.23	76.77
		rTBFmin	2.017	0.87	0.59	0.652173913	79.78038157	70.21943574
	7, (astrocytoma)	rTBFmean	1.857	0.93	0.83	0.82	96.29	71.54
	7, (astrocytoma)	rTBFmax	2.258	0.93	0.83	0.82	96.29	71.54
	7, (astrocytoma)	rTBFmin	2.164	0.79	0.83	0.824	95.645	45.478
Yang, X. et al.; 2016	11, (multiple TIs), (astrocytoma)	rTBFmean	8.55	0.77	0.73	0.464	71.19	78.55
	11, (single TI), (astrocytoma)	rTBFmean	6.64	0.46	0.73	0.464	59.62	60.93

Table 5A Sensitivity, specificity, negative predictive values (NPV) and positive predictive values (PPV) of published ASL-derived biomarkers cut-off values between HGGs and LGGs.

HGG vs LGG								
Author/ year	Study No.	ASL parameter	Cut-off	Sensitivity	Specificity	prevalence	PPV	NPV
Kim, H.S. et al.; 2007	5	rTBFmean	1.24	0.955	0.818	0.667	91.30	90.088
Fudaba, H. et al.; 2014	7	rTBFmean	2.562	0.652	0.778	0.719	88.243	46.661
	7	rTBFmax	2.845	0.609	0.778	0.719	87.516	43.776
	7	rTBFmin	2.017	0.739	0.667	0.719	85.0105	50
	7, (astrocytoma)	rTBFmean	1.8	0.824	0.667	0.85	93.343	40.076
	7, (astrocytoma)	rTBFmax	2.258	0.765	0.667	0.85	92.866	33.372
	7, (astrocytoma)	rTBFmin	1.254	0.882	0.667	0.85	93.753	49.937
Shen, N. et al.; 2016	9	TBFmax	52.21	0.889	0.826	0.5192	84.664	87.317
	9	rTBFmax	1.32	0.926	0.957	0.519	95.831	92.279
Yang, X. et al.; 2016	11, (multiple TIs), (astrocytoma)	rTBFmean	2.43	1	0.54	0.6511	80.229	100
	11, (single TI), (astrocytoma)	rTBFmean	3.01	0.6	0.88	0.651	90.323	54.098
	11, (bolus arrival time (BAT)), (astrocytoma)		0.97	0.71	0.88	0.651	91.697	61.914
Furtner, J. et al.; 2014	12		1.48	0.85	1	0.788	100	64.220
Cebeci, H. et al.; 2014	13	rTBFmax	2.1	1	0.92	0.606	95.057	100
	13	rSImax	2.19	1	0.92	0.606	95.057	100
Kim, M J. et al.; 2008	14, (astrocytoma)	rTBFmax	1.28	0.829	0.962	0.5738	96.707	80.691
Canale, S. et al.; 2011	16, (oligodendroglioma)	rTBFmean	1.8	0.88	0.6	0.762	87.562	60.976

Appendix B

Table 1-B. Statistically significant histogram parameters from ASL, DSC, ADC and DCE MRI in glioma patients based on Kruskal-Wallis test

Modality	Grading (19LGGs, 12 HGGs)			Sub-grading (19-II, 6-III, 6-IV)			IDH-status (7-wt, 25-mut)			Groping (7-wt, 12-astro, 13-oligo)			Sub-groping (7-wt-GBM, 10-astro-LG, 2-astro-HG, 9-oligo-LG, 4-oligo-HG)		
	parameter	Rank	p-value	parameter	Rank	p-value	parameter	Rank	p-value	parameter	Rank	p-value	parameter	Rank	p-value
ASL	rTBF-mean	13 vs 21.6	0.011	rTBF-mean	13, 18.7, 24	0.022	aTBF & rTBF-skew	9.9 vs 18.9	0.028	aTBF & rTBF-skew	9.9, 23.3, 13.9	0.005	aTBF & rTBF-skew	9.9, 22.6, 26.5, 11.6, 19	0.012
	rTBF-SD	13.6 vs 20.6	0.040	rTBF-95tile	12.8, 20, 23.3	0.024	aTBF & rTBF-zscore	9.14 vs 19	0.015	aTBF & rTBF-zscore	9, 24.2, 13.4	0.001	aTBF & rTBF-zscore	9, 23.9, 25.5, 11.4, 17.8	0.005
	rTBF-95tile	12.8 vs 21.9	0.008	rTBF-median	13.3, 18, 23.9	0.036	rTBF-mean	26.7 vs 14.4	0.003	rTBF-mean	25.7, 9.5, 18	0.001	rTBF-mean	25.7, 8.30, 15.5, 17, 19.8	0.005
	rTBF-median	13.3 vs 21	0.020	rTBF-max	12.9, 20.5, 22.7	0.032	rTBF-95tile	24.6 vs 14.96	0.020	**rTBF-95tile**	**23.9, 10.9, 17.7**	**0.012**	rTBF-95tile	23.9, 9.5, 18, 16, 21.3	0.027
	rTBF-iqr	**13.7 vs 20.6**	**0.040**				rTBF-median	26.9 vs 14.4	0.002	rTBF-median	25.9, 9.2, 18	0.001	rTBF-median	25.9, 8, 13.5, 17.7, 19.5	0.004
	rTBF-mode	13.7 vs 20.5	0.044				rTBF-mode	24.9 vs 14.9	0.015	**rTBF-mode**	**23.9, 10.8, 17.8**	**0.012**	rTBF-mode	23.7, 9, 18, 17, 18.8	0.034
	rTBF-slope	**19.3 vs 12.5**	**0.044**				**rTBF-min**	**23.4 vs 15.3**	**0.048**	**rTBF-min**	**22.4, 11.8, 17.7**	**0.048**	rTBF-min	22, 9.6, 22.5, 19.6, 13.5	0.034
	rTBF-entropy	13.7 vs 20.6	0.040				**rTBF-max**	**23.4 vs 15.3**	**0.048**	**rTBF-max**	**22.7, 11, 17.9**	**0.030**			
	rTBF-max	12.95 vs 21.96	0.010												
	rTBF-var	**13.7 vs 20.6**	**0.040**												
DSC	rCBV-mean	12.3 vs 22.6	0.002	rCBV-mean	12.3, 19.5, 25.3	0.005	rCBF-mean	24.9 vs 14.9	0.015	rCBF-mean	23.9, 8.7, 19.8	0.001	rCBF-mean	23.9, 8, 11, 21, 16.5	0.004
	rCBV-SD	12.3 vs 22.6	0.002	rCBV-SD	12.3, 20, 24.9	0.006	rCBF-skew	9.4 vs 19	0.020	rCBF-SD	14.6, 10.7, 22.9	0.004	rCBF-SD	14.6, 9.5, 16.5, 24, 20.5	0.015

	rCBV-95tile	12.2 vs 22.8	0.002	rCBV-95tile	12, 20, 25	0.005	rCBF-median	24.4 vs 15	0.022	rCBF-95tile	22, 9, 20.5	0.002	rCBF-95tile	22,7.8, 15, 22.9, 15	0.004
	rCBV-median	13.3 vs 21.1	0.020	rCBV-median	13.3, 17.3, 24.4	0.027	rCBV-mean	26.9 vs 14.4	0.002	rCBF-kur	19.9, 20.8, 10.7	0.015	rCBF-kur	19.9, 22, 14.5, 12, 7	0.034
	rCBV-iqr	12.5 vs 22.3	0.004	rCBV-iqr	12.5, 20.5, 23.9	0.012	**rCBV-SD**	**25 vs 14.9**	**0.014**	rCBF-skew	9.4, 23.5, 13.9	0.003	rCBF-skew	9, 23.8, 22, 13, 15	0.019
	rCBV-slope	**20.7 vs 10.4**	**0.002**	rCBV-slope	20.7, 12.7, 8.4	0.007	rCBV-95tile	26 vs 14.5	0.005	rCBF-median	23.4, 8.8, 19.9	0.001	rCBF-median	23, 8, 11, 20.9, 17.5	0.007
	rCBV-entropy	11.6 vs 23.6	<0.001	rCBV-entropy	11.6, 21.5, 25.4	0.001	rCBV-median	25.6 vs 14.7	0.008	rCBF-iqr	14.9, 10, 23	0.002	rCBF-iqr	14.9, 8.7, 17.5, 23, 23.3	0.008
	rCBV-min	20 vs 11.4	0.011	rCBV-min	20, 13.3, 9.7	0.030	**rCBV-iqr**	**23.7 vs 15.2**	**0.038**	rCBF-mode	22.3, 10, 19.4	0.08	rCBF-mode	22.3, 10.7, 6.5, 21, 14.8	0.022
	rCBV-max	12 vs 23	0.001	rCBV-max	12, 19.8, 25.9	0.002	**rCBV-mode**	**24 vs 15**	**0.028**	rCBF-slope	20.9, 20.3, 10.7	0.015	rCBF-max	21, 8, 16.5, 23, 13.8	0.007
	rCBV-var	**12.3 vs 22.6**	**0.002**	rCBV-var	12.3, 20, 24.9	0.006	**rCBV-slope**	**9 vs 19.2**	**0.014**	rCBF-max	21.4, 9.7, 20	0.006	rCBF-zscore	10.7, 24.6, 20.5, 12, 14	0.013
							rCBV-entropy	27.4 vs 14.2	0.001	rCBF-zscore	10.7, 23.9, 12.8	0.002	rCBF-var	14.6, 9.5, 16.5, 24, 20.5	0.015
							rCBV-min	7.3 vs 19.6	0.003	rCBF-var	14.6, 10.7, 22.9	0.004	rCBV-mean	25.9, 7, 12.5, 17.9, 22.8	0.001
							rCBV-max	26.7 vs 14.4	0.003	rCBV-mean	25.9, 7.9, 19.4	<0.001	rCBV-SD	24.7, 9, 16, 15.9, 22	0.012
							rCBV-var	**25 vs 14.9**	**0.014**	rCBV-SD	24.7, 10, 17.7	0.005	rCBV-95tile	25.3, 8, 14.5, 16.7, 22.8	0.003
										rCBV-95tile	25.3, 9.2, 18.5	0.001	rCBV-median	24.6, 8, 9, 18.8, 21.3	0.004
										rCBV-skew	14.9, 22.7, 11.7	0.012	rCBV-iqr	23.7, 8.5, 14.5, 17, 23.5	0.008
										rCBV-median	24.6, 8.5, 19.5	<0.001	rCBV-mode	23, 9, 7, 18, 24.3	0.006
										rCBV-iqr	23.7, 9.5, 19	0.003	rCBV-slope	8, 24, 17.5, 17, 10.3	0.009

										rCBV-mode	23, 8.9, 19.9	0.001	rCBV-entropy	26, 7, 15.5, 16, 24	<0.001
										rCBV-slope	8.4, 22.9, 14.9	0.004	rCBV-min	7.3, 21.8, 16.5, 19, 13	0.023
										rCBV-entropy	26.4, 8.5, 18.5	<0.001	rCBV-max	25.9, 8, 15.5, 16, 22	0.003
										rCBV-min	7.3, 20.9, 17.4	0.009	rCBV-var	24.7, 9, 16, 15.9, 22	0.012
										rCBV-max	25.9, 9, 18	0.001			
										rCBV-var	24.7, 10, 17.8	0.005			
ADC	ADC-min	19.89 vs 11.5	0.013	**ADC-median**	**19, 17.3, 8.9**	**0.047**	ADC-mean	6.3 vs 19.9	0.001	ADC-mode	12.5, 22.3, 13	0.029	ADC-mean	6.3, 22.6, 25.5, 16.6, 14.5	0.006
	ADC-60tile	**19.2 vs 12.5**	**0.048**	ADC-min	19.9, 9.3, 13.4	0.034	ADC-SD	10 vs 18.9	0.035	ADC-mean	6.3, 23, 15.9	<0.001	ADC-SD	10, 23.9, 20.5, 13, 15	0.024
	rADC-mean	20.11 vs 11.2	0.009	**ADC-60tile**	**19.2, 16.7, 9**	**0.048**	ADC-median	6 vs 19.9	0.001	**ADC-SD**	**10, 23, 13.7**	**0.004**	ADC-median	6, 22.7, 26, 16.7, 14	0.004
	rADC-median	**20.2 vs 11**	**0.007**	**ADC-80tile**	**19.11, 17.5, 8.6**	**0.038**	**ADC-20tile**	**8.6 vs 19.3**	**0.009**	ADC-median	6, 23, 15.9	0.001	ADC-30tile	7.9, 20.7, 25, 17.6, 14.3	0.043
	rADC-min	20.3 vs 10.9	0.005	**rADC-mean**	**20.1, 14.2, 8.7**	**0.018**	**ADC-30tile**	**7.9 vs 19.4**	**0.005**	**ADC-20tile**	**8.6, 20.8, 16.9**	**0.024**	ADC-40tile	7.3, 21.6, 26, 17, 13.9	0.017
	rADC-30tile	**19.6 vs 11.9**	**0.022**	**rADC-median**	**20.2, 14.2, 8.4**	**0.014**	ADC-40tile	7.3 vs 19.6	0.003	**ADC-30tile**	**7.9, 21.4, 16.6**	**0.010**	ADC-60tile	6.3, 23, 25.3, 16, 13.9	0.003
	rADC-40tile	**20 vs 11.4**	**0.011**	rADC-min	20.3, 9.3, 12.3	0.018	ADC-60tile	6.3 vs 19.9	0.001	ADC-40tile	7.3, 22, 16	0.003	ADC-70tile	6.6, 23, 26, 15.8, 13.6	0.004
	rADC-60tile	20.26 vs 11	0.006	**rADC-30tile**	**19.6, 15, 9.3**	**0.041**	ADC-70tile	6.6 vs 19.8	0.001	ADC-60tile	6.3, 23.6, 15	<0.001	ADC-80tile	6, 23.5, 24.5, 15.8, 15	0.003
	rADC-70tile	20.21 vs 11	0.007	**rADC-40tile**	**20, 14.5, 8.7**	**0.021**	ADC-80tile	6 vs 19.96	0.001	ADC-70tile	6.6, 23.8, 15	<0.001	ADC-90tile	6.4, 23.6, 23, 15, 16.3	0.005
	rADC-80tile	19.8 vs 11.6	0.015	**rADC-60tile**	**20.3, 14.2, 8.3**	**0.012**	ADC-90tile	6.4 vs 19.9	0.001	ADC-80tile	6, 23.7, 15.5	<0.001	ADC-95tile	7.6, 23, 22, 14.9, 16.3	0.015

	rADC-90tile	19.5 vs 12	0.027	**rADC-70tile**	**20.2, 14, 8.6**	**0.015**	ADC-95tile	7.6 vs 19.5	0.004	ADC-90tile	6, 23.5, 15.5	0.001	ADC & rADC-skew	22.6, 10.4, 7, 17.9, 22.8	0.025
	rADC-95tile	**19.2 vs 12.5**	**0.048**	**rADC-90tile**	**19.5, 14.5, 10**	**0.030**	**ADC and rADC-zscore**	**23.7 vs 15.2**	**0.038**	ADC-95tile	7.6, 23, 15	0.002	ADC & rADC-zscore	22.9, 10.4, 7.5, 17.8, 22.3	0.027
							ADC-var	**10 vs 18.9**	**0.035**	ADC and rADC-skew	22.6, 9.8, 19	0.006	ADC-var	10, 23, 20.5, 13, 15	0.024
							rADC-mean	6 vs 19.96	0.001	**ADC-slope**	**22, 10.8, 18.8**	**0.022**	rADC-mean	6, 24, 23.5, 17, 10.5	0.001
							rADC-SD	10 vs 18.9	0.031	ADC and rADC-zscore	22.9, 9.9, 19	0.006	rADC-SD	9.9, 24, 19, 13.6, 14.3	0.021
							rADC-median	5.7 vs 20	0.001	**ADC-var**	**10, 23, 13.7**	**0.004**	rADC-median	5.7, 24, 25, 17, 9.8	<0.001
							rADC-10tile	**8.7 vs 19**	**0.011**	rADC-mode	12.3, 22, 13.4	0.024	rADC-20tile	7.3, 21.8, 24, 18, 11.3	0.012
							rADC-20tile	7.3 vs 19.6	0.003	rADC-mean	6, 24, 15	<0.001	rADC-30tile	6, 22, 24.5, 18, 11.3	0.005
							rADC-30tile	6 vs 19.9	0.001	**rADC-SD**	**9.9, 23, 13.8**	**0.004**	rADC-40tile	6, 23.5, 25, 17.8, 10.3	0.001
							rADC-40tile	6 vs 19.96	0.001	rADC-median	5.7, 24.5, 14.9	<0.001	rADC-60tile	5.6, 24.9, 24, 16.8, 10.3	<0.001
							rADC-60tile	5.6 vs 20	<0.001	**rADC-10tile**	**8.7, 21, 16.4**	**0.020**	rADC-70tile	5.9, 25, 23.5, 16.6, 10.3	<0.001
							rADC-70tile	5.9 vs 20	0.001	rADC-20tile	7.3, 22.2, 16	0.004	rADC-80tile	6.6, 24.4, 24, 16.33, 10.8	0.001
							rADC-80tile	6.6 vs 19.8	0.001	rADC-30tile	6, 22.7, 16	0.001	rADC-90tile	7.4, 25, 22.5, 15, 11.5	0.002
							rADC-90tile	7.4 vs 19.6	0.003	rADC-40tile	6, 23.8, 15.5	<0.001	rADC-95tile	8, 24.8, 20.5, 14.8, 12.3	0.005
							rADC-95tile	8 vs 19.4	0.006	rADC-60tile	5.6, 24.8, 14.8	<0.001	rADC-var	9.9, 24, 19, 13.6, 14.3	0.021
							rADC-var	**10 vs 18.9**	**0.041**	rADC-70tile	5.9, 24.8, 14.6	<0.001			

										rADC-80tile	6.6, 24, 14.6	<0.001			
										rADC-90tile	7.4, 24.6, 13.9	<0.001			
										rADC-95tile	8, 24, 14	0.001			
										rADC-slope	**22.4, 10.8, 18.6**	**0.019**			
										rADC-var	**9.9, 23, 13.8**	**0.004**			
DCE-mTK	**Ve-mode**	**12.8 vs 21**	**0.012**	**Ve-mode**	**12.8, 22.9, 19**	**0.033**	**rK-mean**	**10 vs 18.9**	**0.031**	Ve-mode	20, 11, 19.5	0.035	Ve-mode	20, 9.95, 16.8, 16, 27	0.022
	rVe-mode	**13.11 vs 21.5**	**0.013**	**rVe-mode**	**13, 23.7, 19.6**	**0.034**				rVe-mode	20.6, 10.5, 19.9	0.019	rVe-mode	20.6, 9.6, 15, 16, 28	0.012
	rVp-mean	13.9 vs 20.4	0.042							rVp-mean	17, 11.3, 21	0.032	rVp-mean	17, 10, 16, 18, 27.5	0.037
										rVp-median	17.4, 10, 21.7	0.010	rVp-median	17, 9, 15.5, 19, 26.8	0.020
										rVp-min	19.6, 10.2, 20.7	0.012	rVp-min	19.6, 9, 15.5, 19, 23.5	0.039
DCE-L	TC-SD	18.58 vs 11.9	0.047	**rVp-95tile**	**13.7, 24.5, 17**	**0.049**	**rK-mean**	**9.7 vs 18.96**	**0.025**	**rK-mean**	**9.7, 15, 21**	**0.028**	**rK-iqr**	**9.3, 15.9, 12, 19.6, 26**	**0.046**
	TC-95tile	18.7 vs 11.8	0.041	**rVp-slope**	**19, 8, 16.6**	**0.04**	**rK-SD**	**10 vs 18.9**	**0.035**	**rK-SD**	**10, 15.5, 20.9**	**0.046**	rFb-mean	14, 10, 14, 20, 29.3	0.008
	TC-slope	**13.3 vs 20**	**0.035**	rVp-max	13.89, 25.5, 15.9	0.030	**rK-95tile**	**10 vs 18.9**	**0.035**	**rK-iqr**	**9.3, 15.3, 21.5**	**0.017**	rFb-SD	13, 10.9, 17.5, 20, 27.8	0.021
	TC-var	**18.6 vs 11.9**	**0.047**	rFb-max	15.2, 25, 12.7	0.040	**rK-median**	**10.6 vs 18.7**	**0.048**	**rK-var**	**10, 15.5, 20.9**	**0.046**	rFb-95tile	13, 10.7, 17.5, 20, 28.5	0.014
	rKep-max	19.4 vs 12.2	0.033	rTC-mode	14.3, 13.3, 25.3	0.019	**rK-iqr**	**9.3 vs 19**	**0.017**	rFb-mean	14, 11, 22.9	0.005	**rFb-median**	**15, 12, 13.5, 17, 30**	**0.028**
	rVp-mean	13.7 vs 20.5	0.044	rTC-min	14.5, 13.2, 24.9	0.027	rK-slope	23.7 vs 15.2	0.038	rFb-SD	13, 12, 22.5	0.012	rFb-iqr	13.3, 10.7, 14.5, 20, 28.8	0.011

	rVp-95tile	13.7 vs 20.5	0.044				**rK-var**	**10 vs 18.9**	**0.035**	rFb-95tile	13, 11.8, 22.6	0.009	rFb-slope	20, 22, 15, 13, 5	0.020
	rVp-entropy	13.7 vs 20.6	0.040				**rPS-mean**	**10.3 vs 18.8**	**0.038**	rFb-iqr	13.3, 11, 23	0.005	rFb-max	12, 11.5, 19.5, 19.6, 27.8	0.026
	rFb-entropy	**13.7 vs 20.6**	**0.040**				**rPS-median**	**10 vs 18.9**	**0.035**	rFb-slope	20, 20.8, 10.6	0.013	rFb-var	13, 10.9, 17.5, 20, 27.8	0.021
	rTC-SD	**19.5 vs 12**	**0.027**				**rTC-mode**	**24.6 vs 14.96**	**0.020**	rFb-min	12.9, 12.9, 21.8	0.032	**rPS-mean**	**10.3, 19.3, 4, 17, 25**	**0.028**
	rTC-95tile	19.6 vs 11.9	0.022				rTC-min	24 vs 15	0.028	rFb-max	12.4, 12.8, 22	0.021	**rPS-median**	**10, 18.4, 5, 17.7, 26**	**0.028**
	rTC-kur	13.5 vs 20.9	0.027							rFb-var	13, 12, 22.5	0.012	rTC-mean	17, 21.5, 20, 14.7, 4.8	0.044
	rTC-skew	13.7 vs 20.6	0.04							rTC-mean	17, 21.3, 11.6	0.036			
	rTC-iqr	**19.6 vs 12**	**0.025**							rTC-mode	24.6, 15.6, 13	0.029			
	rTC-slope	**13.3 vs 21.1**	**0.02**							rTC-min	23.9, 17, 12	0.027			
	rTC-entropy	19.3 vs 12.5	0.044												
	rTC-var	**19.5 vs 12**	**0.027**												
The bolded variables even they revealed statistically significant differences they could not predict the assigned classes (p-value>0.05), using the multinomial logistic regression.															

