Imperial College London

Department of Medicine

Magnetic Resonance Spectroscopy in Hypoxic Ischaemic Encephalopathy

Peter James Lally

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy
at Imperial College London

DECLARATION OF ORIGINALITY

The material presented in this thesis is my own work, unless otherwise acknowledged and appropriately referenced.

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

ABSTRACT

Hypoxic-ischaemic encephalopathy results from an acute period of reduced oxygen (hypoxia) and blood flow (ischaemia) at the time of birth, and typically comes as the devastating end to an otherwise normal pregnancy. The most severely affected infants will die within days, whereas the majority who survive have a high chance of sustaining significant lifelong disabilities including cerebral palsy. This therefore has profound implications for the affected infant, their family, and wider society.

Magnetic resonance spectroscopy allows the non-invasive investigation of cerebral metabolism, and provides quantitative measures of brain injury. Since hypoxic-ischaemic encephalopathy is associated with acute metabolic changes in the perinatal period, magnetic resonance spectroscopy has the potential to give insight into the ongoing injury, and potentially predict the neurodevelopmental prognosis of affected infants at an early stage.

This thesis tests the hypothesis that magnetic resonance spectroscopy, acquired within two weeks of birth, is effective at predicting neurodevelopmental outcomes years after the initial hypoxic-ischaemic injury. This is achieved via a prospective multi-centre observational study across several institutions across the UK and USA, the results of which indicate exceptional prognostic value.

The following work tests the hypothesis that accurate magnetic resonance spectroscopy measurements can be performed in the context of routine clinical settings to assist prognostication. A novel acquisition and analysis strategy is developed and validated, demonstrating precise and accurate results in a clinical setting within a modest acquisition time.

ACKNOWLEDGEMENTS

I would like to thank Sudhin Thayyil (supervisor), who always provided me with invaluable insight and feedback, and who gave me many opportunities for me to explore new ideas. In particular, Sudhin took me on at the start of my research career, and took every opportunity to support me, encourage my development and promote my independence. His positivity and drive were a constant source of inspiration, along with that of Paolo Montaldo, Vânia Oliveira, and Josephine Mendoza, and everyone else who I have worked with at the CPN, making it a fantastic team to be a part of.

I would also like to particularly thank Alan Bainbridge, who has always been a great mentor and source of enthusiasm, and who gave his time to provide many opportunities for invaluable feedback, advice and stimulating discussions. I would also like to thank David Sharp, who provided me with additional valuable mentorship and advice.

In addition, I would like to thank all of the following people, who have all given me invaluable advice and support over the past few years: Ern Cady, David Price, Mani Chandrasekharan, Ravi Swamy, Paul Matthews, Gary Zhang, and Enrico de Vita.

Most of all I wish to thank my wonderful family, and I dedicate this to Rossella, who I was so fortunate to meet at the start of this journey, and who will always be the best result of Imperial's PhD programme.

CONTENTS

1 INTRODUCTION 19
1.1 Hypoxic ischaemic encephalopathy 19
1.2 Magnetic resonance spectroscopy in HIE 19
1.3 MARBLE 20
1.4 Thesis contributions 21
2 NUCLEAR MAGNETIC RESONANCE AND *IN-VIVO* SPECTROSCOPY 23
2.1 Nuclear magnetic resonance 23
2.1.1 Magnetic moment and the Larmor frequency 23
2.1.2 Bulk magnetisation 23
2.1.3 Excitation 24
2.1.4 Localisation 24
2.1.5 T_1 relaxation 25
2.1.6 T_2 and T_2^* relaxation 25
2.1.7 The Free Induction Decay 25
2.1.8 Spin echoes 26
2.1.9 The Discrete Fourier Transform 26
2.2 *In-vivo* magnetic resonance spectroscopy 27
2.2.1 Chemical shift 27
2.2.2 Scalar coupling 27
2.2.3 Point Resolved Spectroscopy (PRESS) 28
2.2.4 Stimulated Echo Acquisition Mode (STEAM) 29
2.3 Metabolites of interest 30
2.3.1 N-acetylaspartate (NAA) 30
2.3.2 N-acetylaspartylglutamate (NAAG) 32
2.3.3 Choline (Cho), Glycerophosphocholine (GPC) and Phosphocholine (PCh) 32
2.3.4 Creatine (Cr) and Phosphocreatine (PCr) 32
2.3.5 Lactate (Lac) 33
2.3.6 Threonine (Thr) 33
2.3.7 Macromolecules (MM) and Lipids (Lip) 33
2.4 Acquisition considerations 34
2.4.1 Linewidth and shimming 34
2.4.2 Eddy currents 35
2.4.3 Chemical shift displacement 36
2.4.4 Anomalous $\boldsymbol{J}$ modulation 37
2.4.5 Unwanted coherences 37
2.4.6 Motion 38

2.5 Analysis considerations 38
2.5.1 Pre-processing 38
2.5.2 Analysis 38
2.5.3 Baseline 39
2.5.4 Uncertainty 40
2.5.5 Concentration determination 41
3 APPLICATIONS OF MR SPECTROSCOPY TO HYPOXIC ISCHAEMIC ENCEPHALOPATHY 42
3.1 The healthy developing brain 42
3.2 Hypoxic ischaemic encephalopathy 44
3.2.1 Primary energy failure 44
3.2.2 Secondary energy failure 44
3.2.3 Therapeutic hypothermia 45
3.3 Prognostication 46
3.3.1 ^{31}P and ^{1}H MRS before the introduction of hypothermia 46
3.3.2 ^{1}H MRS after hypothermic neuroprotection 47
3.3.3 Other MR biomarkers in HIE 47
3.3.4 Amplitude-integrated electroencephalography (aEEG) 48
3.3.5 ^{1}H MRS in multicentre trials of HIE 49
4 THE PROGNOSTIC ACCURACY OF MRS IN HYPOXIC ISCHAEMIC ENCEPHALOPATHY 50
4.1 Introduction 50
4.2 MRS protocol 50
4.2.1 Step 1: Metabolite peak area ratios (long TE) 50
4.2.2 Steps 1-3: Metabolite T_1 and T_2 Relaxometry 51
4.2.3 Steps 4-9: Absolute quantification 51
4.3 Phantom validation 54
4.3.1 Methods 54
4.3.2 Results 54
4.3.3 Discussion 55
4.4 Adult volunteers 58
4.4.1 Outline 58
4.4.2 Methods 58
4.4.3 Results 58
4.4.4 Discussion 59
4.5 Clinical study 60
4.5.1 Study design 60
4.5.2 MR protocol 61
4.5.3 Statistical analysis 62
4.6 Results 63

4.6.1 Clinical characteristics 63
4.6.2 MR data quality 63
4.6.3 Prognostic accuracy 65
4.6.4 Linear regression model 71
4.6.5 Sensitivity analysis 71
4.7 Discussion 71
4.7.1 Clinical characteristics 72
4.7.2 Interpretation of MRS changes 72
4.7.3 Limitations 73
4.7.4 The potential of [tNAA] 74
4.8 Conclusions 75
5 DESIGNING A RAPID ACQUISITION PROTOCOL 76
5.1 Outline 76
5.2 Variance due to metabolite relaxivity 76
5.2.1 Outline 76
5.2.2 Study design 76
5.2.3 Methods 76
5.2.4 Results 77
5.2.5 Discussion 78
5.2.6 Conclusions 80
5.3 Temporal changes in metabolite concentrations 80
5.3.1 Outline 80
5.3.2 Study design 80
5.3.3 Methods 80
5.3.4 Results 81
5.3.5 Discussion 81
5.3.6 Conclusions 82
5.4 Variance due to parenchymal water concentration 82
5.5 Proposed solution: two-point saturation recovery STEAM 82
5.5.1 Acquisition details 82
5.5.2 Choice of acquisition scheme 83
5.5.3 Steps 1-4: two-point saturation recovery STEAM 85
5.5.4 Steps 5-12: brain water T_2 relaxometry 85
5.5.5 Lipids and macromolecules 85
5.6 Phantom Validation 85
5.6.1 Outline 85
5.6.2 Methods 85
5.6.3 Results 86

5.6.4 Discussion 87
5.7 Optimising two-point saturation recovery 88
5.7.1 Outline 88
5.7.2 Methods 88
5.7.3 Results 89
5.7.4 Discussion 92
5.7.5 Conclusion 92
6 VALIDATING THE CLINICAL PROTOCOL *IN-VIVO* 93
6.1.1 Outline 93
6.1.2 Study design 93
6.1.3 Methods 93
6.1.4 Results 94
6.1.5 Discussion 96
6.1.6 Conclusion 97
7 DEVELOPING AN OPTIMAL ANALYSIS STRATEGY 98
7.1 Demonstrating proof of principle 98
7.1.1 Outline 98
7.1.2 The TARQUIN signal model for a single spectrum 101
7.1.3 Hard and soft constraints in TARQUIN 102
7.1.4 An adapted TARQUIN model 102
7.1.5 Validation study 103
7.1.6 Methods: Simulated acquisition 103
7.1.7 Methods: Analysis 104
7.1.8 Results 105
7.1.9 Discussion 107
7.1.10 Conclusion 108
7.2 Developing an optimal analysis strategy 108
7.2.1 TARQUIN macromolecular baseline 108
7.2.2 Alternative: AQSES macromolecular baseline 109
7.2.3 Time domain vs frequency domain fitting 109
7.2.4 Efficiency 110
7.3 Proposed fitting strategy 110
7.3.1 Signal model 110
7.3.2 Spline baseline 111
7.3.3 Formulation of the fitting problem 112
7.3.4 Soft constraints 113
7.3.5 Hard constraints 113
7.3.6 Variance estimation 113