Table 2-B. The diagnostic performance of significant features from ASL, DSC, DCE-mTK, DCE-L&L, ADC for glioma grading.

MRI-method	Features	Prediction performance	Grading		Internal validation LOOCV	
			LGGs	HGGs	LGGs	HGGs
ASL	rTBF_mean	sensitivity	79%	62%	79%	59%
		specificity	62%	79%	59%	79%
		F	77%	64%	76%	62%
		Accuracy	72%		72%	
		Kappa	41%		41%	
	rTBF_SD	sensitivity	84%	62%	89%	26%
		specificity	62%	84%	26%	89%
		F	80%	67%	74%	36%
		Accuracy	75%		64%	
		Kappa	47%		17%	
	rTBF_95tile	sensitivity	84%	62%	84%	59%
		specificity	62%	84%	59%	84%
		F	80%	67%	79%	65%
		Accuracy	75%		74%	
		Kappa	47%		44%	
	rTBF_median	sensitivity	79%	62%	79%	54%
		specificity	62%	79%	54%	79%
		F	77%	64%	75%	85%
		Accuracy	72%		67%	
		Kappa	41%		34%	
	rTBF_iqr	sensitivity	84%	31%	86%	21%
		specificity	31%	84%	21%	86%
		F	73%	40%	72%	29%
		Accuracy	63%		59%	
		Kappa	16%		7%	
	rTBF_mode	sensitivity	79%	69 %	79%	67%
		specificity	69%	79%	67%	79%
		F	77%	69%	78%	68%
		Accuracy	75%		74%	
		Kappa	48%		46%	
	rTBF_slope	sensitivity	84%	62%	81%	36%

		specificity	62%	84%	36%	81%
		F	80%	67%	72%	44%
		Accuracy	75%		63%	
		Kappa	47%		18%	
	rTBF_entropy	sensitivity	84%	62%	86%	41%
		specificity	62%	84%	41%	86%
		F	80%	67%	76%	51%
		Accuracy	75%		68%	
		Kappa	47%		29%	
	rTBF_max	**sensitivity**	**84%**	**62%**	**84%**	**62%**
		specificity	**62%**	**84%**	**62%**	**84%**
		F	**80%**	**67%**	**80%**	**67%**
		Accuracy	**75%**		**75%**	
		Kappa	**47%**		**47%**	
	rTBF_var	sensitivity	84%	39%	89%	15%
		specificity	39%	84%	15%	89%
		F	74%	45%	72%	24%
		Accuracy	66%		59%	
		Kappa	24%		5%	
DSC	rCBV_mean	sensitivity	95%	69%	91%	56%
		specificity	69%	95%	56%	91%
		F	88%	78%	83%	67%
		Accuracy	84%		77%	
		Kappa	66%		50%	
	rCBV_SD	sensitivity	84%	69%	88%	54%
		specificity	69%	84%	54%	88%
		F	82%	72%	80%	63%
		Accuracy	78%		74%	
		Kappa	54%		44%	
	rCBV_95tile	sensitivity	89%	77%	89%	69%
		specificity	77%	89%	69%	89%
		F	87%	80%	85%	75%
		Accuracy	84%		81%	
		Kappa	67%		60%	
	rCBV_median	sensitivity	84%	54%	81%	49%
		specificity	54%	84%	49%	81%

		F	78%	61%	75%	55%
		Accuracy	72%		68%	
		Kappa	40%		31%	
	rCBV_iqr	sensitivity	89%	69%	89%	62%
		specificity	69%	89%	62%	89%
		F	85%	75%	80%	70%
		Accuracy	81%		78%	
		Kappa	60%		53%	
	rCBV_slope	sensitivity	79%	69%	79%	67%
		specificity	69%	79%	67%	79%
		F	79%	69%	78%	68%
		Accuracy	75%		74%	
		Kappa	48%		46%	
	rCBV_entropy	**sensitivity**	**89%**	**77%**	**89%**	**77%**
		specificity	**77%**	**89%**	**77%**	**89%**
		F	**87%**	**80%**	**87%**	**80%**
		Accuracy	**84%**		**84%**	
		Kappa	**67%**		**67%**	
	rCBV_min	sensitivity	79%	69%	86%	41%
		specificity	69%	79%	41%	86%
		F	79%	69%	76%	51%
		Accuracy	75%		68%	
		Kappa	48%		29%	
	rCBV_max	sensitivity	89%	77%	89%	77%
		specificity	77%	89%	77%	89%
		F	87%	80%	87%	80%
		Accuracy	84%		84%	
		Kappa	67%		67%	
	rCBV_var	sensitivity	84%	69%	89%	51%
		specificity	69%	84%	51%	89%
		F	82%	72%	80%	62%
		Accuracy	78%		74%	
		Kappa	54%		43%	
DCE_mTK	**Ve_mode**	**sensitivity**	**84%**	**62%**	**84%**	**54%**
		specificity	**62%**	**84%**	**54%**	**84%**
		F	**80%**	**67%**	**78%**	**61%**

		Accuracy	**75%**		**72%**	
		Kappa	**47%**		**40%**	
	rVe_mode	sensitivity	89%	54%	93%	36%
		specificity	54%	89%	36%	93%
		F	81%	64%	79%	49%
		Accuracy	75%		70%	
		Kappa	46%		32%	
	rVp_mean	sensitivity	79%	46%	75%	33%
		specificity	46%	79%	33%	75%
		F	73%	52%	68%	39%
		Accuracy	66%		58%	
		Kappa	26%		9%	
DCE_L&L	TC_SD	sensitivity	79%	46%	75%	38%
		specificity	46%	79%	38%	75%
		F	73%	52%	69%	44%
		Accuracy	66%		60%	
		Kappa	26%		14%	
	TC_95tile	sensitivity	79%	38%	79%	38%
		specificity	38%	79%	38%	79%
		F	71%	45%	71%	45%
		Accuracy	63%		63%	
		Kappa	18%		18%	
	TC_slope	sensitivity	84%	38%	81%	33%
		specificity	38%	84%	33%	81%
		F	74%	48%	71%	41%
		Accuracy	66%		61%	
		Kappa	24%		15%	
	TC_var	sensitivity	74%	46%	68%	44%
		specificity	46%	74%	44%	68%
		F	70%	50%	66%	46%
		Accuracy	63%		58%	
		Kappa	20%		12%	
	rKep_max	sensitivity	100%	0%	100%	0%
		specificity	0%	100%	0%	100%
		F	75%	0%	75%	0%
		Accuracy	59%		59%	

		Kappa	0%		0%	
	rVp_mean	**sensitivity**	**89%**	**46%**	**86%**	**41%**
		specificity	**46%**	**89%**	**41%**	**86%**
		F	**79%**	**57%**	**76%**	**51%**
		Accuracy	**72%**		**68%**	
		Kappa	**38%**		**29%**	
	rVp_95tile	sensitivity	79%	31%	79%	31%
		specificity	31%	79%	31%	79%
		F	70%	38%	70%	38%
		Accuracy	59%		59%	
		Kappa	10%		2%	
	rVp_entropy	sensitivity	79%	23%	72%	5%
		specificity	23%	79%	5%	72%
		F	68%	30%	61%	7%
		Accuracy	56%		45%	
		Kappa	22%		2%	
	rFb_entropy	sensitivity	84%	38%	86%	21%
		specificity	38%	84%	21%	86%
		F	74%	48%	72%	29%
		Accuracy	66%		59%	
		Kappa	24%		7%	
	rTC_SD	sensitivity	79%	31%	79%	21%
		specificity	31%	79%	21%	79%
		F	70%	38%	68%	27%
		Accuracy	5%		55%	
		Kappa	10%		1%	
	rTC_95tile	sensitivity	84%	31%	84%	21%
		specificity	31%	84%	21%	84%
		F	73%	40%	71%	29%
		Accuracy	62%		58%	
		Kappa	16%		5%	
	rTC_kur	sensitivity	84%	46%	84%	36%
		specificity	46%	84%	36%	84%
		F	76%	55%	74%	45%
		Accuracy	69%		65%	
		Kappa	32%		22%	

	rTC_skew	sensitivity	84%	46%	81%	41%
		specificity	46%	84%	41%	81%
		F	76%	55%	73%	48%
		Accuracy	69%		65%	
		Kappa	32%		23%	
	rTC_iqr	sensitivity	89%	38%	82%	38%
		specificity	38%	89%	38%	82%
		F	77%	50%	73%	47%
		Accuracy	69%		65%	
		Kappa	31%		22%	
	rTC_slope	sensitivity	89%	15%	86%	15%
		specificity	15%	89%	15%	86%
		F	72%	24%	71%	23%
		Accuracy	59%		57%	
		Kappa	5%		1%	
	rTC_entropy	sensitivity	84%	46%	88%	33%
		specificity	46%	84%	33%	88%
		F	76%	55%	75%	44%
		Accuracy	69%		66%	
		Kappa	32%		23%	
	rTC_var	sensitivity	79%	31%	75%	28%
		specificity	31%	79%	28%	75%
		F	70%	38%	67%	34%
		Accuracy	59%		56%	
		Kappa	10%		4%	
ADC	ADC_min	sensitivity	89%	46%	91%	46%
		specificity	46%	89%	46%	91%
		F	79%	57%	80%	58%
		Accuracy	72%		73%	
		Kappa	38%		40%	
	ADC_60tile	sensitivity	84%	62%	89%	38%
		specificity	62%	84%	38%	89%
		F	80%	62%	77%	50%
		Accuracy	75%		69%	
		Kappa	47%		30%	
	rADC_mean	sensitivity	89%	69%	89%	44%

		specificity	69%	89%	44%	89%
		F	85%	75%	78%	55%
		Accuracy	81%		71%	
		Kappa	60%		35%	
	rADC_median	sensitivity	84%	69%	89%	49%
		specificity	69%	84%	49%	89%
		F	82%	72%	80%	59%
		Accuracy	78%		73%	
		Kappa	54		40%	
	rADC_min	sensitivity	84%	46%	88%	44%
		specificity	46%	84%	44%	88%
		F	76%	55%	78%	54%
		Accuracy	69%		70%	
		Kappa	32%		33%	
	rADC_30tile	sensitivity	79%	69%	86%	38%
		specificity	69%	79%	38%	86%
		F	79%	69%	75%	48%
		Accuracy	75%		67%	
		Kappa	48%		26%	
	rADC_40tile	sensitivity	84%	69%	86%	49%
		specificity	69%	84%	49%	86%
		F	82%	72%	78%	58%
		Accuracy	78%		71%	
		Kappa	45%		36%	
	rADC_60tile	sensitivity	84%	69%	86%	49%
		specificity	69%	84%	49%	86%
		F	82%	72%	78%	58%
		Accuracy	78%		71%	
		Kappa	54%		36%	
	rADC_70tile	sensitivity	79%	77%	86%	54%
		specificity	77%	79%	54%	86%
		F	81%	74%	79%	62%
		Accuracy	78%		73%	
		Kappa	55%		41%	
	rADC_80tile	**sensitivity**	**84%**	**77%**	**88%**	**51%**
		specificity	**77%**	**84%**	**51%**	**88%**

		F	**84%**	**77%**	**79%**	**61%**
		Accuracy	**81%**		**73%**	
		Kappa	**61%**		**41%**	
	rADC_90tile	sensitivity	84%	46%	84%	33%
		specificity	46%	84%	33%	84%
		F	76%	55%	73%	43%
		Accuracy	69%		64%	
		Kappa	32%		19%	
	rADC_95tile	sensitivity	84%	46%	88%	33%
		specificity	46%	84%	33%	88%
		F	76%	55%	75%	44%
		Accuracy	69%		66%	
		Kappa	32%		23%	

Table 3-B. The diagnostic performance of significant features from ASL, DSC, DCE-mTK, DCE-L&L, ADC for glioma sub-grading

MRI-method	Features	Prediction performance	Sub-grading			Internal validation LOOCV		
			II	III	IV	II	III	IV
ASL	**rTBF_mean**	**sensitivity**	**95%**	**0%**	**57%**	**93%**	**0%**	**43%**
		specificity	**46%**	**100%**	**88%**	**33%**	**100%**	**89%**
		F	**81%**	**0%**	**57%**	**78%**	**0%**	**47%**
		Accuracy	**69%**			**65%**		
		Kappa	**36%**			**25%**		
	rTBF_95tile	sensitivity	89%	0%	57%	93%	0%	29%
		specificity	54%	100%	80%	38%	99%	84%
		F	81%	0%	50%	79%	0%	31%
		Accuracy	66%			61%		
		Kappa	33%			20%		
	rTBF_median	sensitivity	89%	0%	57%	91%	0%	43%
		specificity	46%	100%	84%	33%	100%	88%
		F	79%	0%	53%	77%	0%	46%
		Accuracy	66%			64%		
		Kappa	31%			24%		
	rTBF_max	sensitivity	95%	0%	43%	89%	0%	0%
		specificity	46%	100%	84%	28%	95%	83%
		F	82%	0%	43%	75%	0%	0%
		Accuracy	67%			53%		
		Kappa	30%			1%		
DSC	rCBV_mean	sensitivity	95%	0%	57%	95%	0%	29%
		specificity	46%	100%	88%	31%	100%	88%
		F	82%	0%	57%	78%	0%	33%
		Accuracy	69%			63%		
		Kappa	36%			19%		
	rCBV_SD	sensitivity	95%	0%	67%	96%	0%	0%
		specificity	54%	100%	84%	36%	90%	89%
		F	84%	0%	53%	80%	0%	0%
		Accuracy	69%			57%		
		Kappa	38%			9%		
	rCBV_95tile	sensitivity	95%	0%	71%	96%	0%	19%

		specificity	62%	100%	84%	31%	100%	87%
		F	86%	0%	63%	79%	0%	23%
		Accuracy	72%			61%		
		Kappa	45%			16%		
	rCBV_median	sensitivity	95%	0%	29%	95%	0%	19%
		specificity	31%	100%	88%	23%	100%	89%
		F	78%	0%	33%	77%	0%	24%
		Accuracy	63%			60%		
		Kappa	19%			13%		
	rCBV_iqr	sensitivity	95%	33%	43%	93%	0%	0%
		specificity	38%	100%	96%	31%	92%	87%
		F	80%	50%	55%	77%	0%	0%
		Accuracy	72%			55%		
		Kappa	41%			5%		
	rCBV_slope	sensitivity	100%	0%	29%	100%	0%	0%
		specificity	38%	100%	88%	23%	100%	88%
		F	83%	0%	33%	79%	0%	0%
		Accuracy	66%			59%		
		Kappa	26%			8%		
	rCBV_entropy	**sensitivity**	**100%**	**0%**	**86%**	**96%**	**0%**	**62%**
		specificity	**69%**	**100%**	**88%**	**59%**	**97%**	**87%**
		F	**90%**	**0%**	**75%**	**86%**	**0%**	**59%**
		Accuracy	**78%**			**71%**		
		Kappa	**57%**			**42%**		
	rCBV_min	sensitivity	84%	0%	71%	89%	0%	48%
		specificity	54%	100%	80%	36%	100%	87%
		F	78%	0%	59%	77%	0%	49%
		Accuracy	66%			64%		
		Kappa	34%			25%		
	rCBV_max	sensitivity	100%	0%	71%	93%	0%	43%
		specificity	62%	100%	88%	46%	100%	83%
		F	88%	0%	67%	81%	0%	42%
		Accuracy	75%			65%		
		Kappa	5%			28%		
	rCBV_var	sensitivity	95%	50%	57%	96%	0%	0%
		specificity	54%	100%	96%	36%	90%	89%

		F	84%	67%	67%	80%	0%	0%
		Accuracy	78%			57%		
		Kappa	57%			9%		
DCE_mTK	**Ve_mode**	**sensitivity**	**95%**	**0%**	**14%**	**95%**	**0%**	**0%**
		specificity	**38%**	**100%**	**80%**	**21%**	**96%**	**89%**
		F	**80%**	**0%**	**15%**	**76%**	**0%**	**0%**
		Accuracy	**59%**			**56%**		
		Kappa	**15%**			**3%**		
	rVe_mode	sensitivity	95%	17%	0%	93%	0%	0%
		specificity	31%	96%	88%	26%	96%	85%
		F	78%	25%	0%	76%	0%	0%
		Accuracy	59%			55%		
		Kappa	13%			3%		
DCE_L&L	rVp_95tile	sensitivity	89%	67%	0%	93%	50%	0%
		specificity	31%	92%	100%	23%	95%	100%
		F	76%	67%	0%	76%	58%	0%
		Accuracy	66%			65%		
		Kappa	29%			23%		
	rVp_slope	sensitivity	89%	67%	0%	91%	50%	0%
		specificity	31%	92%	100%	23%	94%	100%
		F	76%	67%	0%	75%	56%	0%
		Accuracy	66%			64%		
		Kappa	29%			23%		
	rVp_max	**sensitivity**	**95%**	**67%**	**0%**	**95%**	**50%**	**0%**
		specificity	**31%**	**96%**	**100%**	**23%**	**96%**	**100%**
		F	**78%**	**73%**	**0%**	**77%**	**60%**	**0%**
		Accuracy	**67%**			**66%**		
		Kappa	**33%**			**25%**		
	rFb_max	sensitivity	95%	17%	0%	95%	11%	0%
		specificity	8%	96%	100%	5%	96%	100%
		F	73%	25%	0%	73%	17%	0%
		Accuracy	59%			58%		
		Kappa	5%			3%		
	rTC_mode	sensitivity	95%	0%	29%	93%	0%	9%
		specificity	31%	100%	88%	15%	100%	89%
		F	78%	0%	33%	74%	0%	13%

		Accuracy	63%			57%		
		Kappa	19%			4%		
	rTC_min	sensitivity	95%	0%	29%	96%	0%	19%
		specificity	23%	100%	92%	15%	100%	95%
		F	77%	0%	36%	76%	0%	28%
		Accuracy	63%			61%		
		Kappa	17%			11%		
ADC	ADC_median	sensitivity	89%	0%	43%	93%	0%	29%
		specificity	31%	100%	88%	23%	100%	91%
		F	76%	0%	46%	76%	0%	35%
		Accuracy	63%			62%		
		Kappa	21%			16%		
	ADC_min	**sensitivity**	**100%**	**67%**	**0%**	**96%**	**67%**	**0%**
		specificity	**46%**	**92%**	**100%**	**44%**	**91%**	**100%**
		F	**84%**	**67%**	**0%**	**82%**	**65%**	**0%**
		Accuracy	**72%**			**70%**		
		Kappa	**42%**			**38%**		
	ADC_60tile	sensitivity	89%	0%	43%	93%	0%	29%
		specificity	31%	100%	88%	23%	100%	91%
		F	76%	0%	46%	76%	0%	35%
		Accuracy	63%			61%		
		Kappa	21%			16%		
	ADC_80tile	sensitivity	95%	0%	43%	93%	0%	29%
		specificity	31%	100%	92%	23%	100%	91%
		F	78%	0%	50%	76%	0%	35%
		Accuracy	66%			61%		
		Kappa	26%			16%		
	rADC_mean	sensitivity	95%	0%	57%	96%	0%	29%
		specificity	38%	100%	92%	21%	100%	95%
		F	80%	0%	62%	77%	0%	39%
		Accuracy	69%			64%		
		Kappa	34%			18%		
	rADC_median	sensitivity	100%	0%	43%	100%	0%	29%
		specificity	31%	100%	96%	21%	100%	97%
		F	81%	0%	55%	79%	0%	41%
		Accuracy	69%			66%		

		Kappa	31%			21%		
	rADC_min	sensitivity	100%	67%	0%	96%	44%	0%
		specificity	38%	96%	100%	28%	96%	97%
		F	83%	73%	0%	79%	55%	0%
		Accuracy	72%			66%		
		Kappa	40%			25%		
	rADC_30tile	sensitivity	100%	0%	71%	100%	0%	48%
		specificity	54%	100%	92%	36%	100%	95%
		F	86%	0%	71%	82%	0%	57%
		Accuracy	75%			70%		
		Kappa	49%			34%		
	rADC_40tile	sensitivity	100%	0%	71%	93%	0%	38%
		specificity	54%	100%	92%	31%	100%	89%
		F	86%	0%	71%	77%	0%	43%
		Accuracy	75%			64%		
		Kappa	49%			22%		
	rADC_60tile	sensitivity	95%	0%	57%	96%	0%	29%
		specificity	38%	100%	92%	21%	100%	95%
		F	80%	0%	62%	77%	0%	39%
		Accuracy	69%			64%		
		Kappa	34%			18%		
	rADC_70tile	sensitivity	95%	0%	57%	96%	0%	29%
		specificity	38%	100%	92%	21%	100%	95%
		F	80%	0%	62%	77%	0%	39%
		Accuracy	69%			64%		
		Kappa	34%			18%		
	rADC_90tile	sensitivity	95%	0%	57%	96%	0%	29%
		specificity	38%	100%	92%	21%	100%	95%
		F	80%	0%	62%	77%	0%	39%
		Accuracy	69%			64%		
		Kappa	34%			18%		

Table 4-B. The diagnostic performance of significant features from ASL, DSC, DCE-mTK, DCE-L&L, ADC for IDH-status prediction in gliomas.

MRI-method	Features	Prediction performance	IDH_status		Internal validation LOOCV	
			wt	mut	wt	mut
ASL	rTBF_skew	sensitivity	43%	96%	14%	95%
		specificity	96%	43%	95%	14%
		F	55%	91%	21%	87%
		Accuracy	84%		77%	
		Kappa	46%		12%	
	rTBF_zscore	sensitivity	57%	96%	5%	96%
		specificity	96%	57%	96%	5%
		F	67%	92%	80%	86%
		Accuracy	88%		76%	
		Kappa	59%		1%	
	rTBF_mean	**sensitivity**	**43%**	**88%**	**29%**	**92%**
		specificity	**88%**	**43%**	**92%**	**29%**
		F	**46%**	**86%**	**36%**	**87%**
		Accuracy	**78%**		**78%**	
		Kappa	**33%**		**24%**	
	rTBF_95tile	sensitivity	14%	92%	0%	93%
		specificity	92%	14%	93%	0%
		F	20%	85%	0%	84%
		Accuracy	75%		73%	
		Kappa	8%		5%	
	rTBF_median	sensitivity	57%	92%	19%	92%
		specificity	92%	57%	92%	19%
		F	62%	90%	26%	86%
		Accuracy	84%		76%	
		Kappa	52%		14%	
	rTBF_mode	sensitivity	14%	88%	10%	89%
		specificity	88%	14%	89%	10%
		F	18%	83%	13%	83%
		Accuracy	72%		72%	
		Kappa	3%		1%	
	rTBF_min	sensitivity	0%	100%	0%	97%

		specificity	100%	0%	97%	0%
		F	0%	88%	0%	86%
		Accuracy	78%		76%	
		Kappa	0%		0%	
	rTBF_max	sensitivity	0%	96%	0%	93%
		specificity	96%	0%	93%	0%
		F	0%	86%	0%	84%
		Accuracy	75%		73%	
		Kappa	0%		0%	
DSC	rCBF_mean	sensitivity	14%	96%	10%	96%
		specificity	96%	14%	96%	10%
		F	22%	87%	15%	87%
		Accuracy	78%		77%	
		Kappa	14%		8%	
	rCBF_skew	sensitivity	29%	96%	14%	93%
		specificity	96%	29%	93%	14%
		F	40%	89%	21%	86%
		Accuracy	81%		76%	
		Kappa	31%		10%	
	rCBF_median	sensitivity	14%	96%	10%	96%
		specificity	96%	14%	96%	10%
		F	22%	87%	15%	87%
		Accuracy	78%		77%	
		Kappa	14%		8%	
	rCBV_mean	sensitivity	29%	88%	10%	91%
		specificity	88%	29%	91%	10%
		F	33%	85%	13%	84%
		Accuracy	75%		73%	
		Kappa	18%		0.2%	
	rCBV_SD	sensitivity	0%	92%	0%	92%
		specificity	92%	0%	92%	0%
		F	0%	84%	0%	84%
		Accuracy	72%		72%	
		Kappa	0%		0%	
	rCBV_95tile	sensitivity	14%	92%	0%	92%
		specificity	92%	14%	92%	0%

		F	20%	85%	0%	84%
		Accuracy	75%		72%	
		Kappa	8%		0%	
	rCBV_median	sensitivity	29%	88%	10%	91%
		specificity	88%	29%	91%	10%
		F	33%	85%	13%	84%
		Accuracy	75%		73%	
		Kappa	18%		0.2%	
	rCBV_iqr	sensitivity	0%	92%	0%	95%
		specificity	92%	0%	95%	0%
		F	0%	84%	0%	85%
		Accuracy	72%		74%	
		Kappa	0%		0%	
	rCBV_mode	sensitivity	0%	96%	0%	92%
		specificity	96%	0%	92%	0%
		F	0%	86%	0%	84%
		Accuracy	75%		72%	
		Kappa	0%		0%	
	rCBV_slope	sensitivity	0%	88%	0%	89%
		specificity	88%	0%	89%	0%
		F	0%	81%	0%	82%
		Accuracy	69%		70%	
		Kappa	0%		0%	
	rCBV_entropy	**sensitivity**	**43%**	**92%**	**29%**	**92%**
		specificity	**92%**	**43%**	**92%**	**29%**
		F	**50%**	**88%**	**36%**	**87%**
		Accuracy	**81%**		**78%**	
		Kappa	**39%**		**24%**	
	rCBV_min	sensitivity	43%	88%	19%	92%
		specificity	88%	43%	92%	19%
		F	46%	86%	26%	86%
		Accuracy	78%		76%	
		Kappa	32%		14%	
	rCBV_max	sensitivity	14%	88%	10%	89%
		specificity	88%	14%	89%	10%
		F	18%	83%	13%	83%

		Accuracy	72%		72%	
		Kappa	3%		0%	
	rCBV_var	sensitivity	0%	92%	0%	95%
		specificity	92%	0%	95%	0%
		F	0%	84%	0%	85%
		Accuracy	72%		74%	
		Kappa	0%		0%	
DCE_mTK	**rK_mean**	**sensitivity**	**14%**	**100%**	**0%**	**92%**
		specificity	**100%**	**14%**	**92%**	**0%**
		F	**25%**	**89%**	**0%**	**84%**
		Accuracy	**81%**		**72%**	
		Kappa	**21%**		**0%**	
DCE_L&L	rK_mean	sensitivity	0%	100%	0%	100%
		specificity	100%	0%	100%	0%
		F	0%	88%	0%	88%
		Accuracy	78%		78%	
		Kappa	0%		0%	
	rK_SD	sensitivity	0%	100%	0%	100%
		specificity	100%	0%	100%	0%
		F	0%	88%	0%	88%
		Accuracy	78%		78%	
		Kappa	0%		0%	
	rK_95tile	sensitivity	0%	100%	0%	100%
		specificity	100%	0%	100%	0%
		F	0%	88%	0%	88%
		Accuracy	78%		78%	
		Kappa	0%		0%	
	rK_median	sensitivity	14%	100%	0%	97%
		specificity	100%	14%	97%	0%
		F	25%	89%	0%	86%
		Accuracy	81%		76%	
		Kappa	21%		0%	
	rK_iqr	sensitivity	0%	100%	0%	100%
		specificity	100%	0%	100%	0%
		F	0%	88%	0%	88%
		Accuracy	78%		78%	

		Kappa	0%		0%	
	rK_slope	sensitivity	29%	92%	0%	92%
		specificity	92%	29%	92%	0%
		F	36%	87%	0%	84%
		Accuracy	78%		72%	
		Kappa	24%		0%	
	rK_var	sensitivity	0%	100%	0%	100%
		specificity	100%	0%	100%	0%
		F	0%	88%	0%	88%
		Accuracy	78%		78%	
		Kappa	0%		0%	
	rPS_mean	sensitivity	0%	100%	0%	97%
		specificity	100%	0%	97%	0%
		F	0%	88%	0%	86%
		Accuracy	78%		76%	
		Kappa	0%		0%	
	rPS_median	sensitivity	0%	96%	0%	97%
		specificity	96%	0%	97%	0%
		F	0%	86%	0%	86%
		Accuracy	75%		76%	
		Kappa	0%		0%	
	rTC_mode	sensitivity	14%	92%	0%	95%
		specificity	92%	14%	95%	0%
		F	20%	85%	0%	85%
		Accuracy	75%		74%	
		Kappa	8%		0%	
	rTC_min	**sensitivity**	**29%**	**92%**	**19%**	**95%**
		specificity	**92%**	**29%**	**95%**	**19%**
		F	**36%**	**87%**	**28%**	**87%**
		Accuracy	**78%**		**78%**	
		Kappa	**24%**		**18%**	
ADC	ADC_mean	sensitivity	14%	96%	10%	95%
		specificity	96%	14%	95%	10%
		F	22%	87%	15%	86%
		Accuracy	78%		76%	
		Kappa	14%		6%	

	ADC_SD	**sensitivity**	**29%**	**100%**	**10%**	**100%**
		specificity	**100%**	**29%**	**100%**	**10%**
		F	**44%**	**91%**	**17%**	**89%**
		Accuracy	**84%**		**80%**	
		Kappa	**38%**		**14%**	
	ADC_median	sensitivity	14%	96%	0%	92%
		specificity	96%	14%	92%	0%
		F	22%	87%	0%	84%
		Accuracy	78%		72%	
		Kappa	14%		0%	
	ADC_20tile	sensitivity	14%	96%	0%	97%
		specificity	96%	14%	97%	0%
		F	22%	87%	0%	86%
		Accuracy	78%		76%	
		Kappa	14%		0%	
	ADC_30tile	sensitivity	0%	96%	0%	97%
		specificity	96%	0%	97%	0%
		F	0%	86%	0%	86%
		Accuracy	75%		76%	
		Kappa	0%		0%	
	ADC_40tile	sensitivity	0%	96%	0%	95%
		specificity	96%	0%	95%	0%
		F	0%	86%	0%	85%
		Accuracy	75%		74%	
		Kappa	0%		0%	
	ADC_60tile	sensitivity	29%	96%	10%	96%
		specificity	96%	29%	96%	10%
		F	40%	89%	15%	87%
		Accuracy	81%		77%	
		Kappa	31%		8%	
	ADC_70tile	sensitivity	29%	96%	10%	96%
		specificity	96%	29%	96%	10%
		F	40%	89%	15%	87%
		Accuracy	81%		77%	
		Kappa	31%		8%	
	ADC_80tile	sensitivity	14%	96%	10%	96%

		specificity	96%	14%	96%	10%
		F	22%	87%	15%	87%
		Accuracy	78%		77%	
		Kappa	14%		8%	
	ADC_90tile	sensitivity	14%	96%	10%	97%
		specificity	96%	14%	97%	10%
		F	22%	87%	16%	87%
		Accuracy	78%		78%	
		Kappa	14%		10%	
	ADC_95tile	sensitivity	29%	96%	10%	97%
		specificity	96%	29%	97%	10%
		F	40%	89%	16%	87%
		Accuracy	81%		78%	
		Kappa	31%		10%	

Table 5-B. Diagnostic performance of significant features from ASL, DSC, DCE-mTK, DCE-L&L, ADC for glioma grouping according to the predefined histomolecular types.

MRI-method	Features	Prediction performance	Grouping			Internal validation LOOCV		
			GBM-wt	Astro-mut	Oligo-mut	GBM-wt	Astro-mut	Oligo-mut
ASL	**rTBF_skew**	**sensitivity**	**43%**	**83%**	**69%**	**10%**	**83%**	**64%**
		specificity	**96%**	**80%**	**74%**	**95%**	**78%**	**61%**
		F	**55%**	**77%**	**67%**	**15%**	**76%**	**58%**
		Accuracy	**69%**			**59%**		
		Kappa	**50%**			**34%**		
	rTBF_zscore	sensitivity	43%	83%	69%	0%	78%	44%
		specificity	96%	80%	74%	91%	70%	54%
		F	55%	77%	67%	0%	68%	41%
		Accuracy	69%			47%		
		Kappa	50%			15%		
	rTBF_mean	sensitivity	57%	83%	54%	29%	83%	56%
		specificity	88%	85%	74%	92%	82%	63%
		F	57%	80%	56%	36%	78%	54%
		Accuracy	66%			60%		
		Kappa	47%			37%		
	rTBF_95tile	sensitivity	14%	83%	54%	0%	58%	59%
		specificity	92%	80%	58%	95%	78%	39%
		F	20%	77%	50%	0%	60%	47%
		Accuracy	56%			46%		

		Kappa	3%			12%		
	rTBF_median	sensitivity	57%	83%	46%	29%	83%	51%
		specificity	88%	80%	74%	92%	78%	63%
		F	57%	77%	50%	36%	76%	50%
		Accuracy	63%			58%		
		Kappa	42%			34%		
	rTBF_mode	sensitivity	57%	67%	54%	10%	67%	51%
		specificity	88%	75%	74%	89%	73%	54%
		F	57%	64%	56%	13%	63%	47%
		Accuracy	59%			48%		
		Kappa	37%			17%		
	rTBF_min	sensitivity	0%	75%	46%	0%	50%	33%
		specificity	100%	55%	58%	97%	52%	40%
		F	0%	60%	44%	0%	43%	30%
		Accuracy	47%			32%		
		Kappa	13%			0%		
	rTBF_max	sensitivity	0%	75%	62%	0%	67%	51%
		specificity	96%	70%	58%	94%	65%	53%
		F	0%	67%	55%	0%	59%	47%
		Accuracy	53%			46%		
		Kappa	24%			12%		
DSC	rCBF_mean	sensitivity	14%	67%	62%	5%	67%	62%
		specificity	96%	80%	47%	97%	78%	44%

		F	22%	67%	52%	8%	66%	50%
		Accuracy	53%			51%		
		Kappa	24%			20%		
	rCBF_SD	sensitivity	0%	75%	85%	0%	61%	87%
		specificity	100%	70%	68%	100%	77%	54%
		F	0%	67%	73%	0%	61%	69%
		Accuracy	62%			58%		
		Kappa	38%			31%		
	rCBF_95tile	sensitivity	0%	75%	85%	0%	72%	79%
		specificity	100%	85%	53%	97%	87%	49%
		F	0%	75%	67%	0%	74%	63%
		Accuracy	63%			59%		
		Kappa	38%			33%		
	rCBF_kur	sensitivity	0%	67%	77%	0%	58%	77%
		specificity	100%	65%	63%	97%	65%	61%
		F	0%	59%	67%	0%	54%	66%
		Accuracy	56%			53%		
		Kappa	28%			23%		
	rCBF_skew	sensitivity	29%	67%	69%	10%	58%	64%
		specificity	96%	75%	63%	93%	75%	51%
		F	40%	64%	62%	14%	58%	54%
		Accuracy	59%			50%		
		Kappa	35%			19%		

	rCBF_median	sensitivity	14%	67%	69%	0%	67%	64%
		specificity	96%	85%	47%	97%	80%	42%
		F	22%	70%	56%	0%	67%	52%
		Accuracy	56%			51%		
		Kappa	29%			20%		
	rCBF_iqr	sensitivity	0%	75%	85%	0%	72%	79%
		specificity	100%	70%	68%	100%	67%	67%
		F	0%	67%	73%	0%	63%	70%
		Accuracy	63%			59%		
		Kappa	38%			33%		
	rCBF_mode	sensitivity	0%	75%	69%	0%	75%	51%
		specificity	100%	75%	53%	95%	68%	54%
		F	0%	69%	58%	0%	66%	47%
		Accuracy	56%			49%		
		Kappa	28%			17%		
	rCBF_slope	sensitivity	0%	67%	85%	0%	56%	79%
		specificity	100%	65%	68%	100%	62%	61%
		F	0%	59%	73%	0%	51%	67%
		Accuracy	59%			53%		
		Kappa	33%			23%		
	rCBF_max	sensitivity	0%	83%	85%	0%	64%	79%
		specificity	100%	90%	53%	97%	90%	40%
		F	0%	83%	67%	0%	71%	60%

		Accuracy	66%			56%		
		Kappa	43%			28%		
	rCBF_zscore	sensitivity	0%	75%	77%	0%	75%	69%
		specificity	100%	75%	58%	96%	75%	58%
		F	0%	69%	65%	0%	69%	60%
		Accuracy	59%			56%		
		Kappa	33%			29%		
	rCBF_var	sensitivity	0%	83%	85%	0%	53%	79%
		specificity	100%	70%	74%	100%	72%	49%
		F	0%	71%	76%	0%	53%	63%
		Accuracy	66%			52%		
		Kappa	44%			21%		
	rCBV_mean	sensitivity	29%	75%	54%	10%	75%	59%
		specificity	88%	85%	58%	92%	83%	51%
		F	33%	75%	50%	14%	74%	51%
		Accuracy	56%			54%		
		Kappa	31%			26%		
	rCBV_SD	sensitivity	0%	67%	62%	0%	64%	49%
		specificity	92%	85%	42%	95%	73%	40%
		F	0%	70%	50%	0%	61%	41%
		Accuracy	50%			44%		
		Kappa	19%			8%		
	rCBV_95tile	sensitivity	14%	75%	54%	0%	75%	36%

		specificity	92%	80%	53%	92%	68%	47%
		F	20%	72%	48%	0%	66%	34%
		Accuracy	53%			43%		
		Kappa	25%			8%		
	rCBV_skew	sensitivity	0%	75%	85%	0%	75%	85%
		specificity	100%	80%	58%	100%	80%	58%
		F	0%	72%	69%	0%	72%	69%
		Accuracy	63%			63%		
		Kappa	38%			38%		
	rCBV_median	sensitivity	29%	83%	46%	0%	78%	46%
		specificity	92%	75%	63%	92%	75%	49%
		F	36%	74%	46%	0%	71%	42%
		Accuracy	56%			48%		
		Kappa	31%			16%		
	rCBV_iqr	sensitivity	0%	67%	77%	0%	58%	59%
		specificity	100%	80%	47%	95%	75%	42%
		F	0%	67%	61%	0%	58%	48%
		Accuracy	56%			46%		
		Kappa	28%			12%		
	rCBV_mode	sensitivity	0%	83%	77%	0%	83%	62%
		specificity	100%	80%	58%	97%	73%	58%
		F	0%	77%	65%	0%	73%	55%
		Accuracy	63%			56%		

		Kappa	38%			29%		
	rCBV_slope	sensitivity	0%	58%	62%	0%	58%	67%
		specificity	88%	85%	42%	92%	85%	40%
		F	0%	64%	50%	0%	64%	53%
		Accuracy	47%			49%		
		Kappa	14%			17%		
	rCBV_entropy	**sensitivity**	**57%**	**75%**	**69%**	**29%**	**75%**	**64%**
		specificity	**92%**	**90%**	**68%**	**93%**	**85%**	**58%**
		F	**62%**	**78%**	**64%**	**38%**	**75%**	**57%**
		Accuracy	**69%**			**60%**		
		Kappa	**51%**			**37%**		
	rCBV_min	sensitivity	71%	50%	69%	33%	39%	44%
		specificity	80%	95%	68%	87%	73%	44%
		F	59%	63%	64%	37%	42%	39%
		Accuracy	63%			40%		
		Kappa	43%			5%		
	rCBV_max	sensitivity	29%	75%	77%	10%	69%	56%
		specificity	88%	95%	63%	92%	78%	51%
		F	33%	82%	67%	14%	68%	49%
		Accuracy	66%			51%		
		Kappa	46%			21%		
	rCBV_var	sensitivity	0%	67%	46%	0%	64%	26%
		specificity	96%	70%	42%	95%	58%	40%