7.3.7 Implementation details 114
7.3.8 Hyperparameters 114
7.4 *In-vivo* validation 115
7.4.1 Outline 115
7.4.2 Methods 115
7.4.3 Results 117
7.4.4 Discussion 122
7.4.5 Limitations 122
7.4.6 Conclusion 123
8 DISCUSSION AND FUTURE WORK 124
8.1 The promise of MRS in HIE 124
8.2 Insight into neuronal injury 125
8.3 Two-point saturation recovery STEAM 125
8.4 Future work 125
8.4.1 Incorporation into semi-LASER acquisition schemes 126
8.4.2 Implementation of MRS in other contexts 126
8.4.3 Acquisition of longitudinal data and further follow-up 126
8.4.4 Flexible lineshapes and imperfect RF excitation 126
8.4.5 Linking the behaviour of metabolite resonances 126
8.5 Summary 126

TABLES

Table 4.1: Key acquisition parameters for the full MRS protocol for the MARBLE study. 51

Table 4.2: Differences in cerebral MRS measures between infants with normal and abnormal neurodevelopment at two years of age. P-values are reported for Student t-test. AUC =Area under the receiver operating characteristic curve. *Ratio of metabolite peak areas at TR/TE=2288/288ms only, uncorrected for relaxation. . 67

Table 4.3: The relationship between each of the MRS measures and two-year motor composite score on Bayley-III. This is controlled for gestational age at birth and postnatal age at the time of scan. The coefficient represents the change in motor score for a unit change in the MRS measure. Concentrations are in terms of mmol/kg wet weight, while relaxation rates are in milliseconds. *Ratio of metabolite peak areas at TR/TE=2288/288ms only, uncorrected for relaxation. 67

Table 4.4: Sensitivity and specificity of each MR measure at the optimal cut-off. TP = true positive; FP = false positive; TN = true negative; FN = false negative; PLIC = posterior limbs of the internal capsules; BG = basal ganglia; BGT = basal ganglia + thalamus composite; PPV = positive predictive value; NPV = negative predictive value. *tLac/tNAA is the ratio of metabolite peak areas recorded in the TR/TE=2288/288ms acquisition. 70

Table 5.1: Mean (SD) T_1 and T_2 values for the methyl singlet peaks and parenchyma water across the MARBLE cohort. 77

Table 5.2: Relationship between relaxation rates and birth gestational age, controlling for motor outcome score at two years and postnatal age. 77

Table 5.3: Relationship between relaxation rates and postnatal age, controlling for motor outcome score at two years and gestational age at birth. 78

Table 5.4: Relationship between metabolite concentration and birth gestational age, controlling for postnatal age at scan and two-year motor score on Bayley-III. 81

Table 5.5: Relationship between metabolite concentration and postnatal age at scan, controlling for birth gestational age and two-year motor score on Bayley-III. 81

Table 5.6: Key acquisition parameters for the full STEAM MRS protocol. Each dummy scan represents a single acquisition of all phase cycling steps. 83

Table 5.7: Estimates of relaxivities and concentrations (with standard errors in parentheses) derived from the phantom experiment. 87

Table 7.1: Amplitude ratio soft constraints implemented in TARQUIN. 115

Table 7.2: Parameterisation of the macromolecule peaks with nine Gaussian components. 116

Table 7.3: Lipid parameterisation employed in LCModel. 116

FIGURES

Figure 2.1: Timing diagram for 3D single voxel localisation with the PRESS sequence. The top line depicts the transmit/receive (T/R) events, and the gradient waveforms in three orthogonal directions (G_x, G_y, G_z) are shown on the lower three lines. RF pulses are shown in black, while the detected echo is shown in white. Selection and refocusing gradients (light grey) ensure detection of coherent signals from the volume of interest. Crusher gradients (dark grey) prevent the coherence of unwanted signals. The echo is detected at time TE (=TE_1+TE_2). 29

Figure 2.2: Timing diagram for 3D single voxel localisation with the STEAM pulse sequence. The top line depicts the transmit/receive (T/R) events, and the gradient waveforms in three orthogonal directions (G_x, G_y, G_z) are shown on the lower three lines. RF pulses are shown in black, while the detected echo is shown in white. Selection and refocusing gradients (light grey) ensure detection of coherent signals from the volume of interest. Crusher gradients (dark grey) prevent the coherence of unwanted signals. The echo is detected at time TM+TE. 29

Figure 4.1: Schematic of the post-processing strategy used. I developed a GUI in Python to allow (i) raw data conversion from Philips (DICOM), Siemens (IMA) or GE (P-files) scanners, (ii) manual selection of non-corrupted sub-spectra, and (iii) residual water subtraction with HLSVD. The sub-spectra are then passed to LCModel (iv) for individual fitting (eddy current correction followed by zeroth order phasing and frequency alignment only), and the newly aligned sub-spectra are averaged and entered into a final LCModel fit (v). 52

Figure 4.2: Calculation steps to derive thalamic tNAA concentration, [tNAA], from the data collected in the various steps of the protocol. Black boxes show the input data, either obtained from the LCModel/HLSVD fits or reference values from previous work/literature. Red boxes and text denote calculated quantities from the tNAA resonance, and blue boxes and text denote calculated quantities from the water signal. Equations accompany the different calculation steps. $S_{NAA,step}$ denotes the tNAA signal intensity in the corresponding protocol step, with the relaxation corrected intensities for tNAA and parenchyma water given by S_{NAA} and S_p respectively. 53

Figure 4.3: Measured Lac/NAA and [NAA] values in the phantom at each site. The dashed line indicates ground truth [NAA]. UCL=University College London (UK), LPL=Alder Hey Children's Hospital (Liverpool, UK), ICL=Imperial College London (UK), BRI=Bristol University (UK), COV=University Hospital Coventry (UK), WSU=Wayne State University (Detroit, MI, USA), NNU= Norfolk and Norwich University Hospital (UK), MWY=Medway Maritime Hospital (Gillingham, UK). Sites are grouped by scanner vendor, from left to right: 3 x Philips (deep blue), 2 x Siemens (light blue), 4 x GE (grey). 55

Figure 4.4: Representative phantom spectra and LCModel fits on a GE (top row) and Philips (bottom row) scanner. Left: eight spectra overlaid from the TR/TE=5000/60ms acquisition. Centre: LCModel fit of the NAA signal at TR/TE=5000/60ms. Right: LCModel fit to the NAA and Lac signals at TR/TE=2288/288ms. 56

Figure 4.5: Representative phantom water signals (black) and HLSVD fits (red) on a GE (top row) and Philips (bottom row) scanner. Left: Time domain representation. Right: Frequency domain representation, with fitting residual (black). 57

Figure 4.6: Plots of 100 successive wild bootstrap iterations, performing an exponential fit (black lines) to the HLSVD fitted water signals at different TEs (red dots) on each of a GE (left) and Philips (right) scanner. 57

Figure 4.7: Top: thalamic [tNAA] measured in the five healthy adults on each scanner. Bottom: prescribed thalamic voxel positions in the five adults (red), overlaid on an axial T_2-weighted image (Philips). 60

Figure 4.8: Scatter plot of signal-to-noise ratios and linewidths obtained from all analysed spectra with TR/TE=2288/288ms (left) and TR/TE=5000/60ms (right). Blue points indicate datasets included in the final analysis, with red points excluded due to excessive linewidths outside of the normal range for each acquisition. 64

Figure 4.9: Representative PRESS spectra from a single subject (in black), with the LCModel fit overlaid (in red). These spectra have been pre-processed with manual rejection of motion corrupted sub-spectra, followed by automatic frequency and zeroth order phase alignment, prior to averaging. 64

Figure 4.10: Median spectra (TR=5060ms, TE=60ms) from infants with normal (blue) and abnormal (red) outcomes two years later, with shading denoting the interquartile range. Note the decrease in the region of NAA and increase in the region of Lac with abnormal outcomes. 65

Figure 4.11: Median spectra (TR=2288ms, TE=288ms) from infants with normal (blue) and abnormal (red) outcomes two years later, with shading denoting the interquartile range. Note the decrease in the region of NAA and increase in the region of Lac with abnormal outcomes. 66

Figure 4.12: Mean (top) and standard deviation (bottom) of the correlation matrices across all of the fitted spectra acquired with TR/TE=2288/288ms in the study. 'no CH_3' indicates multiplet resonances fitted independently from the metabolite's methyl singlet peak..... 68