		F	0%	62%	40%	0%	55%	24%
		Accuracy	44%			34%		
		Kappa	8%			0%		
DCE-mTK	Ve_mode	sensitivity	14%	83%	54%	0%	75%	54%
		specificity	96%	60%	74%	96%	60%	63%
		F	22%	67%	56%	0%	62%	52%
		Accuracy	56%			50%		
		Kappa	30%			19%		
	rVe_mode	sensitivity	0%	67%	69%	0%	64%	49%
		specificity	100%	70%	53%	97%	60%	51%
		F		62%	58%	0%	55%	44%
		Accuracy	53%			44%		
		Kappa	23%			8%		
	rVp_mean	sensitivity	0%	83%	62%	0%	83%	62%
		specificity	100%	60%	68%	100%	60%	68%
		F	0%	67%	59%	0%	67%	59%
		Accuracy	56%			56%		
		Kappa	28%			28%		
	rVp_median	**sensitivity**	**0%**	**83%**	**69%**	**0%**	**83%**	**62%**
		specificity	**100%**	**65%**	**68%**	**100%**	**60%**	**68%**
		F	**0%**	**69%**	**64%**	**0%**	**67%**	**59%**
		Accuracy	**59%**			**56%**		
		Kappa	**33%**			**28%**		

	rVp_min	sensitivity	0%	83%	69%	0%	64%	69%
		specificity	100%	70%	63%	100%	70%	51%
		F	0%	71%	62%	0%	60%	57%
		Accuracy	59%			52%		
		Kappa	33%			21%		
DCE-L&L	rK_mean	sensitivity	29%	67%	46%	0%	42%	46%
		specificity	96%	45%	79%	95%	35%	65%
		F	40%	52%	52%	0%	33%	47%
		Accuracy	50%			34%		
		Kappa	21%			0%		
	rK_SD	sensitivity	0%	67%	38%	0%	39%	38%
		specificity	100%	30%	74%	95%	30%	63%
		F	0%	47%	43%	0%	30%	40%
		Accuracy	41%			30%		
		Kappa	3%			0%		
	rK_iqr	sensitivity	0%	75%	46%	0%	50%	38%
		specificity	100%	35%	79%	95%	30%	70%
		F	0%	53%	52%	0%	38%	42%
		Accuracy	47%			34%		
		Kappa	14%			0%		
	rK_var	sensitivity	0%	67%	38%	0%	67%	38%
		specificity	100%	30%	74%	100%	30%	74%
		F	0%	47%	43%	0%	47%	43%

		Accuracy	41%			41%		
		Kappa	3%			3%		
	rFb_mean	**sensitivity**	**0%**	**92%**	**77%**	**0%**	**81%**	**78%**
		specificity	**100%**	**60%**	**84%**	**100%**	**60%**	**77%**
		F	**0%**	**71%**	**77%**	**0%**	**65%**	**73%**
		Accuracy	**66%**			**61%**		
		Kappa	**44%**			**37%**		
	rFb_SD	sensitivity	0%	75%	69%	0%	75%	69%
		specificity	100%	55%	74%	100%	55%	74%
		F	0%	60%	67%	0%	60%	67%
		Accuracy	56%			56%		
		Kappa	28%			28%		
	rFb_95tile	sensitivity	0%	83%	69%	0%	75%	69%
		specificity	100%	55%	79%	100%	55%	74%
		F	0%	65%	69%	0%	60%	67%
		Accuracy	59%			56%		
		Kappa	34%			28%		
	rFb_iqr	sensitivity	0%	92%	77%	0%	75%	77%
		specificity	100%	60%	84%	100%	62%	72%
		F	0%	71%	77%	0%	63%	71%
		Accuracy	66%			59%		
		Kappa	44%			33%		
	rFb_slope	sensitivity	0%	75%	77%	0%	75%	69%

		specificity	100%	65%	68%	100%	60%	68%
		F	0%	64%	69%	0%	62%	64%
		Accuracy	59%			56%		
		Kappa	33%			28%		
	rFb_min	sensitivity	0%	75%	69%	0%	47%	64%
		specificity	100%	60%	68%	100%	63%	44%
		F	0%	62%	64%	0%	45%	52%
		Accuracy	56%			44%		
		Kappa	28			7%		
	rFb_max	sensitivity	0%	75%	69%	0%	75%	69%
		specificity	100%	50%	79%	100%	52%	77%
		F	0%	58%	69%	0%	59%	68%
		Accuracy	56%			56%		
		Kappa	29%			29%		
	rFb_var	sensitivity	0%	83%	69%	0%	75%	69%
		specificity	100%	50%	84%	100%	52%	77%
		F	0%	63%	72%	0%	59%	68%
		Accuracy	59%			56%		
		Kappa	34%			29%		
	rTC_mean	sensitivity	0%	67%	77%	0%	44%	77%
		specificity	100%	70%	58%	100%	70%	44%
		F	0%	62%	65%	0%	46%	59%
		Accuracy	56%%			48%		

		Kappa	28%			14%		
	rTC_mode	sensitivity	29%	8%	54%	19%	6%	49%
		specificity	84%	50%	58%	89%	50%	42%
		F	31%	9%	50%	42%	6%	42%
		Accuracy	31%			26%		
		Kappa	0%			0%		
	rTC_min	sensitivity	29%	42%	69%	19%	6%	69%
		specificity	92%	60%	68%	87%	63%	46%
		F	36%	40%	64%	23%	7%	56%
		Accuracy	50%			34%		
		Kappa	21%			0%		
ADC	ADC_mode	sensitivity	0%	75%	77%	0%	75%	72%
		specificity	100%	85%	47%	100%	80%	49%
		F	0%	75%	61%	0%	72%	58%
		Accuracy	59%			57%		
		Kappa	33%			30%		
	ADC_mean	sensitivity	14%	58%	62%	10%	58%	54%
		specificity	96%	80%	42%	95%	77%	40%
		F	22%	61%	50%	15%	59%	45%
		Accuracy	50%			46%		
		Kappa	19%			12%		
	ADC_SD	sensitivity	43%	75%	85%	19%	69%	56%
		specificity	96%	85%	74%	92%	72%	61%

		F	55%	75%	76%	26%	64%	53%
		Accuracy	72%			53%		
		Kappa	55			25%		
	ADC_median	sensitivity	29%	58%	69%	10%	58%	62%
		specificity	96%	85%	47%	92%	82%	44%
		F	40%	64%	56%	14%	62%	51%
		Accuracy	56%			49%		
		Kappa	30%			18%		
	ADC_20tile	sensitivity	14%	42%	69%	0%	42%	49%
		specificity	96%	85%	32%	95%	73%	26%
		F	22%	50%	51%	0%		
		Accuracy	47%			35%		
		Kappa	13%			0%		
	ADC_30tile	sensitivity	14%	58%	69%	0%	58%	56%
		specificity	96%	85%	42%	95%	78%	37%
		F	22%	64%	55%	0%	60%	45%
		Accuracy	53%			45%		
		Kappa	24%			10%		
	ADC_40tile	sensitivity	14%	58%	69%	0%	58%	62%
		specificity	92%	85%	47%	92%	82%	40%
		F	20%	64%	56%	0%	62%	49%
		Accuracy	53%			47%		
		Kappa	19%			14%		

	ADC_60tile	sensitivity	29%	67%	85%	10%	58%	53%
		specificity	96%	95%	53%	95%	77%	40%
		F	40%	76%	67%	15%	59%	45%
		Accuracy	66%			46%		
		Kappa	45%			12%		
	ADC_70tile	sensitivity	43%	58%	77%	10%	58%	64%
		specificity	96%	90%	53%	95%	83%	40%
		F	55%	67%	63%	15%	63%	51%
		Accuracy	62%			50%		
		Kappa	40%			19%		
	ADC_80tile	sensitivity	29%	67%	77%	10%	67%	59%
		specificity	96%	90%	53%	95%	80%	46%
		F	40%	73%	63%	15%	67%	49%
		Accuracy	63%			51%		
		Kappa	40%			21%		
	ADC_90tile	sensitivity	14%	50%	62%	10%	50%	56%
		specificity	96%	80%	37%	95%	78%	35%
		F	22%	55%	48%	15%	54%	45%
		Accuracy	47%			44%		
		Kappa	14%			9%		
	ADC_95tile	sensitivity	29%	58%	62%	10%	50%	51%
		specificity	96%	80%	47%	95%	75%	35%
		F	40%	61%	52%	15%	52%	42%

		Accuracy	53%			42%		
		Kappa	25%			5%		
	ADC_ and rADC_skew	sensitivity	0%	75%	77%	0%	75%	77%
		specificity	96%	85%	53%	96%	85%	52%
		F	0%	75%	63%	0%	75%	63%
		Accuracy	59%			59%		
		Kappa	34%			34%		
	ADC_slope	sensitivity	29%	67%	62%	19%	44%	62%
		specificity	96%	70%	63%	97%	72%	42%
		F	40%	62%	57%	30%	46%	50%
		Accuracy	56%			46%		
		Kappa	39%			12%		
	ADC_ and rADC_zscore	**sensitivity**	**0%**	**83%**	**69%**	**0%**	**83%**	**69%**
		specificity	**96%**	**80%**	**58%**	**96%**	**80%**	**58%**
		F	**0%**	**77%**	**60%**	**0%**	**77%**	**60%**
		Accuracy	**59%**			**59%**		
		Kappa	**34%**			**34%**		
	ADC_var	sensitivity	43%	67%	77%	19%	58%	67%
		specificity	92%	90%	63%	92%	82%	51%
		F	50%	73%	67%	56%	62%	56%
		Accuracy	66%			53%		
		Kappa	46%			25%		
	rADC_mode	sensitivity	0%	75%	92%	0%	65%	69%

		specificity	100%	95%	47%	100%	80%	47%
		F	0%	82%	69%	0%	72%	56%
		Accuracy	66%			56%		
		Kappa	43%			28%		
	rADC_mean	sensitivity	43%	58%	62%	29%	58%	56%
		specificity	96%	80%	53%	97%	75%	47%
		F	55%	61%	53%	41%	58%	48%
		Accuracy	56%			51%		
		Kappa	30%			21%		
	rADC_SD	sensitivity	43%	75%	85%	10%	47%	82%
		specificity	96%	85%	74%	97%	83%	42%
		F	55%	75%	76%	16%	54%	62%
		Accuracy	72%			53%		
		Kappa	55%			23%		
	rADC_median	sensitivity	43%	67%	69%	19%	67%	72%
		specificity	96%	85%	58%	97%	85%	49%
		F	55%	70%	60%	30%	70%	58%
		Accuracy	63%			58%		
		Kappa	40%			32%		
	rADC_10tile	sensitivity	14%	42%	54%	0%	42%	51%
		specificity	96%	75%	32%	97%	72%	26%
		F	22%	45%	42%	0%	44%	40%
		Accuracy	41%			36%		

		Kappa	3%			0%		
	rADC_20tile	sensitivity	14%	50%	62%	10%	44%	59%
		specificity	96%	80%	37%	96%	78%	32%
		F	22%	55%	48%	15%	49%	46%
		Accuracy	47%			43%		
		Kappa	14%			7%		
	rADC_30tile	sensitivity	57%	67%	69%	0%	56%	62%
		specificity	96%	85%	63%	95%	82%	35%
		F	67%	70%	62%	0%	60%	48%
		Accuracy	66%			46%		
		Kappa	46%			11%		
	rADC_40tile	sensitivity	43%	67%	69%	19%	67%	69%
		specificity	96%	85%	58%	97%	83%	49%
		F	55%	70%	60%	30%	69%	57%
		Accuracy	63%			57%		
		Kappa	40%			31%		
	rADC_60tile	sensitivity	43%	67%	69%	19%	67%	64%
		specificity	96%	85%	58%	97%	67%	49%
		F	55%	70%	60%	30%	67%	54%
		Accuracy	63%			55%		
		Kappa	40%			27%		
	rADC_70tile	sensitivity	43%	75%	77%	19%	58%	56%
		specificity	96%	90%	63%	95%	78%	44%

		F	55%	78%	67%	28%	60%	47%
		Accuracy	69%			49%		
		Kappa	50%			18%		
	rADC_80tile	sensitivity	43%	67%	69%	19%	67%	67%
		specificity	92%	90%	58%	95%	85%	49%
		F	50%	73%	60%	28%	70%	55%
		Accuracy	63%			56%		
		Kappa	41%			29%		
	rADC_90tile	sensitivity	29%	75%	62%	19%	75%	62%
		specificity	92%	85%	58%	95%	82%	54%
		F	36%	75%	55%	28%	73%	54%
		Accuracy	59%			57%		
		Kappa	35%			31%		
	rADC_95tile	sensitivity	29%	67%	62%	19%	67%	59%
		specificity	92%	85%	53%	95%	80%	49%
		F	36%	70%	53%	28%	67%	51%
		Accuracy	56%			53%		
		Kappa	30%			24%		
	rADC_slope	sensitivity	29%	75%	69%	19%	58%	64%
		specificity	96%	75%	68%	97%	70%	54%
		F	40%	69%	64%	30%	56%	56%
		Accuracy	63%			52%		
		Kappa	4%			22%		

	rADC_var	sensitivity	43%	67%	85%	5%	44%	82%
		specificity	96%	90%	63%	96%	90%	33%
		F	55%	73%	71%	8%	55%	59%
		Accuracy	69%			51%		
		Kappa	50%			20%		

Table 6-B.Diagnostic performance of significant features from ASL, DSC, DCE-mTK, DCE-L&L, ADC for glioma sub-grouping according to the predefined histomolecular types and histological grades.

MRI-method	Features	Prediction performance	Sub-grouping					Internal validation LOOCV				
			GBM-wt	Astro-mut-LG	Astro-mut-HG	Oligo-mut-LG	Oligo-mut-HG	GBM-wt	Astro-mut-LG	Astro-mut-HG	Oligo-mut-LG	Oligo-mut-HG
ASL	aTBF_ & rTBF_skew	sensitivity	43%	90%	0%	56%	0%	0%	90%	0%	44%	0%
		specificity	92%	68%	100%	74%	100%	92%	62%	100%	62%	100
		F	50%	69%	0%	50%	0%	0%	66%	0%	37%	0%
		Accuracy	53%					41%				
		Kappa	34%					16%				
	aTBF_ & rTBF_zscore	sensitivity	57%	90%	0%	44%	0%	0%	83%	0%	41%	0%
		specificity	92%	64%	100%	78%	100%	91%	64%	100%	58%	100%
		F	62%	67%	0%	44%	0%	0%	63%	0%	33%	0%
		Accuracy	53%					38%				
		Kappa	34%					12%				
	rTBF_mean	sensitivity	57%	90%	0%	56%	0%	38%	83%	0%	44%	0%
		specificity	84%	82%	100%	74%	100%	87%	76%	100%	64%	100%
		F	53%	78%	0%	50%	0%	41%	70%	0%	38%	0%
		Accuracy	56%					47%				
		Kappa	39					26%				
	rTBF_95tile	sensitivity	57%	80%	0%	44%	0%	38%	70%	0%	37%	0%
		specificity	76%	77%	100%	78%	100%	79%	73%	100%	67%	100%
		F	47%	70%	0%	44%	0%	36%	61%	0%	33%	0%
		Accuracy	50%					41%				
		Kappa	31%					18%				
	rTBF_median	**sensitivity**	**71%**	**90%**	**0%**	**56%**	**0%**	**38%**	**90%**	**0%**	**48%**	**0%**
		specificity	**84%**	**77%**	**100%**	**83%**	**100%**	**87%**	**74%**	**100%**	**70%**	**100%**
		F	**63%**	**75%**	**0%**	**56%**	**0%**	**41%**	**73%**	**0%**	**43%**	**0%**
		Accuracy	**59%**					**50%**				
		Kappa	**44%**					**30%**				
	rTBF_mode	sensitivity	71%	80%	0%	33%	0%	48%	73%	0%	7%	0%
		specificity	84%	73%	100%	74%	100%	80%	67%	100%	64%	100%
		F	63%	67%	0%	33%	0%	43%	59%	0%	7%	0%
		Accuracy	50%%					35%				
		Kappa	31%					10%				

	rTBF_min	sensitivity	0%	90%	0%	56%	0%	0%	93%	0%	37%	0%
		specificity	100%	59%	100%	61%	100%	97%	39%	100%	77%	100%
		F	0%	64%	0%	43%	0%	0%	57%	0%	38%	0%
		Accuracy	45%					40%				
		Kappa	20%					13%				
DSC	rCBF_mean	sensitivity	14%	80%	0%	56%	0%	0%	80%	0%	44%	0%
		specificity	96%	73%	100%	52%	100%	92%	70%	100%	51%	100%
		F	22%	67%	0%	40%	0%	0%	65%	0%	33%	0%
		Accuracy	44%					38%				
		Kappa	21%					12%				
	rCBF_SD	sensitivity	14%	70%	0%	78%	0%	0%	70%	0%	59%	0%
		specificity	88%	68%	100%	70%	100%	92%	55%	100%	67%	100%
		F	18%	58%	0%	61%	0%	0%	52%	0%	48%	0%
		Accuracy	47%					39%				
		Kappa	25%					13%				
	rCBF_95tile	sensitivity	0%	90%	0%	89%	0%	0%	80%	0%	70%	0%
		specificity	100%	82%	100%	52%	100%	95%	80%	100%	48%	100%
		F	0%	78%	0%	57%	0%	0%	72%	0%	46%	0%
		Accuracy	53%					45%				
		Kappa	34%					22%				
	rCBF_kur	sensitivity	0%	70%	0%	56%	25%	0%	70%	0%	59%	8%
		specificity	100%	64%	100%	61%	93%	100%	63%	100%	57%	95%
		F	0%	56%	0%	43%	29%	0%	56%	0%	44%	12%
		Accuracy	41%					40%				
		Kappa	17%					15%				
	rCBF_skew	sensitivity	57%	70%	0%	33%	0%	29%	70%	0%	33%	0%
		specificity	88%	68%	100%	65%	100%	88%	68%	100%	57%	100%
		F	57%	58%	0%	30%	0%	33%	58%	0%	27%	0%
		Accuracy	44%					38%				
		Kappa	22%					13%				
	rCBF_median	sensitivity	14%	80%	0%	56%	0%	0%	73%	0%	44%	0%
		specificity	96%	73%	100%	52%	100%	92%	68%	100%	49%	100%
		F	22%	67%	0%	40%	0%	0%	60%	0%	32%	0%
		Accuracy	44%					35%				
		Kappa	21%					9%				
	rCBF_iqr	sensitivity	0%	80%	0%	78%	0%	0%	73%	0%	74%	0%

		specificity	88%	73%	100%	65%	100%	89%	68%	100%	64%	100%
		F	0%	67%	0%	58%	0%	0%	60%	0%	56%	0%
		Accuracy	47%					44%				
		Kappa	25%					21%				
	rCBF_mode	sensitivity	0%	70%	0%	67%	0%	0%	70%	0%	41%	0%
		specificity	100%	55%	100%	61%	100%	92%	53%	100%	61%	100%
		F	0%	52%	0%	50%	0%	0%	51%	0%	34%	0%
		Accuracy	41%					33%				
		Kappa	15%					6%				
	rCBF_max	sensitivity	0%	90%	0%	89%	0%	0%	87%	0%	59%	0%
		specificity	100%	77%	100%	57%	100%	95%	70%	100%	57%	100%
		F	0%	75%	0%	59%	0%	0%	68%	0%	44%	0%
		Accuracy	53%					44%				
		Kappa	34%					20%				
	rCBF_zscore	sensitivity	29%	70%	0%	56%	0%	0%	70%	0%	22%	0%
		specificity	96%	68%	100%	57%	100%	84%	68%	100%	100%	100%
		F	40%	58%	0%	42%	0%	0%	58%	0%	17%	0%
		Accuracy	44%					28%				
		Kappa	21%					0%				
	rCBF_var	sensitivity	14%	80%	0%	78%	00%	0%	73%	0%	56%	0%
		specificity	92%	68%	100%	70%	100%	92%	50%	100%	71%	100%
		F	20%	64%	0%	61%	0%	0%	52%	0%	48%	0%
		Accuracy	50%					39%				
		Kappa	30%					13%				
	rCBV_mean	sensitivity	71%	80%	0%	67%	0%	38%	83%	0%	52%	0%
		specificity	88%	82%	100%	74%	100%	91%	74%	100%	64%	100%
		F	67%	73%	0%	57%	0%	44%	69%	0%	42%	0%
		Accuracy	59%					49%				
		Kappa	44%					29%				
	rCBV_SD	sensitivity	71%	70%	0%	56%	50%	38%	67%	0%	33%	0%
		specificity	92%	82%	100%	70%	100%	87%	70%	100%	58%	100%
		F	71%	67%	0%	48%	67%	41%	57%	0%	28%	0%
		Accuracy	59%					39%				
		Kappa	44%					15%				
	rCBV_95tile	sensitivity	86%	80%	0%	56%	0%	48%	80%	0%	48%	0%
		specificity	80%	82%	100%	83%	100%	81%	76%	100%	72%	100%