Figure 4.13: Mean (top) and standard deviation (bottom) of the correlation matrices across all of the fitted spectra acquired with TR/TE=5000/60ms in the study. 'no CH_3' indicates multiplet resonances fitted independently from the metabolite's methyl singlet peak..... 69

Figure 4.14: Prognostic accuracies of various measures analysed in the MARBLE study. Parentheses and error bars indicate 95% confidence intervals, using the exact binomial method. TP = true positive; FP = false positive; TN = true negative; FN = false negative; LR+ = positive likelihood ratio; LR- = negative likelihood ratio; aEEG = amplitude integrated electroencephalogram. 70

Figure 4.15: Plot of mean-squared cross-validation error as the lasso regularisation term is increased, reducing the number of terms in the linear model. The leftmost vertical dashed line indicates the linear model with the lowest cross-validation error, including nine parameters. The rightmost vertical dashed line indicates the most parsimonious linear model within one standard error of the lowest cross-validation error, which included [tNAA] as the only variable. 71

Figure 5.1: Top left) Profile of T_2 relaxation corresponding to the cohort mean for tNAA (red) with 95% confidence intervals (grey). Top right) The same 95% confidence intervals in terms of the cohort mean profile, giving an insight into the uncertainty when applying a fixed T_2 correction. Bottom row) The equivalent plots for T_1 relaxation of tNAA..... 79

Figure 5.2: Calculation steps to derive thalamic tNAA concentration, [tNAA], from the data collected in the various steps of the protocol. Black boxes show the input data, either obtained from the LCModel/HLSVD fits or reference values from previous work/literature. Red boxes and text denote calculated quantities from the tNAA resonance, and blue boxes and text denote calculated quantities from the water signal. Equations accompany the different calculation steps. $S_{NAA,step}$ denotes the tNAA signal intensity in the corresponding protocol step, with the relaxation corrected intensities for tNAA and parenchyma water given by S_{NAA} and S_p respectively. 84

Figure 5.3: Spectra obtained using the STEAM (top, TR/TM/TE=3500/20/20ms) and PRESS (bottom, TR/TE=5000/60ms) protocols..... 86

Figure 5.4: Left) Fitted acetate signal from the STEAM protocol (top) and PRESS protocol (bottom) along with fitted T_1 recovery curve. Right) Fitted water signal from the STEAM protocol (top) and PRESS protocol (bottom) along with the fitted T_2 decay..... 87

Figure 5.5: Illustration of the Monte Carlo experiment to quantify the uncertainty in the relaxation corrected signal and T_1 (red) as experimental conditions are varied (blue) for two-point saturation recovery. 88

Figure 5.6: Theoretical signal-to-noise ratio per unit time plotted against repetition time. Plots are shown to represent the behaviour of short T_1 macromolecules (red) and longer T_1 metabolites (blue). 89

Figure 5.7: Left) Mean estimate of the fully relaxed metabolite signal (top) and metabolite T_1 (bottom), as a percentage difference from the ground truth. Right) Coefficient of variation on the mean estimate of the fully relaxed metabolite signal (top), and metabolite T_1. Both are plotted as a function of the second TR in the experiment (y axis) and the coefficient of variation on the signal at both TRs (x axis). 90

Figure 5.8: Left) Mean estimate of the fully relaxed signal (first and third row) and T_1 (second and fourth rows), as a percentage difference from the ground truth. Right) Coefficient of variation on the mean estimate of the fully relaxed signal (first and third row), and T_1 (second and fourth row). The top four panels relate to TR_1/TR_2=1500/3500ms pair, the bottom four to TR_1/TR_2=1500/5000ms. 91

Figure 6.1: Estimated [tNAA] with both PRESS and STEAM protocols in 11 neonatal subjects, analysed with LCModel. 94

Figure 6.2: Bland-Altman plot of PRESS and STEAM [tNAA] results. Positive differences indicate STEAM > PRESS. 95

Figure 6.3: Representative spectral fits from three subjects. NAA fit components (red) are shown separate from Lip/MM fit components (grey). Arrows indicate apparent pathological changes in Subject 9, which appear to impact on NAA quantitation accuracy. 95

Figure 6.4: Illustration of the changes in spectral appearance with increasing TR. The metabolite signals (red) increase in intensity with TR, while the Lip/MM contributions are fully relaxed (grey) due to their short T_1 and so have equal intensity throughout. These differing behaviours could be exploited to improve the fit. 96

Figure 7.1: Illustration of the proposed combined fitting process. The model assumes that Lip/MM components (dark red) have a fixed contribution throughout while the metabolites of interest (light red) undergo recovery with T_1. Both spectra are fitted simultaneously according to the ideal relaxation recovery curve (top right) for each metabolite. This is the process used for [tNAA] calculation as outlined in Figure 7.3. 99

Figure 7.2: Illustration of the standard fitting process, where the two spectra are analysed independently. No fitting information is shared between spectra, and so separate estimates are made for Lip/MM components (dark yellow/magenta) and metabolites of interest (light yellow/magenta) in each spectrum. This strategy may give unfeasible estimates (e.g. T_1<0) when the signals for each metabolite are then fitted to an ideal recovery curve (top right). This is the process used to determine tNAA T_1 in the calculation outlined in Figure 5.2. 99

Figure 7.3: Calculation steps to derive thalamic tNAA concentration, [tNAA], using a simultaneous fit to STEAM spectra recorded in the saturation recovery experiment. Black boxes show the input data, either as raw data, HLSVD fits or reference values from previous work/literature. Red boxes and text denote calculated quantities from the tNAA resonance, and blue boxes and text denote calculated quantities from the water signal. Equations accompany the different calculation steps. The relaxation corrected intensities for tNAA and parenchyma water given by S_{NAA} and S_p respectively. 100

Figure 7.4: Typical STEAM spectra (TR=5000ms) simulated for the experiment, with LCModel fits in red. 104

Figure 7.5: Mean normalised amplitude from 20 simulations +/- 1 standard deviation. Deeper colours correspond to higher SNR at a fixed linewidth of 3Hz (top row of Figure 7.4). LCModel derived results are shown in red, TARQUIN derived results are shown in blue. 106

Figure 7.6: Mean normalised amplitude of the lipid and macromolecule resonances (at 2.0ppm) from 20 simulations +/- 1 standard deviation. The results are for individual TARQUIN fits to the spectrum at each TR (SNR=68). 106

Figure 7.7: Mean normalised amplitude from 20 simulations +/- 1 standard deviation. Deeper colours correspond to sharper linewidth at the highest SNR (left column of Figure 7.4). LCModel derived results are shown in red, TARQUIN derived results are shown in blue. 107

Figure 7.8: 2D surface plot displaying the effect of varying the regularisation parameter, λ, on the spline fit to an ideal sinusoidal function. The y axis is on a $\log_{10}$ scale, and red indicates lower values while yellow indicates higher values. Higher regularisation increases the smoothness of the fit. 112

Figure 7.9: Fitted fully relaxed signals for (from top to bottom row) tNAA, tCr, Cho and Glx, when using Provencher's MM parameterisation either with soft constraints on amplitude ratios (left column) or without the soft constraints (right column). Results are shown on the same scale, with the signal on the y-axis in arbitrary units, along with CRLBs (error bars). 118

Figure 7.10: Estimated [tNAA] using the STEAM protocol data and proposed fitting method in comparison to the assumed ground truth (black) values from the PRESS acquisition in Section 6. Provencher (no) soft con = Provencher's MM model with (without) soft constraints on signal amplitude ratios; 9MM no soft con = nine MM model without soft constraints on signal amplitude ratios. 119

Figure 7.11: Bland-Altman plot illustrating the difference between estimated [tNAA] and the assumed ground truth values from the PRESS acquisition in Section 6. This corresponds to the Provencher MM parameterisation, without soft constraints. 119

Figure 7.12: Mean (top) and standard deviation (bottom) of the correlation matrices across all 11 subjects, using Provencher's MM parameterisation and λ=1000, with no soft constraints on the signal amplitudes. 'no CH_3' indicates multiplet resonances fitted independently from the metabolite's singlet peak. 120

Figure 7.13: Representative spectral fit for the STEAM dataset in one of the subjects, with Lip/MM components (blue) extracted without prior knowledge. Note the resemblance of the Lip/MM component with metabolite-nulled spectra presented elsewhere (Schaller et al., 2014, Snoussi et al., 2015, Terpstra et al., 2016, Giapitzakis et al., 2018). 121

Figure 7.14: Bland-Altman plot illustrating the difference between estimated [tNAA] and the assumed ground truth values from the PRESS acquisition in Section 6. Here the Lip/MM components have been estimated without prior knowledge. 121

ABBREVIATIONS AND ACRONYMS

AcAc	Acetoacetate
Acn	Acetone
Act	Acetate
Ala	Alanine
AQSES	Automated Quantification of Short Echo-Time MRS
Asc	Ascorbate
Asp	Aspartate
AUC	Area under the receiver operating characteristic curve
Bet	Betaine
Cho	Choline
Cr	Creatine
CRLB	Cramér-Rao Lower Bound
ECC	Eddy Current Correction
Etm	Ethanolamine
FID	Free Induction Decay
FWHM	Full width at half-maximum
GABA	Gamma-aminobutyric acid
Glc	Glucose
Gln	Glutamine
Glu	Glutamate
Glx	Glutamine + glutamate
Gly	Glycine
GPC	Glycerophosphocholine
GSH	Glutathione
HIE	Hypoxic Ischaemic Encephalopathy
Ins	Myo-inositol
jMRUI	Java-based Magnetic Resonance User Interface
Lac	Lactate
LCModel	Linear Combination of Model Spectra
Lip	Lipid
MARBLE	Magnetic Resonance Biomarkers in Neonatal Encephalopathy
MM	Macromolecule
MR(I/S)	Magnetic Resonance (Imaging/Spectroscopy)
NAA	N-acetylaspartate
NAAG	N-acetylaspartylglutamate
PCh	Phosphocholine
PCr	Phosphocreatine
PE	Phophoethanolamine
PGC	Propylene glycol
PRESS	Point Resolved Spectroscopy
QUALITY	Quantification Improvement by Converting Lineshapes to the Lorentzian Type
RF	Radiofrequency
Scyllo	Scyllo-inositol
Ser	Serine
SNR	Signal-to-noise ratio
STEAM	Stimulated Echo Acquisition Mode
TARQUIN	Totally Automatic Robust Quantitation in NMR
Tau	Taurine
tCho	Choline + glycerophosphocholine + phosphocholine
tCr	Creatine + phosphocreatine
TD	Longitudinal recovery delay time
TE	Echo Time
Thr	Threonine
tLac	Lactate + threonine
TM	Mixing time
tNAA	N-acetylaspartate + N-acetylaspartyl glutamate
TR	Repetition Time
VARPRO	Variable Projection
VeSPA	Versatile Simulation, Pulses and Analysis

SYMBOLIC NOTATION

$\mathbf{a}$	vector of signal amplitudes
a_k	signal amplitude for metabolite k
B	baseline
$\boldsymbol{B_0}$	main static magnetic field
$\boldsymbol{B_1}$	applied RF magnetic field
$\boldsymbol{B_e}$	effective magnetic field
$\mathcal{B}$	B-spline matrix
$\boldsymbol{c}$	vector of B-spline coefficients
c_m	mth B-spline coefficient
d	delay in the time domain signal
$\boldsymbol{D}$	second order differential operator
E	energy
f	spectrum frequency index
f_L	left spectral frequency limit
f_R	right spectral frequency limit
F	Fisher information matrix
Δf	digital resolution of the spectrum
G	Gaussian component of signal decay
G_x	gradient amplitude in the x direction
J	scalar coupling constant
$\mathbb{J}$	Jacobian
$\mathbb{J}_{TR}$	Jacobian evaluated at a particular TR
L	Lorentzian component of signal decay
$\boldsymbol{M}$	magnetisation vector
M_0	equilibrium magnetisation
m_k	magnetisation density in region k
$\boldsymbol{m_s}$	secondary spin quantum number
M_x	magnetisation in the x direction
M_{xy}	absolute magnetisation in the xy plane
M_y	magnetisation in the y direction
M_z	magnetisation in the z direction
$[met]$	concentration of metabolite met
n_{met}	number of protons in met peak
n_w	number of protons in water peak
p	penalty term
R_{T1met}	relaxation correction for metabolite T_1
R_{T1w}	relaxation correction for water T_1
R_{T2met}	relaxation correction for metabolite T_2
R_{T2w}	relaxation correction for water T_2
r_u	relative proportion of signal u
$\boldsymbol{S}$	rotational angular momentum (spin)
s	spin quantum number
s_k	basis vector for metabolite k
S_{longT2}	amplitude of non-tissue water signal
S_w	amplitude of total water signal
S_{meas}	measured signal amplitude
S_{met}	amplitude of metabolite signal
S_{NAA}	amplitude of NAA methyl singlet
$S_{NAA,k}$	amplitude of NAA singlet in step k
S_p	amplitude of tissue water signal
α_{flip}	flip angle
T	temperature
t	FID time index
t'	FID time index including delays
T_1	longitudinal relaxation time
$T_{1,k}$	T_1 for metabolite k
T_2	transverse relaxation time
T_2^*	apparent transverse relaxation time
$T_{2,p}$	T_2 of parenchyma tissue water
Δt	digital resolution of the FID
t_e	end point for time domain analysis
t_{FID}	duration of the time domain signal
t_s	start point for time domain analysis
V_k	volume of region k
$\mathbb{V}_x$	voxel dimension in the x direction
$[w]$	pure water concentration
$\boldsymbol{x}$	vector of non-linear model parameters
x_0	slice centre position
x_{sc}	scaled non-linear model parameter
x_{unsc}	unscaled non-linear model parameter
Δx	slice thickness in the x direction
$\bar{x}$	expected value of parameter x
$\boldsymbol{y}$	observed time signal
$\boldsymbol{y}_{TRn}$	observed time signal at the nth TR
$\hat{\boldsymbol{y}}$	modelled time signal
α_k	Lorentzian component to decay of time domain signal for metabolite k
β	Gaussian component to decay of all time domain signals
γ	gyromagnetic ratio
$\boldsymbol{\Gamma}$	lineshape vector in the time domain
δ	chemical shift
ϵ	soft constraint weighting factor
η_{TR}	partial saturation factor
$\boldsymbol{\theta}$	vector of 0th & 1st order phasing terms
θ	0th order phasing term
λ	regularisation parameter in spline fit
$\boldsymbol{\mu}$	magnetic moment
ν_0	RF pulse centre frequency
$\Delta\nu$	RF pulse bandwidth
ξ	factor for imperfect RF excitation
ρ	specific gravity of brain tissue
σ	shielding constant
σ_{noise}	standard deviation of the noise
σ_x	expected standard deviation in x
τ	duration of the RF pulse
$\boldsymbol{\psi}$	scaled basis set of signal vectors
$\boldsymbol{\psi}_k$	component of $\boldsymbol{\psi}$ for metabolite k
$\boldsymbol{\psi}_{TRn,\mathrm{k}}$	$\boldsymbol{\psi}_k$ evaluated at the nth TR
$\boldsymbol{\Psi}_k$	$\boldsymbol{\psi}_k$ including baseline terms
ω_0	Larmor frequency
ω_k	Larmor frequency offset for k
ω_{ref}	Larmor frequency offset for reference signal

OPERATORS

real(■)	real part of a complex valued quantity
[■]	concentration
$■^{\dagger}$	Moore-Penrose inverse
$\mathcal{F}(■)$	Fourier transform
{■}	parameter set
$■^{H}$	Hermitian conjugate
$\|\|■\|\|_1$	L1-norm (i.e. the sum of absolute values)
$■^{(k)}$	evaluated at the kth iteration
$\frac{\partial}{\partial \boldsymbol{z}}■$	derivative with respect to each of the components of vector $\boldsymbol{z}$

1 INTRODUCTION

1.1 HYPOXIC ISCHAEMIC ENCEPHALOPATHY

Where the neonatal brain experiences a transient period of substantially reduced oxygen (hypoxia) and blood flow (ischaemia) at the time of birth, the resultant injury often causes long term disability or in the most severe cases, death within days. This is often the unexpected result of an otherwise complication-free pregnancy, and although the incidence is relatively low in the general population of high-income countries (2-3 cases per 1000 live births), it has profound implications for the infant, their family, and wider society. Of the infants who survive, up to 75% will be diagnosed with cerebral palsy or other lifelong disabilities (de Vries and Jongmans, 2010, Fleiss and Gressens, 2012). As a result, each affected infant and family will need approximately £600,000 in lifetime support costs (Kruse et al., 2009), and litigation claims alone account for a National Health Service (NHS) cost of £130m each year (Anderson, 2013).

Currently, there is just one therapeutic option available – whole body hypothermia – which became the standard of care for babies across the NHS in 2009 (National Institute for Clinical Excellence, 2010). The combination of several large international randomised controlled trials demonstrated that this increases the likelihood of disability-free survival, and an estimated saving of approximately £1bn to the NHS in the four years following its introduction (Azzopardi et al., 2012).

Despite therapeutic hypothermia, up to 50% of affected infants will go on to die, or survive with a moderate or severe disability (Papile et al., 2014). As a result, the research community is urgently searching for potential adjuvant neuroprotectants which can improve the prognosis of these infants.

One of the most difficult aspects for families and clinicians in this scenario is the uncertainty surrounding prognosis at birth. The 'gold standard' for diagnosing a neurodevelopmental abnormality is a clinical neurological examination performed at least 18 months after birth – a substantial period for distressed parents to wait for, after which opportunities for specialised care and rehabilitation may also have been missed.

1.2 MAGNETIC RESONANCE SPECTROSCOPY IN HIE

One potential source of early prognostic information is cerebral magnetic resonance spectroscopy (MRS). This provides an insight into the metabolic dysfunction present in HIE, and as it yields quantitative data, it can provide objective measures of the severity of brain injury.