		F	67%	73%	0%	56%	0%	44%	69%	0%	44%	0%
		Accuracy	59%					45%				
		Kappa	44%					29%				
	rCBV_median	sensitivity	43%	90%	0%	56%	0%	19%	83%	0%	56%	0%
		specificity	88%	73%	100%	74%	100%	91%	71%	100%	62%	100%
		F	46%	72%	0%	50%	0%	25%	68%	0%	44%	0%
		Accuracy	53%					46%				
		Kappa	35%					24%				
	rCBV_iqr	sensitivity	57%	60%	0%	44%	50%	38%	63%	0%	37%	0%
		specificity	88%	77%	100%	65%	100%	85%	71%	100%	58%	100%
		F	57%	57%	0%	38%	67%	40%	56%	0%	30%	0%
		Accuracy	50%					39%				
		Kappa	32%					14%				
	rCBV_mode	sensitivity	43%	80%	0%	56%	25%	0%	87%	0%	26%	0%
		specificity	92%	73%	100%	70%	100%	87%	65%	100%	64%	94%
		F	50%	67%	0%	48%	40%	0%	66%	0%	24%	0%
		Accuracy	53%					34%				
		Kappa	0.001					9%				
	rCBV_slope	sensitivity	71%	70%	0%	56%	0%	48%	63%	0%	48%	0%
		specificity	76%	82%	100%	78%	100%	81%	79%	100%	62%	100%
		F	56%	67%	0%	53%	0%	44%	60%	0%	39%	0%
		Accuracy	53%					44%				
		Kappa	36%					22%				
	rCBV_entropy	**sensitivity**	**86%**	**70%**	**0%**	**78%**	**0%**	**81%**	**63%**	**0%**	**67%**	**0%**
		specificity	**80%**	**91%**	**100%**	**78%**	**100%**	**79%**	**86%**	**100%**	**75%**	**100%**
		F	**67%**	**74%**	**0%**	**67%**	**0%**	**63%**	**66%**	**0%**	**58%**	**0%**
		Accuracy	**63%**					**56%**				
		Kappa	**49%**					**40%**				
	rCBV_min	sensitivity	71%	60%	0%	56%	0%	48%	63%	0%	7%	0%
		specificity	76%	86%	100%	70%	100%	84%	50%	100%	71%	100%
		F	56%	63%	0%	48%	0%	47%	46%	0%	8%	0%
		Accuracy	50%					32%				
		Kappa	32%					5%				
	rCBV_max	sensitivity	86%	70%	0%	67%	0%	52%	73%	0%	48%	0%
		specificity	80%	91%	100%	74%	100%	80%	80%	100%	68%	100%
		F	67%	74%	0%	57%	0%	47%	68%	0%	42%	0%

		Accuracy	59%					48%				
		Kappa	44%					28%				
	rCBV_var	sensitivity	71%	70%	0%	56%	50%	38%	77%	0%	30%	0%
		specificity	92%	82%	100%	70%	100%	87%	62%	100%	68%	100%
		F	71%	67%	0%	48%	67%	41%	59%	0%	28%	0%
		Accuracy	59%					41%				
		Kappa	44%					17%				
DCE-mTK	Ve_mode	sensitivity	14%	80%	0%	44%	0%	0%	80%	0%	22%	0%
		specificity	80%	68%	100%	70%	100%	77%	64%	100%	68%	96%
		F	15%	64%	0%	40%	0%	0%	62%	0%	22%	0%
		Accuracy	41%					31%				
		Kappa	17%					4%				
	rVe_mode	sensitivity	14%	70%	0%	33%	25%	0%	80%	0%	22%	0%
		specificity	80%	68%	100%	70%	96%	84%	48%	100%	75%	96%
		F	15%	58%	0%	32%	33%	0%	55%	0%	24%	0%
		Accuracy	38%					31%				
		Kappa	14%					4%				
	rVp_mean	sensitivity	0%	80%	0%	56%	50%	0%	77%	0%	48%	33%
		specificity	100%	59%	100%	65%	100%	100%	59%	100%	61%	98%
		F	0%	59%	0%	45%	67%	0%	58%	0%	39%	44%
		Accuracy	47%					42%				
		Kappa	25%					18%				
	rVp_median	**sensitivity**	**0%**	**90%**	**0%**	**67%**	**50%**	**0%**	**90%**	**0%**	**52%**	**33%**
		specificity	**100%**	**68%**	**100%**	**70%**	**96%**	**100%**	**62%**	**100%**	**67%**	**96%**
		F	**0%**	**69%**	**0%**	**55%**	**57%**	**0%**	**66%**	**0%**	**44%**	**42%**
		Accuracy	**53%**					**47%**				
		Kappa	**35%**					**26%**				
	rVp_min	sensitivity	0%	70%	0%	56%	25%	0%	77%	0%	41%	0%
		specificity	100%	68%	100%	52%	96%	97%	56%	100%	58%	98%
		F	0%	58%	0%	40%	33%	0%	56%	0%	33%	0%
		Accuracy	41%					35%				
		Kappa	17%					9%				
DCE-L&L	rK_iqr	sensitivity	57%	40%	0%	11%	25%	0%	50%	0%	4%	0%
		specificity	84%	45%	100%	78%	96%	89%	26%	100%	74%	94%
		F	53%	31%	0%	13%	33%	0%	31%	0%	4%	0%
		Accuracy	31%					17%				

		Kappa	6%					0%				
	rFb_mean	**sensitivity**	**0%**	**80%**	**0%**	**56%**	**75%**	**0%**	**80%**	**0%**	**56%**	**58%**
		specificity	**100%**	**59%**	**100%**	**74%**	**96%**	**100%**	**59%**	**100%**	**71%**	**96%**
		F	**0%**	**59%**	**0%**	**50%**	**75%**	**0%**	**59%**	**0%**	**48%**	**64%**
		Accuracy	**50%**					**48%**				
		Kappa	**31%**					**28%**				
	rFb_SD	sensitivity	0%	80%	0%	44%	25%	0%	80%	0%	44%	17%
		specificity	100%	55%	100%	65%	96%	100%	55%	100%	64%	96%
		F	0%	57%	0%	38%	33%	0%	57%	0%	38%	24%
		Accuracy	41%					40%				
		Kappa	16%					15%				
	rFb_95tile	sensitivity	0%	80%	0%	44%	25%	0%	80%	0%	44%	17%
		specificity	100%	55%	100%	65%	96%	100%	55%	100%	64%	96%
		F	0%	57%	0%	38%	33%	0%	57%	0%	38%	24%
		Accuracy	41%					40%				
		Kappa	16%					15%				
	rFb_median	sensitivity	0%	70%	0%	56%	75%	0%	70%	0%	41%	67%
		specificity	100%	64%	100%	70%	93%	100%	56%	100%	68%	94%
		F	0%	56%	0%	48%	67%	0%	53%	0%	37%	64%
		Accuracy	47%					42%				
		Kappa	27%					19%				
	rFb_iqr	sensitivity	0%	80%	0%	56%	50%	0%	80%	0%	56%	17%
		specificity	100%	64%	100%	65%	96%	100%	63%	100%	59%	96%
		F	0%	62%	0%	45%	57%	0%	62%	0%	43%	24%
		Accuracy	47%					43%				
		Kappa	5%					19%				
	rFb_slope	sensitivity	0%	80%	0%	56%	25%	0%	63%	0%	37%	17%
		specificity	100%	64%	100%	61%	96%	100%	62%	100%	52%	92%
		F	0%	62%	0%	43%	33%	0%	51%	0%	29%	19%
		Accuracy	44%					32%				
		Kappa	21%					6%				
	rFb_max	sensitivity	0%	80%	0%	44%	25%	0%	80%	0%	44%	0%
		specificity	100%	55%	100%	65%	96%	100%	55%	100%	61%	96%
		F	0%	57%	0%	38%	33%	0%	57%	0%	36%	0%
		Accuracy	41%					38%				
		Kappa	16%					11%				

	rFb_var	sensitivity	0%	80%	0%	44%	25%	0%	80%	0%	44%	0%
		specificity	100%	55%	100%	65%	96%	100%	55%	100%	61%	96%
		F	0%	57%	0%	38%	33%	0%	57%	0%	36%	0%
		Accuracy	41%					38%				
		Kappa	16%					11%				
	rPS_mean	sensitivity	57%	60%	50%	11%	25%	29%	40%	0%	0%	0%
		specificity	76%	59%	97%	91%	96%	81%	41%	98%	71%	96%
		F	47%	48%	50%	17%	33%	29%	30%	0%	0%	0%
		Accuracy	41%					19%				
		Kappa	21%					0%				
	rPS_median	sensitivity	57%	60%	50%	22%	50%	29%	37%	0%	0%	17%
		specificity	80%	59%	100%	91%	96%	84%	35%	98%	78%	94%
		F	50%	48%	67%	31%	57%	31%	26%	0%	0%	21%
		Accuracy	47%					20%				
		Kappa	29%					0%				
	rTC_mean	sensitivity	0%	60%	0%	44%	50%	0%	53%	0%	44%	33%
		specificity	100%	64%	100%	57%	93%	100%	61%	100%	51%	95%
		F	0%	50%	0%	35%	50%	0%	44%	0%	33%	40%
		Accuracy	38%					33%				
		Kappa	14%					7%				
ADC	ADC_mean	sensitivity	86%	60%	0%	67%	0%	48%	67%	0%	33%	0%
		specificity	88%	77%	100%	74%	100%	89%	62%	100%	65%	100%
		F	75%	57%	0%	57%	0%	51%	53%	0%	30%	0%
		Accuracy	56%					41%				
		Kappa	40%					17%				
	ADC_SD	**sensitivity**	**71%**	**80%**	**0%**	**67%**	**0%**	**29%**	**83%**	**0%**	**59%**	**0%**
		specificity	**88%**	**73%**	**100%**	**83%**	**100%**	**89%**	**68%**	**100%**	**71%**	**100%**
		F	**67%**	**67%**	**0%**	**63%**	**0%**	**34%**	**66%**	**0%**	**51%**	**0%**
		Accuracy	**59%**					**49%**				
		Kappa	**44%**					**28%**				
	ADC_median	sensitivity	71%	60%	0%	67%	0%	48%	63%	0%	48%	0%
		specificity	88%	77%	100%	70%	100%	92%	67%	100%	62%	100%
		F	67%	57%	0%	55%	0%	54%	54%	0%	39%	0%
		Accuracy	53%					44%				
		Kappa	35%					21%				
	ADC_30tile	sensitivity	71%	60%	0%	67%	0%	48%	60%	0%	41%	0%

		specificity	84%	77%	100%	74%	100%	89%	62%	100%	65%	100%
		F	63%	57%	0%	57%	0%	51%	49%	0%	35%	0%
		Accuracy	53%					41%				
		Kappa	35%					17%				
	ADC_40tile	sensitivity	71%	60%	0%	67%	0%	48%	60%	0%	41%	0%
		specificity	88%	77%	100%	70%	100%	92%	64%	100%	61%	100%
		F	67%	57%	0%	55%	0%	54%	50%	0%	34%	0%
		Accuracy	53%					41%				
		Kappa	35%					17%				
	ADC_60tile	sensitivity	71%	70%	0%	56%	0%	48%	63%	0%	33%	0%
		specificity	88%	82%	100%	65%	100%	89%	67%	100%	59%	100%
		F	67%	67%	0%	45%	0%	51%	53%	0%	28%	0%
		Accuracy	53%					40%				
		Kappa	35%					16%				
	ADC_70tile	sensitivity	86%	70%	0%	56%	0%	29%	67%	0%	37%	0%
		specificity	84%	86%	100%	70%	100%	89%	70%	100%	54%	100%
		F	71%	70%	0%	48%	0%	34%	57%	0%	29%	0%
		Accuracy	56%					38%				
		Kappa	40%					13%				
	ADC_80tile	sensitivity	86%	70%	0%	56%	0%	38%	73%	0%	41%	0%
		specificity	88%	86%	100%	61%	100%	92%	71%	100%	57%	100%
		F	75%	70%	0%	45%	0%	46%	62%	0%	32%	0%
		Accuracy	53%					43%				
		Kappa	35%					20%				
	ADC_90tile	sensitivity	71%	60%	0%	56%	0%	38%	63%	0%	41%	0%
		specificity	88%	77%	100%	65%	100%	92%	62%	100%	61%	100%
		F	67%	57%	0%	45%	0%	46%	51%	0%	34%	0%
		Accuracy	50%					40%				
		Kappa	31%					15%				
	ADC_95tile	sensitivity	57%	70%	0%	44%	0%	29%	70%	0%	30	0%
		specificity	88%	73%	100%	65%	100%	92%	58%	100%	61%	100%
		F	57%	61%	0%	38%	0%	36%	53%	0%	26%	0%
		Accuracy	47%					36%				
		Kappa	26%					11%				
	ADC_skew	sensitivity	57%	70%	0%	67%	25%	29%	70%	0%	52%	0%
		specificity	84%	77%	100%	78%	100%	79%	76%	100%	71%	96%

		F	53%	64%	0%	60%	40%	28%	63%	0%	46%	0%
		Accuracy	56%					43%				
		Kappa	40%					21%				
	ADC_zscore	sensitivity	57%	80%	0%	56%	0%	14%	80%	0%	37%	0%
		specificity	84%	73%	100%	78%	100%	77%	73%	100%	65%	100%
		F	53%	67%	0%	53%	0%	15%	67%	0%	33%	0%
		Accuracy	53%					39%				
		Kappa	35%					14%				
	ADC_var	sensitivity	71%	80%	0%	56%	0%	38%	73%	0%	41%	0%
		specificity	84%	73%	100%	83%	100%	81%	71%	100%	68%	100%
		F	63%	67%	0%	56%	0%	37%	62%	0%	37%	0%
		Accuracy	56%					43%				
		Kappa	39%					21%				
	rADC_mean	sensitivity	86%	60%	0%	44%	0%	57%	67%	0%	22%	0%
		specificity	84%	82%	100%	65%	100%	89%	65%	100%	61%	100%
		F	71%	60%	0%	38%	0%	59%	55%	0%	20%	0%
		Accuracy	50%					40%				
		Kappa	31%					16%				
	rADC_SD	sensitivity	100%	80%	0%	67%	0%	33%	83%	0%	30%	0%
		specificity	88%	82%	100%	70%	100%	89%	56%	100%	72%	100%
		F	57%	73%	0%	55%	0%	39%	60%	0%	30%	0%
		Accuracy	56%					42%				
		Kappa	39%					18%				
	rADC_median	sensitivity	100%	70%	0%	44%	0%	67%	73%	0%	37%	0%
		specificity	80%	82%	100%	78%	100%	87%	73%	100%	68%	100%
		F	74%	67%	0%	44%	0%	62%	63%	0%	34%	0%
		Accuracy	56%					48%				
		Kappa	40%					28%				
	rADC_20tile	sensitivity	71%	60%	0%	56%	0%	48%	50%	0%	22%	0%
		specificity	84%	77%	100%	70%	100%	87%	59%	100%	59%	100%
		F	63%	57%	0%	48%	0%	49%	42%	0%	20%	0%
		Accuracy	50%					32%				
		Kappa	31%					45%				
	rADC_30tile	sensitivity	71%	60%	0%	56%	0%	48%	57%	0%	37%	0%
		specificity	84%	82%	100%	65%	100%	89%	68%	100%	57%	100%
		F	63%	60%	0%	45%	0%	51%	50%	0%	30%	0%

		Accuracy	50%					39%				
		Kappa	31%					14%				
	rADC_40tile	sensitivity	100%	70%	0%	56%	0%	48%	70%	0%	41%	0%
		specificity	84%	82%	100%	78%	100%	89%	70%	100%	62%	100%
		F	78%	67%	0%	53%	0%	51%	59%	0%	34%	0%
		Accuracy	59%					44%				
		Kappa	44%					22%				
	rADC_60tile	sensitivity	100%	70%	0%	56%	0%	67%	73%	0%	33%	0%
		specificity	80%	86%	100%	78%	100%	87%	71%	100%	68%	100%
		F	74%	70%	0%	53%	0%	62%	62%	0%	31%	0%
		Accuracy	59%					47%				
		Kappa	45%					26%				
	rADC_70tile	sensitivity	100%	60%	0%	56%	0%	48%	70%	0%	33%	0%
		specificity	88%	86%	100%	65%	100%	92%	71%	100%	55%	100%
		F	82%	63%	0%	45%	0%	54%	60%	0%	27%	0%
		Accuracy	56%					42%				
		Kappa	40%					19%				
	rADC_80tile	sensitivity	71%	70%	0%	56%	0%	38%	77%	0%	33%	0%
		specificity	84%	91%	100%	61%	100%	89%	74%	100%	55%	100%
		F	63%	74%	0%	43%	0%	43%	66%	0%	27%	0%
		Accuracy	53%					42%				
		Kappa	36%					19%				
	rADC_90tile	sensitivity	57%	80%	0%	67%	0%	38%	77%	0%	33%	0%
		specificity	92%	82%	100%	65%	100%	89%	65%	100%	64%	100%
		F	62%	73%	0%	52%	0%	43%	61%	0%	30%	0%
		Accuracy	56%					42%				
		Kappa	39%					18%				
	rADC_95tile	sensitivity	57%	70%	0%	67%	0%	38%	77%	0%	41%	0%
		specificity	88%	82%	100%	65%	100%	89%	65%	100%	67%	100%
		F	57%	67%	0%	52%	0%	43%	61%	0%	36%	0%
		Accuracy	53%					44%				
		Kappa	35%					21%				
	rADC_var	sensitivity	57%	80%	0%	67%	0%	33%	77%	0%	30%	0%
		specificity	88%	82%	100%	70%	100%	89%	52%	100%	74%	100%
		F	57%	73%	0%	55%	0%	39%	54%	0%	30%	0%
		Accuracy	56%					40%				

		Kappa	39%	15%

Table 7-B. Diagnostic performance of the combined features from ASL, DSC, DCE-mTK, DCE-L&L, ADC for glioma grading

MRI-method	Features combinations	Prediction performance	Grading		Internal validation LOOCV	
			LGGs	HGGs	LGGs	HGGs
ASL		sensitivity	84%	62%	79%	54%
Backwards-elimination	rTBF_mean_95tile_mode_var	specificity	62%	84%	54%	79%
		F	80%	67%	75%	58%
		Accuracy	75%		67%	
		Kappa	47%		34%	
	rTBF_mean_95tile_var	**sensitivity**	**89%**	**62%**	**74%**	**59%**
		specificity	**62%**	**89%**	**59%**	**74%**
		F	**83%**	**70%**	**73%**	**60%**
		Accuracy	**78%**		**68%**	
		Kappa	**53%**		**33%**	
Pairwise-correlation	**rTBF_iqr_max_min**	**sensitivity**	**89%**	**69%**	**84%**	**56%**
		specificity	**69%**	**89%**	**56%**	**84%**
		F	**85%**	**75%**	**79%**	**63%**
		Accuracy	**81%**		**73%**	
		Kappa	**60%**		**42%**	
DSC		sensitivity	100%	100%	84%	72%
Backwards-elimination	rCBV_mean_SD_95tile_median_iqr_slope_entropy_min_max_var	specificity	100%	100%	72%	84%
		F	100%	100%	83%	74%
		Accuracy	100%		79%	
		Kappa	100%		56%	
	rCBV_mean_SD_95tile_median_iqr_slope_entropy_max_var	sensitivity	100%	100%	84%	72%
		specificity	100%	100%	72%	84%
		F	100%	100%	83%	74%
		Accuracy	100%		79%	
		Kappa	100%		56%	
	rCBV_mean_SD_95tile_median_slope_entropy_max_var	sensitivity	100%	100%	84%	72%
		specificity	100%	100%	72%	84%
		F	100%	100%	83%	74%
		Accuracy	100%		79%	
		Kappa	100%		56%	
		sensitivity	100%	100%	86%	77%

		specificity	100%	100%	77%	86%
	rCBV_mean_SD_median_slope_entropy_max_var	F	100%	100%	85%	76%
		Accuracy	100%		82%	
		Kappa	100%		63%	
	rCBV_mean_SD_median_entropy_max_var	sensitivity	100%	92%	84%	74%
		specificity	92%	100%	74%	84%
		F	97%	96%	83%	75%
		Accuracy	97%		80%	
		Kappa	93%		59%	
	rCBV_mean_SD_median_entropy_var	**sensitivity**	**95%%**	**92%**	**84%**	**85%**
		specificity	**92%%**	**95%**	**85%**	**84%**
		F	**95%**	**92%**	**86%**	**81%**
		Accuracy	**94%**		**84%**	
		Kappa	**87%**		**68%**	
Pairwise-correlation	rCBF_kur_max_rCBV_min_var	sensitivity	89%	85%	93%	51%
		specificity	85%	89%	51%	93%
		F	89%	85%	82%	63%
		Accuracy	88%		76%	
		Kappa	74%		47%	
	rCBF_kur_rCBV_SD_min	sensitivity	89%	85%	86%	62%
		specificity	85%	89%	62%	86%
		F	89%	85%	81%	68%
		Accuracy	88%		76%	
		Kappa	74%		49%	
	rCBF_kur_rCBV_slope_min	sensitivity	89%	85%	74%	74%
		specificity	85%	89%	74%	74%
		F	89%	85%	77%	70%
		Accuracy	88%		74%	
		Kappa	74%		47%	
	rCBF_kur_rCBV_min_max	sensitivity	89%	85%	79%	72%
		specificity	85%	89%	72%	79%
		F	89%	85%	80%	71%
		Accuracy	88%		76%	
		Kappa	74%		51%	
	rCBF_zscore_rCBV_95tile	sensitivity	95%	77%	79%	74%
		specificity	77%	95%	74%	79%

		F	90%	83%	80%	73%
		Accuracy	88%		77%	
		Kappa	73%		53%	
		sensitivity	**89%**	**85%**	**86%**	**74%**
		specificity	**85%**	**89%**	**74%**	**86%**
	rCBV_iqr_min_max	**F**	**89%**	**85%**	**84%**	**76%**
		Accuracy	**88%**		**81%**	
		Kappa	**74%**		**61%**	
DCE_mTK		**sensitivity**	**84%**	**62%**	**84%**	**44%**
Backwards-elimination		**specificity**	**62%**	**84%**	**44%**	**84%**
	Ve_mode_rVe_mode_rVp_mean	**F**	**80%**	**67%**	**76%**	**52%**
		Accuracy	**75%**		**68%**	
		Kappa	**47%**		**29%**	
Pairwise-correlation		**sensitivity**	**89%**	**62%**	**84%**	**54%**
		specificity	**62%**	**89%**	**54%**	**84%**
	Ve_mode_rK_mean	**F**	**83%**	**70%**	**78%**	**61%**
		Accuracy	**78%**		**72%**	
		Kappa	**53%**		**40%**	
DCE_L&L		sensitivity	95%	77%	49%	21%
		specificity	77%	95%	21%	49%
	TC-SD_95tile_slope_var_rKep_max_rVp_mean_95tile_entropy_rFb_entropy_rTC_SD_95tile_kur_skew_iqr_slope_var	F	90%	83%	48%	21%
		Accuracy	88%		38%	
		Kappa	73%		31%	
		sensitivity	**79%**	**62%**	**46%**	**10%**
		specificity	**62%**	**79%**	**10%**	**46%**
	TC-SD_95tile_slope_var_rKep_max_rVp_95tile_entropy_rFb_entropy_rTC_SD_95tile_kur_skew_iqr_slope_var	**F**	**77%**	**64%**	**44**	**11%**
		Accuracy	**72%%**		**31%**	
		Kappa	**41%**		**45%**	
		sensitivity	79%	62%	53%	23%
		specificity	62%	79%	23%	53%
	TC-SD_95tile_slope_var_rKep_max_rVp_entropy_rFb_entropy_rTC_SD_95tile_kur_skew_slope	F	77%	64%	51%	24%
		Accuracy	72%		41%	
Backwards-elimination		Kappa	41%		25%	
		sensitivity	79%	69%	63%	38%
		specificity	69%	79%	38%	63%
Pairwise-correlation	TC_95tile_rK_slope_rKep_max_rVp_95tile_rPS_median	F	77%	69%	62%	40

		Accuracy	75%		53%	
		Kappa	48%		2%	
	TC_slope_rKep_max_rVp_95tile_rPS_mean	**sensitivity**	**79%**	**38%**	**70%**	**36%**
		specificity	**38%**	**79%**	**36%**	**70%**
		F	**71%**	**45%**	**66%**	**40%**
		Accuracy	**63%**		**56%**	
		Kappa	**18%**		**6%**	
ADC	**ADC-min_60tile_rADC-mean_min_30tile_40tile_80tile_90tile_95tile**	**sensitivity**	**89%**	**77%**	**77%**	**56%**
Backwards-elimination		**specificity**	**77%**	**89%**	**56%**	**77%**
		F	**87%**	**80%**	**75%**	**59%**
		Accuracy	**84%**		**69%**	
		Kappa	**67%**		**34%**	
Pairwise-correlation	rADC_median_min_max_var_zscore	sensitivity	89%	77%	79%	51%
		specificity	77%	89%	51%	79%
		F	87%	80%	74%	56%
		Accuracy	84%		68%	
		Kappa	67%		31%	
	rADC_median_min_max_zscore	**sensitivity**	**89%**	**77%**	**79%**	**56%**
		specificity	**77%**	**89%**	**56%**	**79%**
		F	**87%**	**80%**	**76%**	**60%**
		Accuracy	**84%**		**70%**	
		Kappa	**67%**		**46%**	
ASL&ADC	ADC-min_60tile_rADC-mean_median_min_30tile_40tile_60tile_70tile_80tile_90tile_95tile_rTBF_mean_SD_95tile_median_iqr_mode_slope_entropy_max_var	sensitivity	95%	85%	70%	54%
Backwards-elimination		specificity	85%	95%	54%	70%
		F	92%	88%	70%	56%
		Accuracy	91%		64%	
		Kappa	80%		24%	
	ADC-min_rADC-mean_median_min_30tile_40tile_60tile_70tile_80tile_90tile_95tile_rTBF_mean_SD_95tile_median_iqr_mode_slope_entropy_max_var	sensitivity	89%	85%	70%	51%
		specificity	85%	89%	51%	70%
		F	89%	85%	69%	53%
		Accuracy	89%		63%	
		Kappa	74%		22%	
	ADC-min_rADC-mean_median_min_30tile_40tile_60tile_70tile_80tile_90tile_95tile_rTBF_mean_SD_95tile_median_iqr_mode_entropy_max_var	sensitivity	89%	85%	72%	54%
		specificity	85%	89%	54%	72%
		F	89%	84%	71%	55%

		Accuracy	88%		65%	
		Kappa	74%		26%	
	ADC-min_rADC-mean_median_min_30tile_40tile_60tile_70tile_80tile_90tile_95tile_rTBF_mean_SD_95tile_median_iqr_mode_entropy_max	sensitivity	89%	85%	72%	54%
		specificity	85%	89%	54%	72%
		F	89%	85%	71%	55%
		Accuracy	88%		65%	
		Kappa	74%		26%	
	ADC-min_rADC-mean_median_min_30tile_40tile_60tile_70tile_80tile_90tile_95tile_rTBF_mean_SD_95tile_median_mode_entropy_max	sensitivity	89%	85%	72%	54%
		specificity	85%	89%	54%	72%
		F	89%	85%	71%	55%
		Accuracy	88%		65%	
		Kappa	74%		26%	
	ADC-min_rADC-mean_median_min_30tile_40tile_60tile_70tile_80tile_95tile_rTBF_mean_SD_95tile_median_mode_entropy_max	sensitivity	95%	85%	74%	56%
		specificity	85%	95%	56%	74%
		F	92%	88%	72%	58%
		Accuracy	91%		67%	
		Kappa	80%		30%	
	ADC-min_rADC-mean_median_min_40tile_60tile_70tile_80tile_95tile_rTBF_mean_SD_95tile_median_mode_entropy_max	sensitivity	89%	85%	74%	56%
		specificity	85%	89%	56%	74%
		F	89%	85%	72%	58%
		Accuracy	88%		67%	
		Kappa	74%		30%	
	ADC-min_rADC-mean_median_min_40tile_60tile_70tile_80tile_95tile_rTBF_mean_SD_95tile_median_mode_max	**sensitivity**	**89%**	**85%**	**77%**	**56%**
		specificity	**85%**	**89%**	**56%**	**77%**
		F	**89%**	**85%**	**75%**	**59%**
		Accuracy	**88%**		**67%**	
		Kappa	**74%**		**34%**	
	ADC-min_rADC-mean_median_min_40tile_60tile_70tile_80tile_95tile_rTBF_mean_95tile_median_mode_max	sensitivity	95%	85%	72%	54%
		specificity	85%	95%	54%	72%
		F	92%	88%	71%	55%
		Accuracy	90%		65%	
		Kappa	80%		26%	

Table 8-B. Diagnostic performance of the combined features from ASL, DSC, DCE-mTK, DCE-L&L, and ADC for glioma sub-grading

MRI-method	Features combinations	Prediction performance	Sub-grading			Internal validation LOOCV		
			II	III	IV	II	III	IV
ASL	**rTBF-mean_95tile_median_max**	**sensitivity**	**84%**	**67%**	**71%**	**84%**	**11%**	**33%**
Backwards-elimination		**specificity**	**69%**	**96%**	**92%**	**36%**	**95%**	**87%**
		F	**82%**	**73%**	**72%**	**74%**	**17%**	**37%**
		Accuracy	**78%**			**59%**		
		Kappa	**60%**			**18%**		
Pairwise-correlation	**rTBF_SD_skew_mode_min**	**sensitivity**	**89%**	**50%**	**57%**	**72%**	**22%**	**38%**
		specificity	**62%**	**96%**	**92%**	**41%**	**92%**	**81%**
		F	**83%**	**60%**	**62%**	**68%**	**29%**	**37%**
		Accuracy	**75%**			**55%**		
		Kappa	**53%**			**16%**		
DSC	rCBV-mean_SD_95tile_median_iqr_slope_entropy_min_max_var	sensitivity	100%	100%	100%	82%	28%	57%
		specificity	100%	100%	100%	67%	88%	87%
		F	100%	100%	100%	80%	31%	56%
		Accuracy	100%			67%		
		Kappa	100%			40%		
	rCBV-mean_SD_95tile_median_slope_entropy_min_max_var	sensitivity	100%	100%	100%	82%	33%	62%
		specificity	100%	100%	100%	74%	87%	87%
		F	100%	100%	100%	82%	35%	59%
		Accuracy	100%			67%		
		Kappa	100%			45%		
	rCBV-mean_95tile_median_slope_entropy_min_max_var	sensitivity	100%	100%	100%	82%	28%	57%
		specificity	100%	100%	100%	72%	87%	85%
		F	100%	100%	100%	82%	30%	55%
		Accuracy	100%			67%		
		Kappa	100%			40%		
Backwards-elimination	**rCBV-mean_median_slope_entropy_min_max_var**	**sensitivity**	**100%**	**100%**	**100%**	**88%**	**28%**	**67%**
		specificity	**100%**	**100%**	**100%**	**69%**	**91%**	**89%**
		F	**100%**	**100%**	**100%**	**84%**	**33%**	**65%**
		Accuracy	**100%**			**72%**		
		Kappa	**100%**			**48%**		

	rCBV-mean_median_slope_entropy_max_var	sensitivity	100%	100%	100%	87%	28%	67%
		specificity	100%	100%	100%	69%	91%	89%
		F	100%	100%	100%	84%	33%	65%
		Accuracy	100%			72%		
		Kappa	100%			48%		
	rCBV-mean_median_slope_entropy_max	sensitivity	100%	67%	57%	88%	11%	52%
		specificity	85%	92%	96%	72%	91%	80%
		F	95%	67%	67%	85%	15%	47%
		Accuracy	84%			66%		
		Kappa	71%			14%		
Pairwise-correlation	**rCBF_kur_max_rCBV_min_var**	**sensitivity**	**95%**	**67%**	**86%**	**93%**	**22%**	**52%**
		specificity	**85%**	**92%**	**100%**	**54%**	**95%**	**92%**
		F	**92%**	**67%**	**92%**	**83%**	**31%**	**58%**
		Accuracy	**88%**			**71%**		
		Kappa	**77%**			**42%**		
DCE_mTK	**Ve-mode_rVe_mode**	**sensitivity**	**95%**	**33%**	**14%**	**91%**	**17%**	**0%**
Backwards-elimination		**specificity**	**38%**	**96%**	**92%**	**18%**	**91%**	**97%**
		F	**80%**	**44%**	**20%**	**74%**	**21%**	**0%**
		Accuracy	**66%**			**57%**		
		Kappa	**28%**			**6%**		
Pairwise-correlation	**Ve_mode_rK_mean**	**sensitivity**	**89%**	**33%**	**57%**	**91%**	**17%**	**19%**
		specificity	**54%**	**100%**	**88%**	**33%**	**96%**	**89%**
		F	**81%**	**50%**	**57%**	**77%**	**25%**	**24%**
		Accuracy	**72%**			**61%**		
		Kappa	**24%**			**19%**		
DCE_L&L	**rVp-95tile_slope_max_rFb-max_rTC_mode_min**	**sensitivity**	**84%**	**67%**	**57%**	**72%**	**39%**	**10%**
Backwards-elimination		**specificity**	**62%**	**96%**	**92%**	**33%**	**91%**	**83%**
		F	**80%**	**72%**	**62%**	**66%**	**44%**	**11%**
		Accuracy	**75%**			**52%**		
		Kappa	**54%**			**9%**		
Pairwise-correlation	TC_SD_rK_median_rKep_max_rFb_slope_rPS_mean	sensitivity	95%	67%	57%	86%	28%	14%
		specificity	77%	96%	92%	41%	92%	87%
		F	90%	72%	62%	76%	34%	18%
		Accuracy	81%			59%		
		Kappa	65%			19%		
		sensitivity	**95%**	**67%**	**57%**	**88%**	**28%**	**14%**

	TC_95tile_rK_median_rKep_max_rFb_slope_rPS_mean	**specificity**	**77%**	**96%**	**92%**	**41%**	**94%**	**87%**
		F	**90%**	**72%**	**62%**	**77%**	**36%**	**18%**
		Accuracy	**81%**			**60%**		
		Kappa	**65%**			**21%**		
	TC_slope_rK_median_rKep_max_rFb_slope_rPS_mean	sensitivity	95%	77%	57%	89%	22%	14%
		specificity	77%	96%	92%	41%	94%	87%
		F	90%	73%	62%	78%	30%	18%
		Accuracy	81%			60%		
		Kappa	65%			20%		
ADC	ADC-median_min_60tile_80tile_rADC-mean_median_min_30tile_40tile_60tile_70tile_90tile	sensitivity	100%	50%	86%	72%	39%	19%
		specificity	85%	96%	96%	54%	85%	81%
		F	95%	60%	86%	71%	38%	21%
		Accuracy	88%			54%		
		Kappa	77%			18%		
	ADC-median_min_60tile_rADC-mean_median_min_30tile_40tile_60tile_70tile_90tile	**sensitivity**	**100%**	**50%**	**86%**	**82%**	**39%**	**19%**
		specificity	**85%**	**96%**	**96%**	**54%**	**88%**	**85%**
		F	**95%**	**60%**	**86%**	**77%**	**41%**	**22%**
		Accuracy	**88%**			**61%**		
		Kappa	**77%**			**26%**		
	ADC-median_min_60tile_rADC-mean_median_min_30tile_40tile_60tile_70tile	sensitivity	100%	50%	57%	81%	39%	19%
		specificity	77%	96%	92%	54%	90%	83%
		F	93%	60%	62%	76%	42%	21%
		Accuracy	81%			59%		
		Kappa	64%			24%		
	ADC-median_min_60tile_rADC-mean_median_min_30tile_40tile_60tile	sensitivity	100%	50%	43%	82%	39%	19%
		specificity	69%	92%	96%	54%	91%	83%
		F	90%	55%	55%	77%	44%	21%
		Accuracy	78%			60%		
		Kappa	58%			26%		
	ADC-median_60tile_rADC-mean_median_min_30tile_40tile_60tile	sensitivity	100%	50%	43%	84%	22%	19%
		specificity	69%	92%	96%	46%	91%	84%
		F	90%	55%	55%	76%	28%	22%
		Accuracy	78%			58%		
		Kappa	58%			19%		
Backwards-elimination								
Pairwise-correlation	**rADC_median_min_max_var_zscore**	**sensitivity**	**89%**	**67%**	**71%**	**81%**	**33%**	**29%**
		specificity	**69%**	**96%**	**96%**	**51%**	**88%**	**87%**

		F	**85%**	**72%**	**77%**	**75%**	**36%**	**32%**
		Accuracy	**81%**			**60%**		
		Kappa	**65%**			**26%**		
	rADC_median_min_max_zscore	sensitivity	89%	67%	71%	77%	33%	29%
		specificity	69%	96%	96%	51%	88%	84%
		F	85%	72%	77%	73%	36%	31%
		Accuracy	81%			58%		
		Kappa	56%			23%		

Table 9-B.Diagnostic performance of the combined features from ASL, DSC, DCE-mTK, DCE-L&L, and ADC for glioma identification according to IDH-status.

MRI-method	Features combinations	Prediction performance	IDH_status		Internal validation LOOCV	
			wt	mut	wt	mut
ASL	**rTBF_zscore_95tile_median_mode_min_max**	**sensitivity**	**71%**	**100%**	**48%**	**81%**
Backwards-elimination		**specificity**	**100%**	**71%**	**81%**	**48%**
		F	**83%**	**96%**	**44%**	**83%**
		Accuracy	**94%**		**74%**	
		Kappa	**80%**		**28%**	
Pairwise-correlation	**rTBF_mean_slope_min**	**sensitivity**	**71%**	**92%**	**48%**	**91%**
		specificity	**92%**	**71%**	**91%**	**48%**
		F	**71%**	**92%**	**53%**	**88%**
		Accuracy	**88%**		**81%**	
		Kappa	**63%**		**41%**	
DSC	rCBF_mean_skew_median_rCBV_mean_SD_95tile_median_iqr_mode_slope_entropy_min_max_var	sensitivity	100%	100%	52%	84%
		specificity	100%	100%	84%	52%
		F	100%	100%	50%	85%
		Accuracy	100%		77%	
		Kappa	100%		35%	
	rCBF_mean_skew_median_rCBV_mean_95tile_median_iqr_mode_slope_entropy_min_max_var	sensitivity	100%	100%	52%	84%
		specificity	100%	100%	84%	52%
		F	100%	100%	50%	85%
		Accuracy	100%		77%	
		Kappa	100%		35%	
	rCBF_mean_skew_median_rCBV_mean_95tile_median_iqr_slope_entropy_min_max_var	sensitivity	100%	100%	62%	83%
		specificity	100%	100%	83%	62%
		F	100%	100%	55%	86%
		Accuracy	100%		78%	
		Kappa	100%		41%	
Backwards-elimination	**rCBF_mean_skew_median_rCBV_mean_95tile_median_slope_entropy_min_max_var**	**sensitivity**	**100%**	**100%**	**62%**	**83%**
		specificity	**100%**	**100%**	**83%**	**62%**
		F	**100%**	**100%**	**55%**	**86%**
		Accuracy	**100%**		**78%**	
		Kappa	**100%**		**41%**	
		sensitivity	100%	100%	52%	87%

	rCBF_mean_skew_median_rCBV_mean_95tile_ median_entropy_min_max_var	specificity	100%	100%	87%	52%
		F	100%	100%	52%	87%
		Accuracy	100%		79%	
		Kappa	100%		39%	
	rCBF_mean_skew_median_rCBV_mean_95tile_ median_entropy_min_max	sensitivity	86%	100%	52%	85%
		specificity	100%	86%	85%	52%
		F	92%	98%	51%	86%
		Accuracy	97%		78%	
		Kappa	90%		37%	
	rCBF_mean_median_rCBV_mean_95tile_median _entropy_min_max	sensitivity	86%	100%	48%	87%
		specificity	100%	86%	87%	48%
		F	92%	98%	49%	86%
		Accuracy	97%		78%	
		Kappa	90%		35%	
	rCBF_mean_median_rCBV_mean_95tile_median _entropy_min	sensitivity	100%	100%	38%	88%
		specificity	100%	100%	88%	38%
		F	100%	100%	42%	86%
		Accuracy	100%		77%	
		Kappa	100%		28%	
	rCBF_mean_median_rCBV_mean_95tile_median _min	sensitivity	71%	100%	48%	89%
		specificity	100%	71%	89%	48%
		F	83%	96%	51%	88%
		Accuracy	94%		80%	
		Kappa	80%		39%	
	rCBF_mean_median_rCBV_mean_95tile_min	sensitivity	71%	92%	38%	91%
		specificity	92%	71%	91%	38%
		F	71%	92%	44%	87%
		Accuracy	88%		79%	
		Kappa	63%		32%	
Pairwise-correlation	rCBF_kur_mode_rCBV_mean_min	sensitivity	86%	96%	48%	91%
		specificity	96%	86%	91%	48%
		F	86%	96%	53%	88%
		Accuracy	94%		81%	
		Kappa	82%		41%	
	rCBF_kur_rCBV_mean_min	sensitivity	86%	96%	48%	92%
		specificity	96%	86%	92%	48%

		F	86%	96%	54%	89%
		Accuracy	94%		82%	
		Kappa	82%		43%	
	rCBF_kur_rCBV_slope_min	**sensitivity**	**86%**	**96%**	**67%**	**92%**
		specificity	**96%**	**86%**	**92%**	**67%**
		F	**86%**	**96%**	**68%**	**91%**
		Accuracy	**94%**		**86%**	
		Kappa	**82%**		**60%**	
	rCBF_kur_rCBV_min_max	sensitivity	86%	96%	57%	92%
		specificity	96%	86%	92%	57%
		F	86%	96%	62%	90%
		Accuracy	94%		84%	
		Kappa	82%		52%	
	rCBF_kur_mode_rCBV_min_max	sensitivity	86%	96%	57%	89%
		specificity	96%	86%	89%	57%
		F	86%	96%	59%	89%
		Accuracy	94%		82%	
		Kappa	82%		47%	
DCE_mTK	**rK_mean_rVp_min**	**sensitivity**	**43%**	**92%**	**10%**	**92%**
Pairwise-correlation		**specificity**	**92%**	**43%**	**92%**	**10%**
		F	**50%**	**88%**	**14%**	**85%**
		Accuracy	**81%**		**74%**	
		Kappa	**39%**		**2%**	
DCE_L&L	rK_mean_SD_95tile_median_iqr_slope_var_rPS_mean_median_rTC_mode_min	sensitivity	86%	96%	33%	85%
Backwards-elimination		specificity	96%	86%%	85%	33%
		F	86%	96%	36%	84%
		Accuracy	94%		74%	
		Kappa	82%		20%	
	rK_mean_95tile_median_iqr_slope_var_rPS_mean_median_rTC_mode_min	sensitivity	86%	96%	33%	81%
		specificity	96%	86%%	81%	33%
		F	86%	96%	33%	81%
		Accuracy	94%		71%	
		Kappa	82%		15%	
	rK_95tile_median_iqr_slope_var_rPS_mean_median_rTC_mode_min	sensitivity	71%	96%	38%	81%
		specificity	96%	71%	81%	38%
		F	77%	94%	37%	82%