The application of MRS to HIE has underpinned some of the most important discoveries concerning the associated brain injury: in its nature and time-course as well as its response to neuroprotective agents. Lorek et al. (1994) demonstrated with ^{31}P MRS that there was a rapid metabolic recovery from hypoxic-ischaemia in newborn piglets, which was proceeded by a steady decline after a delay of a few hours. With this concept of a delayed 'secondary energy failure' came the possibility of providing neuroprotection to prevent it altogether. Soon after, Thoresen et al. (1995) used ^{31}P MRS to demonstrate that hypothermia reduced the extent of this delayed injury, providing the first evidence for what is now the only effective therapy in clinical use.

Given the clear metabolic sequelae of HIE, and the widespread availability of MRS on clinical MRI scanners, various investigators have explored the prognostic utility of cerebral MRS when performed soon after the initial injury. A meta-analysis of such studies by Thayyil et al. (2010) highlighted its superior prognostic accuracy when compared to other early measures of brain injury in small single-centre studies, predicting death or moderate/severe disability at 18 months with an area under the receiver operating characteristic curve of 0.96 (standard error 0.02).

Although MRS is clearly a useful prognostic tool, it is not being used widely in clinical settings. A recent review from the international MRS Consensus Group (Öz et al., 2014) highlighted several key areas where MRS can have a substantial impact on the clinical management of patients, including HIE, but concluded that there are still significant barriers that need to be overcome to achieve this. Particularly, they identified a 'lack of standardisation and quality assurance of MR spectroscopy data acquisition and analysis methods', as well as a need for 'robust prospective multicentre clinical trials, where patient selection and treatment meets prespecified criteria and the statistical methodology is set before the trial commences'. Similar concerns were raised by Bhogal et al. (2017), who highlight the variation in analysis practices and the results derived by experts at various institutions, from the same source data. In this thesis, I present work which seeks to address both of these issues: demonstrating the prognostic value of MRS in HIE in the context of a multi-centre study, and proposing a solution for its rapid, accurate and robust measurement in typical clinical settings.

1.3 MARBLE

The Magnetic Resonance Biomarkers in Neonatal Encephalopathy (MARBLE) study was a prospective, multi-centre study with the aim of evaluating the prognostic accuracy of MR biomarkers in HIE within pragmatic clinical settings. This focused primarily on MRS biomarkers which had shown considerable promise in various small and retrospective studies, but also included the collection of diffusion weighted MRI alongside conventional T_1- and T_2-weighted structural MRI, as well as prognostic scoring based on clinical examinations and amplitude integrated electroencephalography. A flowchart of all

patients and their inclusion in various stages of MR data analysis is shown in Appendix II. In this thesis, I focus on the information derived from MRS, although this will form part of a larger publication incorporating all of the prognostic information obtained. In Chapter 4 I refer to these various prognostic measures alongside MRS, in order to highlight the relative accuracy of each measure in clinical context.

Patients were recruited consecutively, to avoid selection bias. However, the full MRS protocol (including absolute metabolite concentrations) had a duration of approximately 25 minutes, meaning that some degree of selectivity may have been present, depending on the degree to which infants were settled during the scans. Despite this, the characteristics of those included in the analysis showed no evidence of selection bias, and so appear representative of the data from the whole cohort.

1.4 THESIS CONTRIBUTIONS

I conducted all of the work presented in this thesis, with the following support:

- I was supported throughout by an NIHR Healthcare Scientist Doctoral Research Fellowship award (HCS DRF-2014-05-013).
- For the MARBLE study, the imaging protocol had already been established at University College London as part of their routine neonatal imaging, and the MR spectroscopy protocol was initially drafted by Dr. Alan Bainbridge, Mr. David Price and Mr. Ern Cady (all Senior MR Physicists at University College London). I further developed the MR spectroscopy protocol under their guidance. I then adapted the protocol from University College London to give comparable data at all other sites. Dr. Bainbridge also constructed two phantoms for use across the UK sites, while I constructed an equivalent phantom for the study site in the USA (Wayne State University, Detroit).
- I trained all local radiographer teams to perform study protocols accurately, and they collected all *in-vivo* MR data in sites outside of Imperial College London. All MR data collected at Imperial College London was acquired by either myself, Ms. Julie Fitzpatrick (Research MR Radiographer, Imperial College London) or Mr. Solomon Adjakloe (Research MR Radiographer, Imperial College London).
- T_1- and T_2-weighted MR imaging was reported by Dr Sudhin Thayyil (Head of Academic Neonatology, Imperial College London)
- Dr Paolo Montaldo (MRC Doctoral Fellow, Imperial College London) performed the neurological outcome assessments at 2 years of age, along with local teams at sites where another trained examiner was available.

- The recruitment of study subjects and their clinical assessments were performed by local clinical teams.
- The MARBLE study was conceived and developed by Dr. Thayyil, and was funded by his NIHR Clinician Scientist award.
- Dr. Montaldo, Ms. Vânia Oliveira (NIHR Doctoral Fellow, Imperial College London), and Ms. Josephine Mendoza (Research Nurse, Imperial College London) maintained the study database and entered all clinical data.

I developed all novel software using freely available open-source software packages or in-built functions in MATLAB, Python, BASH or C++.

2 NUCLEAR MAGNETIC RESONANCE AND *IN-VIVO* SPECTROSCOPY

In this chapter I summarise theoretical aspects of magnetic resonance (MR) and *in-vivo* MR spectroscopy which are pertinent to the research presented in later chapters.

2.1 NUCLEAR MAGNETIC RESONANCE

The phenomenon of nuclear magnetic resonance (or magnetic resonance, MR) was first demonstrated in condensed matter in 1946 (Bloch et al., 1946, Purcell et al., 1946).

2.1.1 Magnetic moment and the Larmor frequency

Nuclei with a non-zero spin quantum number (s) possess a quantised rotational angular momentum (or spin) $\boldsymbol{S}$, which is directly proportional to its magnetic moment, $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = \gamma \boldsymbol{S} \tag{2.1}$$

where γ is a unique proportionality constant for each nucleus, termed the gyromagnetic ratio. When an external magnetic field $\boldsymbol{B_0}$ is applied to such nuclei, the discrete values of $\boldsymbol{\mu}$ ensure that that they can only possess one of the discrete magnetic energy eigenstates E due to Zeeman splitting:

$$E = -\boldsymbol{\mu} \cdot \boldsymbol{B_0} = -\mu_z B_0 = -\gamma m_s \hbar B_0 \tag{2.2a}$$

where

$$m_s = -s, (-s+1), \dots, (s-1), s \tag{2.2b}$$

and where we have applied $\boldsymbol{B_0}$ along the z direction, and μ_z denotes the z-component of $\boldsymbol{\mu}$. In the case of ^{1}H nuclei, where $s = 1/2$, there are two energy levels corresponding to $m_s = \pm 1/2$ which are separated by ΔE:

$$\Delta E = \gamma \hbar B_0 = \hbar \omega_0 \tag{2.3}$$

where ω_0 is termed the Larmor frequency. In the classical description of the magnetic moment, these two energy states correspond to precession of the spins at a frequency equal to ω_0, either parallel ($m_s = +1/2$) or anti-parallel ($m_s = -1/2$) to the direction of $\boldsymbol{B_0}$. In order to visualise the spin precession more clearly a rotating reference frame is typically used, which accounts for the constant ω_0 rotation.

2.1.2 Bulk magnetisation

When describing the behaviour of larger systems, the magnetic moment of each nucleus is instead replaced by a local magnetic moment per unit volume, termed the magnetisation $\boldsymbol{M}$. Within a given

volume there will be a collection of nuclei, each with their own discrete $\boldsymbol{\mu}_i$, which will each have a discrete E under an external magnetic field $\boldsymbol{B_0}$. In the ^{1}H nucleus, the population of the two possible energy states at thermal equilibrium is described by a Boltzmann distribution:

$$\frac{N_{parallel}}{N_{anti-parallel}} = \exp\left(\frac{\gamma\hbar B_0}{k_B T}\right) \tag{2.4}$$

where k_B is the Boltzmann constant and T is the sample temperature (in Kelvin). The small excess of nuclei in the lower energy parallel configuration results in $\boldsymbol{M}$ having the same direction as $\boldsymbol{B_0}$ at thermal equilibrium (i.e. along the +z-axis). It is important to note that, since there is no coherence between the precession of the individual nuclear spins, there is no component of $\boldsymbol{M}$ perpendicular to $\boldsymbol{B_0}$ (i.e. in the x-y plane).

2.1.3 Excitation

In order to observe a resonance signal, it is necessary to disrupt the thermal equilibrium state. A second magnetic field, $\boldsymbol{B_1}$, is therefore applied perpendicular to $\boldsymbol{B_0}$, and oscillating at frequency ω_0. The $\boldsymbol{B_1}$ field increases the phase coherence in the precession of the nuclear spins, while also exciting a proportion of the nuclei from the parallel state to the anti-parallel state, reducing the net longitudinal magnetisation. The net effect is that $\boldsymbol{B_1}$ exerts a classical torque on $\boldsymbol{M}$, causing it to rotate away from the +z direction so that it has some component in the x-y plane.

In practice the $\boldsymbol{B_1}$ field is achieved with a radiofrequency (RF) pulse, which is applied for a duration τ sufficient to rotate $\boldsymbol{M}$ through a flip angle α_{flip}:

$$\alpha_{flip} = \gamma B_1 \tau \tag{2.5}$$

where $\alpha_{flip} = \pi/2$ (i.e. a 90° pulse) describes rotation into the x-y plane, and $\alpha_{flip} = \pi$ (i.e. a 180° pulse) describes complete inversion of the magnetisation.

2.1.4 Localisation

If a linear magnetic field gradient is applied along an object in a given direction, the frequency of the constituent spins will also vary linearly in the same direction. By applying the RF pulse over a finite range of frequencies (the 'bandwidth', $\Delta\nu$), only a selection of the spins is excited.

The thickness of the excited region, Δx, is given by:

$$\Delta x = \frac{2\pi}{\gamma G_x}\Delta\nu \tag{2.6}$$

where G_x is the magnitude of the magnetic field gradient. Similarly, the slice's position offset (x_0) can be chosen by selecting an appropriate RF central frequency offset (ν_0):

$$x_0 = \frac{2\pi}{\gamma G_x}\nu_0 \tag{2.7}$$

To obtain an ideal RF rect function in the frequency domain, the temporal variation of the RF amplitude must be described by its Fourier transform: a sinc function. In practice, the use of short duration RF waveforms means that frequency selectivity is imperfect.

2.1.5 T_1 relaxation

Following excitation, $\boldsymbol{M}$ returns to its thermal equilibrium state via separate relaxation processes which affect the longitudinal (M_z) and transverse (M_{xy}) components.

Longitudinal (T_1) relaxation occurs as nuclei in the anti-parallel state move to the parallel state at an exponentially decreasing rate, eventually restoring thermal equilibrium after a sufficient time. The expression for M_z during T_1 relaxation after a 90° pulse is:

$$M_z(t) = M_0\left(1 - e^{-t/T_1}\right) \tag{2.8}$$

where M_0 denotes the longitudinal magnetisation at thermal equilibrium. The time between successive excitations is called the repetition time (TR).

2.1.6 T_2 and T_2^* relaxation

Transverse (T_2) relaxation occurs as the phase coherence of the nuclear spins is lost, returning to the fully incoherent state at thermal equilibrium. The expression for M_{xy} during T_2 relaxation after a 90° pulse is:

$$M_{xy}(t) = M_{xy}(0)e^{-t/T_2} \tag{2.9}$$

In reality the $\boldsymbol{B_0}$ field is never perfectly homogeneous across a sample, and M_{xy} decays at a rate faster than T_2 due to the resultant variation in the precession frequencies. This apparent transverse relaxation time (T_2^*) is given by:

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'} \tag{2.10}$$

where T_2' is the decay constant due to B_0 inhomogeneities alone.

2.1.7 The Free Induction Decay

In the magnetic resonance experiment, two orthogonal detection coils are positioned in the x-y plane. Immediately following a 90° excitation, $\boldsymbol{M}$ precesses in the transverse plane and induces a time-varying electromotive force in the detection coils, with the observed magnetisation along each axis as:

$$M_x(t) = M_{xy}(0)\cos[(\omega_0 - \omega)t + \theta(0)]\, e^{-t/T_2^*} \tag{2.11a}$$

$$M_y(t) = M_{xy}(0)\sin[(\omega_0 - \omega)t + \theta(0)]\, e^{-t/T_2{}^*} \qquad (2.11b)$$

where $\theta(0)$ is the initial phase of $\boldsymbol{M}$ in the x-y plane at $t = 0$. This is termed the Free Induction Decay (FID). In practice, this is rarely measured since it comes immediately after the transmitted RF pulse, and so is often distorted due to RF breakthrough. Instead, the magnetisation is usually observed through 'echoes' which can be made to occur later.

2.1.8 Spin echoes

Although M_{xy} decays rapidly due to $T_2{}^*$ relaxation, a 180° pulse can be applied in an orthogonal direction in the x-y plane to partially restore phase coherence in the precession of the excited spins.

During the FID, the inhomogeneities in B_0 cause the accumulation of additional phase dispersions in the spins. The 180° pulse reverses this phase dispersions (from +$\Delta\phi$ to -$\Delta\phi$), so that as they continue to accumulate, the net phase dispersion will reach zero at the echo time (TE). As the phase accumulation occurs at a constant rate, the 180° pulse placed at TE/2 generates a spin echo centred at TE.

The intensity of the observed signal is lower than the FID signal by a constant factor due to T_2 decay, with the observed magnetisation from TE onwards given by:

$$M_x(TE + t) = M_{xy}(0)e^{-TE/T_2}\cos[(\omega_0 - \omega)t + \theta(TE)]\, e^{-t/T_2{}^*} \qquad (2.12a)$$

$$M_y(TE + t) = M_{xy}(0)e^{-TE/T_2}\sin[(\omega_0 - \omega)t + \theta(TE)]\, e^{-t/T_2{}^*} \qquad (2.12b)$$

in the absence of any magnetic field gradients.

2.1.9 The Discrete Fourier Transform

Since the detected signal is circularly polarised, it can be conveniently represented as a time varying complex function according to Euler's formula ($e^{j\theta} = \cos\theta + j\sin\theta$). To determine the frequency components of a given time domain signal, the Fourier transform can then be employed, which models the time signal as a series of sine and cosine functions. The amplitudes of the individual components of the Fourier series can then be expressed as a function of their frequency, resulting in a spectrum.

When the complex signals are recorded over a series of n discrete time points, the discrete Fourier transform returns a complex spectrum with n discrete points. Depending on the time spacing (or 'dwell time' Δt) and total duration of the signal measurement for a given FID (t_{FID}), the frequency spectrum may appear with different resolution (Δf) and bandwidth ($\Delta\nu$):

$$\Delta f = \frac{1}{t_{FID}} = \frac{1}{n\Delta t} \qquad (2.13a)$$

$$\Delta\nu = \frac{n}{t_{FID}} = \frac{1}{\Delta t} \qquad (2.13b)$$

Unlike in imaging applications, spectra are typically interpreted in the absorption mode – the in-phase ('real') component of the frequency spectrum – as this has a sharper linewidth than the absolute frequency spectrum, and so peaks at different frequencies can be more readily separated.

2.2 *IN-VIVO* MAGNETIC RESONANCE SPECTROSCOPY

Soon after the discovery of NMR, it was demonstrated that nuclei in the same molecule resonate at different characteristic frequencies (Dickinson, 1950, Proctor and Yu, 1950). This discovery paved the way towards magnetic resonance spectroscopy (MRS).

2.2.1 Chemical shift

As an external $\boldsymbol{B_0}$ field is applied to a sample, the electrons which surround the constituent nuclei rotate in such a way as to shield (in general) those nuclei from the induced change. Under these circumstances, the moving charge from the electrons generates a magnetic moment $\boldsymbol{\mu_e}$ which acts in the opposite direction to $\boldsymbol{\mu}$ (and hence $\boldsymbol{B_0}$). The effective field at the nucleus $\boldsymbol{B_e}$ is then given by:

$$\boldsymbol{B_e} = \boldsymbol{B_0}(1 - \sigma) \qquad (2.14)$$

where σ is the shielding factor, which depends on the local chemical environment. Due to this effect, nuclei in different chemical environments resonate at different frequencies. These variations in frequency are usually expressed in parts per million (ppm):

$$\delta(\text{ppm}) = \frac{\omega - \omega_{ref}}{\omega_{ref}} \times 10^6 \qquad (2.15)$$

where ω is the frequency of the nuclei of interest and ω_{ref} is the frequency of a reference compound, usually chosen as tetramethylsilane. Metabolites can therefore be identified by their characteristic chemical shifts, using an MRS experiment.

2.2.2 Scalar coupling

The chemical environment of nuclei is also determined by the bonds to neighbouring atoms, and the electron distribution therein. Scalar coupling refers to the interaction between neighbouring magnetic moments through these intervening electrons, and is unlike the dipolar interactions which act directly through space and cause relaxation. Whereas dipole interactions average to zero due to the rapid tumbling motion of molecules over time, this scalar coupling is always present and unchanged at the same locations within a given molecule.

Depending on the frequency difference between the interacting nuclei, scalar coupling has different effects on the spectral appearance. In the case where the frequency difference between the two nuclear spins is much larger than the scalar coupling constant, J, the spins constitute a 'weakly coupled' system. Alternatively, where J is comparable to the frequency difference, the spins are in a 'strongly coupled' system.

In a heteronuclear two-spin system 'AX', the frequency difference between the two spins is sufficiently large (on the order of MHz) that they are weakly coupled. In a homonuclear two-spin system 'AB', the nuclei can be either weakly or strongly coupled. The effects of weak coupling can be predicted by simple rules:

- No splitting occurs between two spins which are 'magnetically equivalent'. That is, where both have the same δ, and have the same J with a shared third spin at a different δ (for example, the ^{1}H nuclei in a -CH_3 group).
- Such magnetically equivalent spins cause splitting on other spins, in a pattern which follows a binomial distribution (for example, the three ^{1}H nuclei in the -CH_3 group of lactate split the -CH group into a 1:3:3:1 pattern)
- A spin interacting with more than one non-magnetically equivalent spins will have a complex pattern, as it is split by both.