		Accuracy	91%		72%	
		Kappa	71%		19%	
	rK_95tile_median_iqr_slope_var_rPS_median_rTC_mode_min	**sensitivity**	**71%**	**96%**	**29%**	**89%**
		specificity	**96%**	**71%**	**89%**	**29%**
		F	**77%**	**94%**	**34%**	**85%**
		Accuracy	**91%**		**76%**	
		Kappa	**71%**		**20%**	
	rK_95tile_median_iqr_slope_var_rTC_mode_min	sensitivity	71%	96%	19%	85%
		specificity	96%	71%	85%	19%
		F	77%	94%	22%	82%
		Accuracy	91%		71%	
		Kappa	71%		5%	
Pairwise-correlation	**TC_var_rK_slope_rKep_max_rVp_mean_rPS_median**	**sensitivity**	**57%**	**100%**	**57%**	**88%**
		specificity	**100%**	**57%**	**88%**	**57%**
		F	**73%**	**94%**	**57%**	**88%**
		Accuracy	**91%**		**81%**	
		Kappa	**68%**		**45%**	
ADC	ADC_mean_SD_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_zscore_var	sensitivity	100%	100%	29%	67%
		specificity	100%	100%	67%	29%
		F	100%	100%	23%	71%
		Accuracy	100%		58%	
		Kappa	100%		0%	
	ADC_mean_SD_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_zscore	sensitivity	100%	100%	29%	71%
		specificity	100%	100%	71%	29%
		F	100%	100%	24%	74%
		Accuracy	100%		61%	
		Kappa	100%		0%	
	ADC_mean_SD_median_30tile_40tile_60tile_70tile_80tile_90tile_95tile_zscore	sensitivity	100%	100%	38%	73%
		specificity	100%	100%	73%	38%
		F	100%	100%	33%	77%
		Accuracy	100%		66%	
		Kappa	100%		10%	
Backwards-elimination	ADC_mean_SD_median_30tile_40tile_60tile_70tile_95tile_zscore	sensitivity	100%	100%	38%	77%
		specificity	100%	100%	77%	38%
		F	100%	100%	35%	79%
		Accuracy	100%		69%	

		Kappa	100%		14%	
	ADC_mean_SD_median_30tile_40tile_70tile_95tile_zscore	**sensitivity**	**100%**	**100%**	**48%**	**83%**
		specificity	**100%**	**100%**	**83%**	**48%**
		F	**100%**	**100%**	**45%**	**84%**
		Accuracy	**100%**		**75%**	
		Kappa	**100%**		**29%**	
	rADC_mean_SD_median_10tile_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_var_zscore	sensitivity	71%	96%	23%	89%
		specificity	96%	71%	89%	23%
		F	77%	94%	29%	85%
		Accuracy	91%		75%	
		Kappa	71%		15%	
	rADC_mean_SD_median_10tile_20tile_30tile_40tile_60tile_70tile_80tile_95tile_var_zscore	sensitivity	71%	96%	19%	89%
		specificity	96%	71%	89%	19%
		F	77%	94%	24%	84%
		Accuracy	91%		74%	
		Kappa	71%		10%	
	rADC_mean_SD_median_10tile_30tile_40tile_60tile_70tile_80tile_95tile_var_zscore	sensitivity	57%	100%	19%	88%
		specificity	100%	57%	88%	19%
		F	73%	94%	24%	84%
		Accuracy	91%		73%	
		Kappa	68%		8%	
	rADC_mean_median_10tile_30tile_40tile_60tile_70tile_80tile_95tile_var_zscore	sensitivity	43%	88%	19%	91%
		specificity	88%	43%	91%	19%
		F	46%	86%	25%	85%
		Accuracy	78%		75%	
		Kappa	33%		12%	
	rADC_mean_median_10tile_30tile_40tile_70tile_80tile_95tile_var_zscore	sensitivity	57%	88%	19%	91%
		specificity	88%	57%	91%	19%
		F	57%	88%	25%	85%
		Accuracy	81%		75%	
		Kappa	45%		12%	
	rADC_mean_median_10tile_30tile_40tile_80tile_95tile_var_zscore	sensitivity	57%	96%	19%	92%
		specificity	96%	57%	92%	19%
		F	67%	92%	26%	86%
		Accuracy	88%		76%	
		Kappa	59%		14%	

	rADC_mean_median_10tile_30tile_40tile_95tile_var_zscore	sensitivity	43%	96%	19%	89%
		specificity	96%	43%	89%	19%
		F	55%	91%	24%	84%
		Accuracy	84%		74%	
		Kappa	46%		10%	
Pairwise-correlation	**rADC_median_min_max_var_zscore**	**sensitivity**	**57%**	**96%**	**29%**	**92%**
		specificity	**96%**	**57%**	**92%**	**29%**
		F	**67%**	**92%**	**36%**	**87%**
		Accuracy	**88%**		**78%**	
		Kappa	**59%**		**24%**	

Table 10-B. Diagnostic performance of the combined features from ASL, DSC, DCE-mTK, DCE-L&L, and ADC for glioma grouping.

MRI-method	Features combinations	Prediction performance	Grouping			Internal validation LOOCV		
			GBM-wt	Astro-mut	Oligo-mut	GBM-wt	Astro-mut	Oligo-mut
ASL	**rTBF-skew_zscore_mean_95tile_median_mode_min_max**	**sensitivity**	**71%**	**92%**	**77%**	**48%**	**75%**	**31%**
Backwards-elimination		**specificity**	**96%**	**90%**	**84%**	**79%**	**78%**	**68%**
		F	**77%**	**88%**	**77%**	**43%**	**71%**	**35%**
		Accuracy	**81%**			**51%**		
		Kappa	**71%**			**26%**		
Pairwise-correlation	**rTBF_SD_min_zscore**	**sensitivity**	**57%**	**92%**	**77%**	**0%**	**89%**	**46%**
		specificity	**96%**	**90%**	**79%**	**87%**	**80%**	**58%**
		F	**67%**	**88%**	**74%**	**0%**	**80%**	**44%**
		Accuracy	**78%**			**52%**		
		Kappa	**66%**			**24%**		
DSC	rCBF-mean_SD_95tile_kur_skew_median_iqr_mode_slope_max_zscore_var_rCBV-mean_SD_95tile_skew_median_iqr_mode_slope_entropy_min_max_var	sensitivity	100%	100%	100%	76%	67%	54%
		specificity	100%	100%	100%	92%	77%	74%
		F	100%	100%	100%	74%	65%	56%
		Accuracy	100%			64%		
		Kappa	100%			44%		
	rCBF-mean_SD_95tile_kur_skew_median_iqr_mode_max_zscore_var_rCBV-mean_SD_95tile_skew_median_iqr_mode_slope_entropy_min_max_var	sensitivity	100%	100%	100%	71%	72%	54%
		specificity	100%	100%	100%	95%	65%	74%
		F	100%	100%	100%	75%	68%	56%
		Accuracy	100%			65%		
		Kappa	100%			45%		
	rCBF-mean_SD_95tile_kur_skew_median_iqr_mode_max_zscore_var_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_min_max_var	**sensitivity**	**100%**	**100%**	**100%**	**81%**	**69%**	**51%**
		specificity	**100%**	**100%**	**100%**	**92%**	**78%**	**74%**
		F	**100%**	**100%**	**100%**	**77%**	**68%**	**54%**
		Accuracy	**100%**			**65%**		
		Kappa	**100%**			**46%**		
	rCBF-mean_SD_95tile_skew_median_iqr_mode_max_zscore_var_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_min_max_var	sensitivity	100%	100%	100%	62%	58%	56%
		specificity	100%	100%	100%	92%	77%	65%
		F	100%	100%	100%	65%	59%	54%
		Accuracy	100%			58%		
		Kappa	100%			35%		
Backwards-elimination	rCBF-mean_SD_95tile_skew_median_iqr_max_zs	sensitivity	100%	100%	100%	67%	61%	62%

	core_var_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_min_max_var	specificity	100%	100%	100%	93%	77%	70%
		F	100%	100%	100%	70%	61%	60%
		Accuracy	100%			63%		
		Kappa	100%			42%		
	rCBF-mean_SD_95tile_skew_median_iqr_max_zscore_var_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_max_var	sensitivity	100%	100%	100%	67%	61%	62%
		specificity	100%	100%	100%	93%	77%	70%
		F	100%	100%	100%	70%	61%	60%
		Accuracy	100%			63%		
		Kappa	100%			42%		
	rCBF-mean_SD_95tile_skew_median_iqr_max_zscore_var_rCBV-mean_95tile_median_iqr_mode_slope_entropy_max_var	sensitivity	100%	100%	100%	67%	61%	62%
		specificity	100%	100%	100%	93%	77%	70%
		F	100%	100%	100%	70%	61%	60%
		Accuracy	100%			63%		
		Kappa	100%			42%		
	rCBF-mean_SD_95tile_skew_median_iqr_max_zscore_var_rCBV-mean_95tile_median_iqr_mode_slope_entropy_max	sensitivity	100%	100%	100%	67%	61%	62%
		specificity	100%	100%	100%	93%	77%	70%
		F	100%	100%	100%	70%	61%	60%
		Accuracy	100%			63%		
		Kappa	100%			42%		
	rCBF-mean_95tile_skew_median_iqr_max_zscore_var_rCBV-mean_95tile_median_iqr_mode_slope_entropy_max	sensitivity	100%	100%	100%	67%	61%	62%
		specificity	100%	100%	100%	93%	77%	70%
		F	100%	100%	100%	70%	61%	60%
		Accuracy	100%			63%		
		Kappa	100%			42%		
	rCBF-mean_95tile_skew_median_iqr_max_zscore_rCBV-mean_95tile_median_iqr_mode_slope_entropy_max	sensitivity	100%	100%	100%	67%	61%	62%
		specificity	100%	100%	100%	93%	77%	70%
		F	100%	100%	100%	70%	61%	60%
		Accuracy	100%			63%		
		Kappa	100%			42%		
	rCBF-mean_95tile_skew_median_iqr_max_zscore_rCBV-mean_95tile_median_iqr_mode_entropy_max	sensitivity	100%	100%	100%	71%	61%	64%
		specificity	100%	100%	100%	93%	77%	74%
		F	100%	100%	100%	73%	61%	63%
		Accuracy	100%			65%		
		Kappa	100%			45%		
	rCBF-mean_95tile_skew_median_iqr_max_zscore	sensitivity	100%	100%	100%	62%	61%	67%
		specificity	100%	100%	100%	93%	78%	70%

	_rCBV-mean_95tile_median_mode_entropy_max	F	100%	100%	100%	67%	62%	63%
		Accuracy	100%			64%		
		Kappa	100%			41%		
	rCBF-mean_95tile_median_iqr_max_zscore_rCBV-mean_95tile_median_mode_entropy_max	sensitivity	100%	100%	100%	67%	64%	54%
		specificity	100%	100%	100%	92%	77%	68%
		F	100%	100%	100%	68%	63%	54%
		Accuracy	100%			61%		
		Kappa	100%			39%		
	rCBF-mean_95tile_median_max_zscore_rCBV-mean_95tile_median_mode_entropy_max	sensitivity	100%	100%	100%	57%	67%	54%
		specificity	100%	100%	100%	92%	77%	67%
		F	100%	100%	100%	62%	65%	53%
		Accuracy	100%			59%		
		Kappa	100%			37%		
	rCBF-mean_95tile_median_max_rCBV-mean_95tile_median_mode_entropy_max	sensitivity	100%	100%	100%	52%	72%	56%
		specificity	100%	100%	100%	93%	77%	68%
		F	100%	100%	100%	59%	68%	56%
		Accuracy	100%			61%		
		Kappa	100%			40%		
Pairwise-correlation	**rCBF_kur_rCBV_mean_min**	**sensitivity**	**86%**	**83%**	**85%**	**57%**	**81%**	**56%**
		specificity	**96%**	**95%**	**84%**	**92%**	**78%**	**75%**
		F	**86%**	**87%**	**81%**	**62%**	**74%**	**59%**
		Accuracy	**84%**			**66%**		
		Kappa	**76%**			**47%**		
DCE_mTK	**rVe_mode_rVp_median_min**	**sensitivity**	**43%**	**92%**	**77%**	**0%**	**83%**	**51%**
Backwards-elimination		**specificity**	**96%**	**80%**	**84%**	**92%**	**63%**	**68%**
		F	**55%**	**81%**	**77%**	**0%**	**68%**	**52%**
		Accuracy	**75%**			**52%**		
		Kappa	**60%**			**23%**		
Pairwise-correlation	**rK_mean_rVp_min**	**sensitivity**	**43%**	**83%**	**62%**	**19%**	**64%**	**46%**
		specificity	**88%**	**75%**	**84%**	**89%**	**63%**	**63%**
		F	**46%**	**74%**	**67%**	**24%**	**57%**	**46%**
		Accuracy	**66%**			**47%**		
		Kappa	**47%**			**16%**		
DCE_L&L	rK-mean_SD_iqr_var_rFb-mean_SD_95tile_iqr_slope_min_max_var_rTC-mean_mode_min	sensitivity	100%	100%	100%	19%	67%	69%
Backwards-elimination		specificity	100%	100%	100%	89%	70%	74%
		F	100%	100%	100%	24%	62%	67%

		Accuracy	100%			57%		
		Kappa	100%			32%		
	rK-mean_iqr_var_rFb-mean_SD_95tile_iqr_slope_min_max_var_rTC-mean_mode_min	sensitivity	100%	100%	100%	24%	61%	72%
		specificity	100%	100%	100%	87%	73%	74%
		F	100%	100%	100%	28%	59%	68%
		Accuracy	100%			57%		
		Kappa	100%			33%		
	rK-mean_iqr_rFb-mean_SD_95tile_iqr_slope_min_max_var_rTC-mean_mode_min	sensitivity	100%	100%	92%	29%	61%	72%
		specificity	100%	95% %	100%	85%	75%	75%
		F	100%	96%	96%	32%	60%	69%
		Accuracy	97%			58%		
		Kappa	95%			35%		
	rK-mean_iqr_rFb-mean_SD_95tile_iqr_slope_max_var_rTC-mean_mode_min	**sensitivity**	**100%**	**100%**	**100%**	**33%**	**64%**	**72%**
		specificity	**100%**	**100%**	**100%**	**87%**	**77%**	**75%**
		F	**100%**	**100%**	**100%**	**37%**	**63%**	**69%**
		Accuracy	**100%**			**60%**		
		Kappa	**100%**			**38%**		
	rK-mean_iqr_rFb-mean_SD_95tile_iqr_slope_max_var_rTC-mode_min	sensitivity	100%	100%	100%	29%	61%	67%
		specificity	100%	100%	100%	84%	70%	79%
		F	100%	100%	100%	31%	58%	68%
		Accuracy	100%			56%		
		Kappa	100%			32%		
	rK-mean_iqr_rFb-mean_SD_95tile_iqr_slope_max_var_rTC-min	sensitivity	86%	83%	92%	19%	61%	67%
		specificity	88%	95%	100%	84%	68%	77%
		F	75%	87%	96%	22%	57%	67%
		Accuracy	88%			54%		
		Kappa	81%			28%		
	rK-mean_iqr_rFb-mean_95tile_iqr_slope_max_var_rTC-min	sensitivity	86%	83%	92%	19%	61%	67%
		specificity	92%	90%	100%	85%	67%	77%
		F	80%	83%	96%	22%	56%	67%
		Accuracy	88%			54%		
		Kappa	81%			28%		
	rK-mean_iqr_rFb-mean_iqr_slope_max_var_rTC-min	sensitivity	86%	83%	92%	23%	61%	72%
		specificity	88%	95%	100%	85%	72%	77%
		F	75%	87%	96%	27%	59%	70%
		Accuracy	88%			57%		

		Kappa	81%			33%		
	rK-mean_iqr_rFb-mean_slope_max_var_rTC-min	sensitivity	86%	83%	92%	29%	67%	67%
		specificity	88%	95%	100%	87%	70%	79%
		F	75%	87%	96%	32%	62%	68%
		Accuracy	88%			58%		
		Kappa	81%			35%		
Pairwise-correlation	rK_slope_rFb_SD_rPS_median_rTC_kur_rTC_mode_rTC_entropy	sensitivity	71%	67%	77%	43%	31%	64%
		specificity	88%	75%	95%	77%	68%	74%
		F	67%	64%	83%	38%	33%	63%
		Accuracy	72%			47%		
		Kappa	57%			19%		
	rK_slope_rFb_var_rPS_median_rTC_kur_rTC_mode	**sensitivity**	**57%**	**75%**	**69%**	**33%**	**53%**	**64%**
		specificity	**92%**	**70%**	**89%**	**83%**	**67%**	**79%**
		F	**62%**	**67%**	**75%**	**34%**	**51%**	**66%**
		Accuracy	**69%**			**53%**		
		Kappa	**51%**			**27%**		
ADC	ADC-mode_mean_SD_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_slope_zscore_var	sensitivity	100%	100%	100%	33%	47%	51%
Backwards-elimination		specificity	100%	100%	100%	73%	67%	79%
		F	100%	100%	100%	29%	47%	56%
		Accuracy	100%			46%		
		Kappa	100%			18%		
	ADC-mode_mean_SD_median_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_slope_zscore_var	sensitivity	100%	100%	100%	43%	50%	62%
		specificity	100%	100%	100%	80%	70%	79%
		F	100%	100%	100%	40%	50%	64%
		Accuracy	100%			53%		
		Kappa	100%			28%		
	ADC-mode_mean_SD_median_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore_var	sensitivity	100%	100%	100%	43%	50%	64%
		specificity	100%	100%	100%	80%	72%	79%
		F	100%	100%	100%	40%	51%	66%
		Accuracy	100%			54%		
		Kappa	100%			30%		
	ADC-mean_SD_median_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore_var	sensitivity	100%	100%	100%	29%	50%	41%
		specificity	100%	100%	100%	73%	72%	67%
		F	100%	100%	100%	26%	51%	43%
		Accuracy	100%			42%		
		Kappa	100%			11%		

	ADC-mean_SD_median_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore_var	sensitivity	100%	92%	92%	24%	53%	38%
		specificity	100%	95%	95%	75%	68%	67%
		F	100%	92%	92%	22%	51%	41%
		Accuracy	94%			41%		
		Kappa	90%			9%		
	ADC-mean_SD_median_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore	sensitivity	100%	92%	100%	38%	56%	44%
		specificity	100%	100%	95%	80%	72%	67%
		F	100%	96%	96%	36%	55%	45%
		Accuracy	97%			47%		
		Kappa	95			18%		
	rADC-mode_mean_SD_median_10tile_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_slope_zscore_var	sensitivity	100%	100%	100%	43%	64%	56%
		specificity	100%	100%	100%	88%	80%	63%
		F	100%	100%	100%	46%	65%	54%
		Accuracy	100%			56%		
		Kappa	100%			32%		
	rADC-mean_SD_median_10tile_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_slope_zscore_var	sensitivity	100%	92%	85%	48%	67%	56%
		specificity	96%	95%	95%	87%	83%	65%
		F	93%	92%	88%	49%	69%	54%
		Accuracy	91%			58%		
		Kappa	86%			35%		
	rADC-mean_SD_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_slope_zscore_var	sensitivity	100%	92%	92%	43%	67%	54%
		specificity	100%	95%	95%	88%	78%	65%
		F	100%	92%	92%	46%	66%	53%
		Accuracy	94%			56%		
		Kappa	90%			32%		
	rADC-mean_SD_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore_var	sensitivity	100%	92%	84%	33%	78%	56%
		specificity	96%	95%	95%	88%	82%	67%
		F	93%	92%	88%	38%	75%	55%
		Accuracy	91%			59%		
		Kappa	86%			36%		
	rADC-mean_SD_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore	sensitivity	100%	92%	85%	33%	78%	56%
		specificity	96%	95%	95%	88%	82%	67%
		F	93%	92%	88%	38%	75%	55%
		Accuracy	91%			59%		
		Kappa	86%			36%		
		sensitivity	100%	92%	85%	33%	78%	54%

	rADC-mean_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore	specificity	96%	95%	95%	87%	80%	68%
		F	93%	92%	88%	37%	74%	54%
		Accuracy	91%			58%		
		Kappa	86%			35%		
	rADC-mean_median_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore	**sensitivity**	**86%**	**92%**	**85%**	**38%**	**78%**	**54%**
		specificity	**96%**	**95%**	**89%**	**87%**	**80%**	**70%**
		F	**86%**	**92%**	**85%**	**41%**	**74%**	**55%**
		Accuracy	**88%**			**59%**		
		Kappa	**81%**			**37%**		
	rADC-mean_median_30tile_40tile_60tile_70tile_80tile_90tile_skew_zscore	sensitivity	86%	92%	85%	38%	78%	54%
		specificity	96%	95%	89%	87%	80%	70%
		F	86%	92%	85%	41%	74%	55%
		Accuracy	88%			59%		
		Kappa	81%			37%		
Pairwise-correlation	**rADC_median_min_max_var_zscore**	**sensitivity**	**57%**	**92%**	**77%**	**24%**	**69%**	**64%**
		specificity	**96%**	**90%**	**79%**	**88%**	**83%**	**61%**
		F	**67%**	**88%**	**74%**	**29%**	**70%**	**58%**
		Accuracy	**78%**			**57%**		
		Kappa	**66%**			**32%**		