On the other hand, the effects of strong coupling are difficult to predict, and require calculations based on the density matrix formalism.

2.2.3 Point Resolved Spectroscopy (PRESS)

The timing diagram for the PRESS pulse sequence (Bottomley, 1984) is shown in Figure 2.1. Following a slice selective 90° pulse, two further 180° pulses are applied to refocus a subset of the spins, so that the intersecting region between all three pulses creates an echo at time TE. The three RF pulses are usually chosen to select mutually orthogonal slices, so that the echo is generated from a cuboid region.

Strong crusher gradients are applied either side of the 180° pulses to dephase any FIDs they produce, avoiding contamination of the spin echo magnetisation. The amplitude of the observed magnetisation at the echo can then be approximated by:

$$M_{xy}(TR, TE) = M_0\left(1 - e^{-TD/T_1}\right)e^{-TE/T_2} \qquad (2.16a)$$

assuming ideal RF pulses, and where the time delay (TD) for PRESS is given by (Knight-Scott, 2011):

$$TD = TR - TE \qquad (2.16b)$$

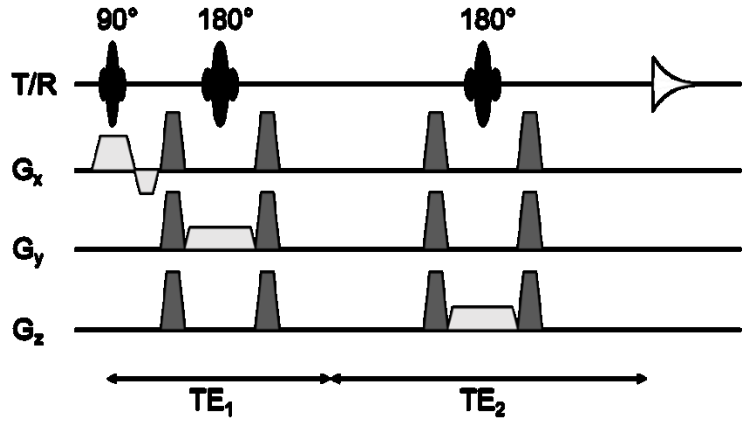

Figure 2.1: Timing diagram for 3D single voxel localisation with the PRESS sequence. The top line depicts the transmit/receive (T/R) events, and the gradient waveforms in three orthogonal directions (G_x, G_y, G_z) are shown on the lower three lines. RF pulses are shown in black, while the detected echo is shown in white. Selection and refocusing gradients (light grey) ensure detection of coherent signals from the volume of interest. Crusher gradients (dark grey) prevent the coherence of unwanted signals. The echo is detected at time TE (=TE_1+TE_2).

2.2.4 Stimulated Echo Acquisition Mode (STEAM)

The timing diagram for the STEAM pulse sequence (Granot, 1986) is shown in Figure 2.2. Unlike PRESS, this uses three 90° pulses, and results in a stimulated echo.

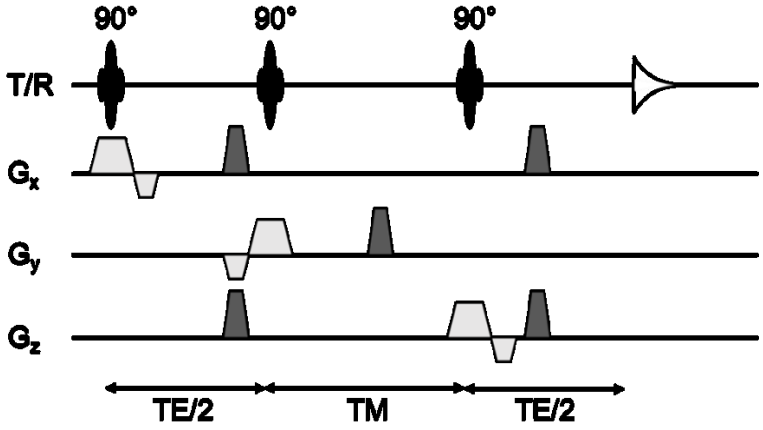

Figure 2.2: Timing diagram for 3D single voxel localisation with the STEAM pulse sequence. The top line depicts the transmit/receive (T/R) events, and the gradient waveforms in three orthogonal directions (G_x, G_y, G_z) are shown on the lower three lines. RF pulses are shown in black, while the detected echo is shown in white. Selection and refocusing gradients (light grey) ensure detection of coherent signals from the volume of interest. Crusher gradients (dark grey) prevent the coherence of unwanted signals. The echo is detected at time TM+TE.

Although FID signals are generated after each 90° pulse along with an additional four spin echoes, the stimulated echo is isolated by using crusher gradients and phase cycling (see Section 2.4.5).

The amplitude of the observed magnetisation at the stimulated echo is given by:

$$M_{xy}(TR, TM, TE) = 0.5 \times M_0 \left(1 - e^{-\mathrm{TD}/T_1}\right) e^{-\mathrm{TM}/T_1}\, e^{-TE/T_2} \qquad (2.17a)$$

assuming ideal RF pulses and no diffusion effects, and where the time delay (TD) for STEAM is:

$$TD = TR - TM - \mathrm{TE}/2 \qquad (2.17b)$$

Typically, the mixing time (TM) is kept short so that the additional saturation term in Equation 2.17a ($e^{-\mathrm{TM}/T_1}$) is approximately equal to unity. It is important to note from the above expressions that STEAM suffers from a theoretical 50% signal loss in comparison to PRESS, due to the incomplete refocusing of the 90° pulses. However, STEAM has some desirable characteristics:

- The minimum achievable TE is shorter than in PRESS, resulting in less signal reduction due to T_2 decay
- The B_1 requirement for 90° pulses is lower than for 180° pulses, so the bandwidth can be higher, resulting in less pronounced chemical shift artefacts (see Section 2.2.1)
- The frequency profile of 90° pulses is sharper than for 180° pulses, so the region producing the signal is more clearly defined

2.3 METABOLITES OF INTEREST

An exhaustive list of observable metabolites in *in-vivo* MRS and their properties is given by Govindaraju et al. (2000) and various others. The following sections outline some of the key metabolites discussed in later chapters, and outline their prominent spectral patterns in the range of 0.0 to 4.2ppm (as is typically used for analysis).

2.3.1 N-acetylaspartate (NAA)

NAA is an amino acid present only in the nervous system, and produces the most prominent peak in the healthy brain. The acetyl CH_3 group of NAA produces a prominent singlet peak at 2.01ppm, while multiplet resonances at 2.49ppm and 2.67ppm appear due to its aspartate CH_2 group.

NAA is produced by the reaction of acetyl-CoA and aspartate in the neuronal mitochondria, with this aspartate produced from glutamate in the tricarboxylic acid cycle. It is then transported out of neurons into oligodendrocytes, where it is catabolised. Since the rate of transport out of neurons is slower than the accumulation, NAA reaches high concentrations inside the neurons (Edden and Harris, 2016).

Reductions in [NAA] have been compared to histology in animal models of ischaemia (Sager et al., 2001) in order to determine its utility as a specific marker of neuronal density. While Sager et al. (2000) found that [NAA] correlates with neuron count in mice, as neuron density tends to zero [NAA] remains at low but finite levels.

The interpretation of [NAA] as solely a neuron density marker is also complicated by several studies which suggest reduction is reversible. De Stefano et al. (1995) outlined apparent reversible NAA changes in mitochondrial encephalopathy with lactic acidosis and stroke-like episodes (MELAS), which could only be plausibly explained by concentration changes. In this condition there is a primary mitochondrial injury which may temporarily interfere with NAA production during each episode. They also demonstrated reversible changes in NAA in multiple sclerosis lesions, which correlated with clinical outcome scores.

In traumatic brain injury, Signoretti et al. (2001) used high-performance liquid chromatography to demonstrate recovery of [NAA] in the rat brain following moderate diffuse injury, and a lack of recovery in severe diffuse injury. More recently, Shannon et al. (2016) used microdialysis to demonstrate an increase in extracellular NAA soon after injury, which gradually reduced over time. These together suggest [NAA] as measured by MRS may additionally reflect physical damage to neuronal membranes, as well as metabolic dysfunction.

In stroke, there is some confusion in the literature over whether [NAA] is reversible. Although Brulatout et al. (1996) present a recovery in metabolism after global ischaemia, the changes in NAA cannot be attributed to concentration alone, and may also be caused by pathological changes in its relaxivity, as well as the other metabolites used as denominators in peak area ratios. Muñoz Maniega et al. (2008) performed serial studies in ischaemic stroke suggesting a rebound effect in NAA, but they also did not perform appropriate corrections for relaxation effects, and so it is unclear whether this reflects true concentration changes. This is particularly problematic as they noted a correlation between the measured NAA levels and the characteristics of water diffusion in the corresponding tissue.