Table 11-B. Diagnostic performance of the combined features from ASL, DSC, DCE-mTK, DCE-L&L, and ADC for glioma sub-grouping

MRI-method	Features combinations	Prediction performance	Sub-grouping					Internal validation LOOCV				
			GBM-wt	Astro-mut-LG	Astro-mut-HG	Oligo-mut-LG	Oligo-mut-HG	GBM-wt	Astro-mut-LG	Astro-mut-HG	Oligo-mut-LG	Oligo-mut-HG
ASL	**rTBF-skew_zscore_mean_95tile_median_mode_min**	**sensitivity**	**71%**	**90%**	**50%**	**56%**	**50%**	**38%**	**83%**	**0%**	**22%**	**0%**
Backwards-elimination		**specificity**	**92%**	**82%**	**100%**	**87%**	**96%**	**73%**	**80%**	**96%**	**80%**	**93%**
		F	**71%**	**78%**	**67%**	**59%**	**57%**	**33%**	**74%**	**0%**	**26%**	**0%**
		Accuracy	**69%**					**41%**				
		Kappa	**58%**					**20%**				
Pairwise-correlation	**rTBF_SD_min_zscore**	**sensitivity**	**43%**	**90%**	**0%**	**67%**	**50%**	**29%**	**90%**	**0%**	**48%**	**0%**
		specificity	**92%**	**81%**	**100%**	**83%**	**93%**	**85%**	**74%**	**98%**	**75%**	**96%**
		F	**50%**	**78%**	**0%**	**63%**	**50%**	**32%**	**73%**	**0%**	**46%**	**0%**
		Accuracy	**63%**					**48%**				
		Kappa	**49%**					**28%**				
DSC	rCBF-mean_SD_95tile_kur_skew_median_iqr_mode_max_zscore_var_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_min_max_var	sensitivity	100%	100%	100%	100%	100%	76%	67%	0%	44%	33%
Backwards-elimination		specificity	100%	100%	100%	100%	100%	95%	76%	98%	77%	93%
		F	100%	100%	100%	100%	100%	78%	61%	0%	44%	36%
		Accuracy	100%					54%				
		Kappa	100%					38%				
	rCBF-mean_SD_95tile_skew_median_iqr_mode_max_zscore_var_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_min_max_var	sensitivity	100%	100%	100%	100%	100%	71%	60%	0%	48%	17%
		specificity	100%	100%	100%	100%	100%	89%	82%	97%	81%	86%
		F	100%	100%	100%	100%	100%	68%	60%	0%	49%	15%
		Accuracy	100%					50%				

		Kappa	100%					34%				
	rCBF-mean_SD_95tile_skew_median_iqr_mode_max_zscore_var_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_max_var	sensitivity	100%	100%	100%	100%	100%	71%	60%	0%	48%	17%
		specificity	100%	100%	100%	100%	100%	89%	82%	97%	81%	86%
		F	100%	100%	100%	100%	100%	68%	60%	0%	49%	15%
		Accuracy	100%					50%				
		Kappa	100%					34%				
	rCBF-mean_SD_95tile_skew_median_iqr_max_zscore_var_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_max_var	sensitivity	100%	100%	100%	100%	100%	76%	50%	0%	62%	25%
		specificity	100%	100%	100%	100%	100%	92%	83%	98%	75%	89%
		F	100%	100%	100%	100%	100%	74%	54%	0%	56%	25%
		Accuracy	100%					53%				
		Kappa	100%					37%				
	rCBF-mean_SD_95tile_median_iqr_max_zscore_var_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_max_var	sensitivity	100%	100%	100%	100%	100%	67%	57%	0%	56%	25%
		specificity	100%	100%	100%	100%	100%	91%	85%	96%	74%	89%
		F	100%	100%	100%	100%	100%	67%	60%	0%	50%	25%
		Accuracy	100%					51%				
		Kappa	100%					35%				
	rCBF-mean_95tile_median_iqr_max_zscore_var_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_max_var	sensitivity	100%	100%	100%	100%	100%	67%	57%	0%	56%	25%
		specificity	100%	100%	100%	100%	100%	91%	85%	97%	73%	90%
		F	100%	100%	100%	100%	100%	67%	60%	0%	49%	26%
		Accuracy	100%					51%				
		Kappa	100%					35%				
	rCBF-mean_95tile_median_iqr_max_zscore_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_max_var	sensitivity	100%	100%	100%	100%	100%	67%	60%	0%	56%	25%
		specificity	100%	100%	100%	100%	100%	91%	85%	97%	74%	90%

		F	100%	100%	100%	100%	100%	67%	62%	0%	50%	26%
		Accuracy	100%					52%				
		Kappa	100%					36%				
	rCBF-mean_95tile_median_iqr_max_rCBV-mean_SD_95tile_median_iqr_mode_slope_entropy_max_var	sensitivity	100%	100%	100%	100%	100%	71%	53%	0%	56%	25%
		specificity	100%	100%	100%	100%	100%	92%	82%	93%	71%	96%
		F	100%	100%	100%	100%	100%	71%	55%	0%	48%%	33%
		Accuracy	100%					51%				
		Kappa	100%					34%				
	rCBF-mean_95tile_median_iqr_max_rCBV-mean_95tile_median_iqr_mode_slope_entropy_max_var	sensitivity	100%	100%	100%	100%	100%	71%	53%	0%	56%	25%
		specificity	100%	100%	100%	100%	100%	91%	82%	93%	72%	96%
		F	100%	100%	100%	100%	100%	70%	55%	0%	49%	33%
		Accuracy	100%					51%				
		Kappa	100%					34%				
	rCBF-mean_95tile_median_iqr_max_rCBV-mean_95tile_median_iqr_mode_slope_entropy_max	sensitivity	100%	100%	100%	100%	100%	71%	53%	0%	52%	25%
		specificity	100%	100%	100%	100%	100%	91%	82%	93%	72%	95%
		F	100%	100%	100%	100%	100%	70%	55%	0%	47%	32%
		Accuracy	100%					50%				
		Kappa	100%					33%				
	rCBF-mean_95tile_median_iqr_max_rCBV-mean_95tile_median_iqr_mode_entropy_max	sensitivity	100%	100%	100%	100%	100%	79%	57%	0%	59%	25%
		specificity	100%	100%	100%	100%	100%	92%	89%	92%	74%	93%
		F	100%	100%	100%	100%	100%	74%	63%	0%	52%	29%
		Accuracy	100%					54%				
		Kappa	100%					39%				

	rCBF-mean_95tile_median_iqr_max_rCBV-mean_95tile_median_mode_entropy_max	sensitivity	100%	100%	100%	100%	100%	81%	57%	0%	63%	25%
		specificity	100%	100%	100%	100%	100%	93%	88%	94%	75%	92%
		F	100%	100%	100%	100%	100%	79%	62%	0%	56%	27%
		Accuracy	100%					56%				
		Kappa	100%					42%				
	rCBF-mean_95tile_median_iqr_rCBV-mean_95tile_median_mode_entropy_max	sensitivity	100%	100%	100%	100%	100%	76%	57%	0%	67%	25%
		specificity	100%	100%	100%	100%	100%	92%	89%	93%	75%	93%
		F	100%	100%	100%	100%	100%	74%	63%	0%	58%	29%
		Accuracy	100%					56%				
		Kappa	100%					42%				
	rCBF-mean_95tile_iqr_rCBV-mean_95tile_median_mode_entropy_max	**sensitivity**	**100%**	**100%**	**100%**	**100%**	**100%**	**71%**	**63%**	**0%**	**63%**	**25%**
		specificity	**100%**	**100%**	**100%**	**100%**	**100%**	**92%**	**89%**	**93%**	**73%**	**95%**
		F	**100%**	**100%**	**100%**	**100%**	**100%**	**71%**	**68%**	**0%**	**54%**	**32%**
		Accuracy	**100%**					**56%**				
		Kappa	**100%**					**42%**				
	rCBF-mean_95tile_iqr_rCBV-mean_95tile_mode_entropy_max	sensitivity	100%	100%	100%	100%	100%	71%	60%	0%	63%	8%
		specificity	100%	100%	100%	100%	100%	92%	88%	94%	78%	87%
		F	100%	100%	100%	100%	100%	71%	64%	0%	58%	8%
		Accuracy	100%					53%				
		Kappa	100%					38%				
	rCBF-95tile_iqr_rCBV-mean_95tile_mode_entropy_max	sensitivity	100%	90%	100%	100%	75%	67%	63%	0%	56%	17%
		specificity	100%	100%	100%	95%	96%	91%	88%	94%	75%	89%
		F	100%	95%	100%	95%	75%	67%	67%	0%	51%	17%

		Accuracy	94%					52%				
		Kappa	92%					36%				
Pairwise-correlation	**rCBF_kur_mode_rCBV_min_max**	**sensitivity**	**100%**	**80%**	**50%**	**89%**	**75%**	**62%**	**73%**	**0%**	**56%**	**17%**
		specificity	**100%**	**95%**	**97%**	**87%**	**100%**	**89%**	**88%**	**56%**	**78%**	**89%**
		F	**100%**	**84%**	**50%**	**80%**	**86%**	**62%**	**73%**	**0%**	**53%**	**17%**
		Accuracy	**84%**					**54%**				
		Kappa	**79%**					**39%**				
	rCBF_kur_mode_slope_rCBV_mean_min	sensitivity	100%	90%	100%	78%	75%	33%	70%	0%	56%	8%
		specificity	96%	100%	100%	91%	96%	92%	76%	97%	74%	89%
		F	93%	95%	100%	78%	75%	41%	63%	0%	50%	9%
		Accuracy	88%					46%				
		Kappa	84%					27%				
DCE_mTK	**Ve-mode_rVe-mode_rVp-mean_median_min**	**sensitivity**	**57%**	**80%**	**0%**	**89%**	**75%**	**10%**	**77%**	**0%**	**19%**	**42%**
Backwards-elimination		**specificity**	**88%**	**82%**	**100%**	**91%**	**100%**	**84%**	**70%**	**100%**	**71%**	**89%**
		F	**57%**	**73%**	**0%**	**84%**	**86%**	**11%**	**63%**	**0%**	**19%**	**38%**
		Accuracy	**72%**					**36%**				
		Kappa	**62%**					**14%**				
Pairwise-correlation	**Ve_mode_rVp_mean**	**sensitivity**	**29%**	**90%**	**0%**	**68%**	**75%**	**10%**	**83%**	**0%**	**48%**	**17%**
		specificity	**92%**	**77%**	**100%**	**78%**	**100%**	**88%**	**73%**	**100%**	**65%**	**96%**
		F	**36%**	**75%**	**0%**	**60%**	**86%**	**13%**	**68%**	**0%**	**41%**	**24%**
		Accuracy	**63%**					**44%**				
		Kappa	**48%**					**22%**				
DCE_L&L		sensitivity	86%	100%	100%	78%	100%	29%	43%	33%	37%	50%

Backwards-elimination	rK-iqr_rFb-mean_SD_95tile_median_iqr_slope_max_var_rPS-mean_median_rTC-mean	specificity	96%	95%	100%	96%	100%	79%	77%	99%	67%	95%
		F	86%	95%	100%	82%	100%	28%	45%	44%	33%	55%
		Accuracy	90%					39%				
		Kappa	88%					18%				
	rFb-mean_SD_95tile_median_iqr_slope_max_var_rPS-mean_median_rTC-mean	sensitivity	86%	100%	100%	89%	100%	19%	37%	33%	41%	85%
		specificity	100%	90%	100%	100%	100%	77%	76%	98%	67%	96%
		F	92%	91%	100%	94%	100%	19%	39%	40%	36%	64%
		Accuracy	94%					36%				
		Kappa	92%					15%				
	rFb-mean_SD_95tile_median_iqr_slope_max_var_rPS-mean_median	**sensitivity**	**86%**	**100%**	**100%**	**89%**	**100%**	**19%**	**53%**	**0%**	**41%**	**58%**
		specificity	**100%**	**91%**	**100%**	**100%**	**100%**	**75%**	**74%**	**100%**	**72%**	**96%**
		F	**92%**	**91%**	**100%**	**94%**	**100%**	**18%**	**51%**	**0%**	**39%**	**64%**
		Accuracy	**94%**					**40%**				
		Kappa	**92%**					**18%**				
	rFb-mean_SD_95tile_median_iqr_slope_max_var_rPS-median	sensitivity	86%	90%	100%	89%	100%	10%	50%	17%	26%	50%
		specificity	100%	91%	100%	96%	100%	79%	74%	98%	68%	90%
		F	92%	86%	100%	89%	100%	10%	48%	22%	25%	46%
		Accuracy	91%					32%				
		Kappa	88%					10%				
Pairwise-correlation	TC_95tile_rK_median_rKep_max_rFb_slope_rPS_mean	sensitivity	86%	70%	50%	67%	100%	33%	43%	0%	52%	50%
		specificity	96%	826%	100%	87%	100%	91%	64%	93%	78%	95%
		F	86%	67%	67%	67%	100%	40%	39%	0%	50%	55%
		Accuracy	75%					42%				

		Kappa	67%					22%				
	TC_slope_rK_slope_rKep_max_rVp_mean_rPS_median	sensitivity	86%	60%	100%	78%	75%	52%	30%	0%	52%	58%
		specificity	92%	82%	100%	91%	100%	84%	70%	93%	83%	94%
		F	80%	63%	100%	84%	86%	50%	31%	0%	53%	58%
		Accuracy	75%					43%				
		Kappa	67%					24%				
	rK_median_rKep_max_rFb_median_rPS_mean_rTC_mode	**sensitivity**	**86%**	**70%**	**100%**	**89%**	**100%**	**48%**	**43%**	**0%**	**52%**	**50%**
		specificity	**96%**	**91%**	**97%**	**96%**	**100%**	**93%**	**65%**	**91%**	**75%**	**100%**
		F	**86%**	**74%**	**80%**	**89%**	**100%**	**56%**	**39%**	**0%**	**48%**	**67%**
		Accuracy	**84%**					**45%**				
		Kappa	**80%**					**26%**				
	rK_slope_rFb_SD_rPS_median_rTC_kur_rTC_mode_rTC_entropy	sensitivity	71%	90%	100%	78%	75%	33%	33%	0%	30%	42%
		specificity	96%	82%	100%	96%	100%	77%	74%	93%	75%	89%
		F	77%	78%	100%	82%	86%	31%	35%	0%	31%	38%
		Accuracy	81%					31%				
		Kappa	75%					10%				
ADC	ADC-mean_SD_median_30tile_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore_var	sensitivity	100%	100%	100%	100%	100%	29%	47%	17%	33%	0%
Backwards-elimination		specificity	100%	100%	100%	100%	100%	75%	85%	92%	77%	83%
		F	100%	100%	100%	100%	100%	26%	52%	14%	35%	0%
		Accuracy	100%					31%				
		Kappa	100%					11%				
	ADC-mean_SD_median_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore_var	sensitivity	100%	100%	100%	100%	100%	14%	50%	50%	37%	0%
		specificity	100%	100%	100%	100%	100%	76%	79%	97%	80%	81%

		F	100%	100%	100%	100%	100%	14%	51%	50%	39%	0%
		Accuracy	100%					32%				
		Kappa	100%					11%				
	ADC-mean_median_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore_var	sensitivity	100%	100%	100%	100%	100%	10%	50%	33%	22%	0%
		specificity	100%	100%	100%	100%	100%	83%	79%	93%	73%	77%
		F	100%	100%	100%	100%	100%	11%	51%	29%	23%	0%
		Accuracy	100%					26%				
		Kappa	100%					4%				
	ADC-mean_median_40tile_60tile_70tile_80tile_90tile_95tile_skew_zscore	sensitivity	100%	100%	100%	100%	100%	10%	57%	0%	22%	0%
		specificity	100%	100%	100%	100%	100%	88%	76%	91%	77%	74%
		F	100%	100%	100%	100%	100%	13%	54%	0%	24%	0%
		Accuracy	100%					26%				
		Kappa	100%					4%				
	ADC-mean_median_40tile_60tile_70tile_80tile_90tile_95tile_skew	sensitivity	100%	100%	100%	100%	100%	10%	53%	0%	26%	0%
		specificity	100%	100%	100%	100%	100%	88%	77%	87%	72%	81%
		F	100%	100%	100%	100%	100%	13%	52%	0%	26%	0%
		Accuracy	100%					26%				
		Kappa	100%					4%				
	rADC-mean_SD_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_var_skew_zscore	**sensitivity**	**86%**	**100%**	**50%**	**89%**	**75%**	**29%**	**73%**	**0%**	**44%**	**0%**
		specificity	**96%**	**95%**	**100%**	**96%**	**96%**	**84%**	**77%**	**96%**	**80%**	**87%**
		F	**86%**	**95%**	**67%**	**89%**	**75%**	**31%**	**66%**	**0%**	**45%**	**0%**
		Accuracy	**88%**					**42%**				
		Kappa	**83%**					**23%**				

	rADC-mean_SD_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_var_skew	sensitivity	86%	90%	50%	89%	75%	29%	73%	0%	44%	0%
		specificity	96%	95%	97%	95%	97%	85%	77%	94%	74%	92%
		F	86%	90%	50%	89%	75%	32%	66%	0%	42%	0%
		Accuracy	84%					42%				
		Kappa	79%					21%				
	rADC-mean_SD_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile_var	sensitivity	86%	90%	50%	89%	50%	43%	60%	0%	37%	0%
		specificity	96%	95%	100%	87%	96%	89%	79%	93%	65%	92%
		F	86%	90%	67%	80%	57%	47%	58%	0%	33%	0%
		Accuracy	81%					39%				
		Kappa	75%					17%				
	rADC-mean_SD_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile	sensitivity	86%	90%	50%	89%	50%	38%	60%	0%	37%	0%
		specificity	96%	95%	100%	87%	96%	89%	77%	94%	65%	90%
		F	86%	90%	67%	80%	57%	43%	57%	0%	33%	0%
		Accuracy	81%					38%				
		Kappa	75%					16%				
	rADC-mean_median_20tile_30tile_40tile_60tile_70tile_80tile_90tile_95tile	sensitivity	86%	90%	50%	67%	50%	24%	60%	0%	30%	0%
		specificity	92%	91%	100%	87%	96%	88%	76%	92%	67%	88%
		F	80%	86%	67%	67%	57%	29%	56%	0%	28%	0%
		Accuracy	75%					32%				
		Kappa	66%					10%				
	rADC-mean_median_30tile_40tile_60tile_70tile_80tile_90tile_95tile	sensitivity	71%	90%	50%	78%	50%	19%	67%	0%	30%	0%
		specificity	92%	95%	97%	87%	96%	88%	73%	94%	67%	89%
		F	71%	90%	50%	74%	57%	24%	59%	0%	28%	0%

		Accuracy	75%					33%				
		Kappa	67%					10%				
	rADC-mean_median_30tile_40tile_60tile_70tile_80tile_90tile	sensitivity	71%	90%	0%	78%	25%	24%	73%	0%	33%	0%
		specificity	88%	91%	100%	83%	96%	88%	73%	97%	68%	90%
		F	67%	86%	0%	70%	33%	29%	63%	0%	31%	0%
		Accuracy	69%					38%				
		Kappa	57%					15%				
	rADC-mean_median_40tile_60tile_70tile_80tile_90tile	sensitivity	86%	90%	0%	78%	25%	26%	73%	0%	37%	0%
		specificity	88%	91%	100%	83%	100%	89%	74%	97%	70%	89%
		F	75%	86%	0%	70%	40%	34%	64%	0%	34%	0%
		Accuracy	72%					40%				
		Kappa	62%					18%				
Pairwise-correlation	**rADC_median_min_max_var_zscore**	**sensitivity**	**71%**	**70%**	**100%**	**56%**	**75%**	**26%**	**57%**	**0%**	**33%**	**25**
		specificity	**92%**	**91%**	**100%**	**78%**	**96%**	**80%**	**76%**	**96%**	**77%**	**88%**
		F	**71%**	**74%**	**100%**	**53%**	**75%**	**29%**	**54%**	**0%**	**35%**	**24%**
		Accuracy	**69%**					**36%**				
		Kappa	**59%**					**15%**				