Such confounding issues in stroke were discussed by Walker et al. (2004) (and previously van der Toorn et al. (1994)), who specifically warn against misinterpretations of [NAA] in stroke based on pathological dynamic changes in the T_2 of NAA and water diffusivity. In some cases, this has instead been misinterpreted as evidence of [NAA] deviating from changes in neuronal density, and being reversible (see for example the review by Moffett et al. (2007) with nearly 800 citations at the time of writing). Although, for example, Bivard et al. (2014) are careful to state that hyperperfused peri-infarct tissue 'preserved NAA compared with controls', this has also been misinterpreted as a reversible [NAA] change (Edden and Harris, 2016), despite more widespread evidence of a general decrease (see for example van der Toorn et al. (1994), Saunders et al. (1995), Sager et al. (1995), Wardlaw et al. (1998)).

Since changes in [NAA] may occur through one of many metabolic and microstructural mechanisms, it is pertinent to interpret this broadly, as a deviation from the normal neuronal status.

2.3.2 N-acetylaspartylglutamate (NAAG)

NAAG is produced by the reaction between NAA and glutamate, and is present in much lower concentrations than NAA in the brain (Moffett and Namboodiri, 2006). The acetyl CH_3 group of NAAG produces a prominent singlet peak at 2.04ppm, which typically overlaps with the NAA peak at 2.01ppm. Although these peaks are separable at high field and with exceptional $\boldsymbol{B_0}$ homogeneity, in practice this is not achievable on clinical MR scanners and the two peaks display high covariance during fitting. As a result, the signals arising from both NAA and NAAG are typically summed during quantification and termed tNAA for brevity.

2.3.3 Choline (Cho), Glycerophosphocholine (GPC) and Phosphocholine (PCh)

Choline containing compounds are essential for the production and breakdown of phospholipids, and are frequently cited as markers of cell membrane turnover and synthesis. The $(CH_3)_3$ group of Cho produces a prominent singlet peak at 3.19ppm, while the equivalent group is observed at 3.21ppm on spectra from each of GPC and PCh. Additional multiplet resonances are observed from these metabolites in the range of 3.5-4.1ppm. Due to the difficulty in separating the components of Cho, GPC and PCh in a given spectrum, the contribution from all three is typically summed during quantification and termed tCho for brevity.

In several types of brain tumours, elevated [tCho] is considered an indicator of malignancy and correlates with cellular markers of proliferation (Herminghaus et al., 2002, Peet et al., 2007), likely due to the increased rate of cell synthesis and tumour density. However, Howe et al. (2003) warn against interpretation of [tCho] changes in the presence of necrosis, where cell density reduces.

Bagory et al. (2007) demonstrated diffuse increases in brain [tCho] in primary-progressive multiple sclerosis in comparison to the relapsing-remitting phenotype, although without correcting for potential pathology-related relaxivity effects. This suggests [tCho] increases may be associated with the ongoing inflammation and active demyelination, corresponding with the decreases in [tCho] observed with myelination in the healthy neonatal brain (Cady et al. (1996), more detail in Section 3).

2.3.4 Creatine (Cr) and Phosphocreatine (PCr)

Both Cr and PCr are contained in the neuronal and glial cells, are both essential in energy metabolism and together act as an energy buffer which maintains constant levels of adenosine triphosphate (ATP). The CH_3 groups of Cr and PCr produce prominent singlets at 3.03ppm, while their CH_2 groups produce singlets at 3.91ppm and 3.93ppm respectively. Due to the strong overlap between these peaks, the sum of Cr and PCr signals is typically reported and termed tCr for brevity.

Following focal ischemia, van der Toorn et al. (1994) demonstrated a significant gradual decrease in [tCr] in the cat brain, coincident with T_1 and T_2 increases over the course of hours. Toorn et al. (1995) then also described similar increases in the T_1 and T_2 of tCr following focal ischaemia in rats. Ke et al. (2002) demonstrated that such changes can be related to the relative proportion of Cr and PCr contributing to the measured signal, as the former has a substantially longer T_2 than the latter. This raises the possibility of using the relaxivity of the tCr resonance as a probe of energy status, in the absence of ^{31}P MRS. Meanwhile, since the energy buffering tCr pool should remain intact during transient changes between aerobic/anaerobic metabolism, decreases in [tCr] would suggest cell death in this context.

2.3.5 Lactate (Lac)

Lac is the end product of anaerobic glycolysis, and is only present in healthy brain at low concentrations. The CH_3 group of Lac produces a prominent doublet at 1.31ppm, while the CH group is split into a quartet at 4.10ppm.

Lac is a sensitive marker of anaerobic conditions, and is therefore increased profoundly in various pathologies, such as stroke (Bruhn et al., 1989), hypoxia (Behar et al., 1983), and high-grade tumours (Howe et al., 2003).

Pellerin and Magistretti (1994) have also suggested that, under increased neuronal activity, astrocytes increase their glucose consumption and release Lac into the extracellular space as a substrate for additional neuronal production of ATP (the 'astroglial-neuronal lactate shuttle (ANLS) hypothesis'). Although additional Lac is reported in areas of increased neuronal activity in visual stimulation during functional MR spectroscopy (for example, Prichard et al. (1991), Bednařík et al. (2015), and Bednarik et al. (2018)), there is controversy over the role of this Lac both in healthy and pathological brain (Mason, 2017).

2.3.6 Threonine (Thr)

Thr is an essential amino acid formed from aspartate, and its CH_3 group produces a prominent doublet at 1.32ppm, overlapping with the Lac doublet at 1.31ppm. Since the scalar coupling interactions of the CH_3 groups have similar J (J_{Thr} = 6.35 Hz, J_{Lac} = 6.93 Hz), the doublets are almost coherent at appropriate TEs, and essentially indistinguishable. This sum is termed tLac, as suggested by Hofmann et al. (2002), for brevity.

2.3.7 Macromolecules (MM) and Lipids (Lip)

MM components apparent in brain ^{1}H MRS correspond to high molecular weight components, and have been attributed largely to methyl and methylene resonances from various cytosolic proteins. Lip

components correspond either to lipids in the healthy scalp, or the pathological accumulation of lipids in brain tissue.

Pathological changes in Lip and MM resonances are observed in high-grade tumours (Opstad et al., 2008b), multiple sclerosis (Mader et al., 2001), and stroke (Graham et al., 2001). In stroke, Saunders et al. (1997) have proposed a number of plausible mechanisms for these changes, including the liberation of free fatty acids or infiltration by macrophages. More recently, imaging methods of MM proton fraction mapping have been observed to agree with histological and immuno-histochemical measures of demyelination in ischemic stroke (Khodanovich et al., 2018) giving some weight to the suggestion of Mader et al. (2001) that increases in MM signals in part represent myelin fragments. Opstad et al. (2008a) also demonstrated with that increased lipid signals evident in tumours agreed with the deposition of lipid droplets as identified by histochemistry, suggesting that these form in response to hypoxic stress and precede necrosis. Finally, Marjańska et al. (2018) have recently demonstrated substantial age-related variation in MM signals in younger (21±1 year old) and older (75±6 year old) healthy adults.

Považan et al. (2018) suggest MM parameterisation with broad peaks at the following locations: 0.90ppm (MM1), 1.21ppm (MM2), 1.43ppm (MM3), 1.67ppm (MM4), 2.04ppm (MM5), 2.26ppm (MM6), 2.99ppm (MM7), 3.21ppm (MM8), 3.77ppm (MM9). An additional peak at 4.3ppm (MM10) is outside of the typical analysis range. Provencher (1993) parameterises cerebral lipid signals by independent broad singlets due to (-CH_2-C**H**$_3$) groups at 0.89ppm and (-CH_2-C**H**$_2$-CH_3) groups at 1.28ppm, as well as a fixed combination of broad singlets at 1.95ppm, 2.08ppm, 2.25ppm, and 3.00ppm.

2.4 Acquisition considerations

When designing an *in-vivo* MRS experiment, there are a number of physical phenomena that must be accounted for. These can otherwise reduce the quality of the resultant data, impacting on the ability to perform accurate quantification.

2.4.1 Linewidth and shimming

The T_2^* decay rate determines both the duration of the signal decay envelope in the time domain representation, and the linewidth of resultant peaks in the frequency domain:

$$FWHM = {}^{1}/_{\pi T_2^*} \tag{2.18}$$

The $\boldsymbol{B_0}$ homogeneity within the region of interest is directly proportional to the rate of T_2^* decay, and so reducing the inhomogeneity both increases the duration of measurable signal in the time domain of

the FID and sharpens the peaks in its frequency domain representation. This can be achieved by iteratively adjusting the current through shim coils positioned around the magnet, optimising either the linewidth (minimum) or the duration of the T_2^* decay envelope (maximum).

Although the metabolite signal linewidth has no effect on its amplitude in the time domain, various investigators have demonstrated that it biases the quantitative results of fitting in typical *in-vivo* spectra (Tkac et al., 2002, Kreis and Boesch, 2003, Mangia et al., 2006, Bartha, 2007). As T_2^* shortens, the complicated splitting patterns present in J coupled nuclei are increasingly smeared out through convolution with the broader lineshape in the frequency domain, so that the high-resolution information which uniquely characterises different metabolites is less discernible. Equivalently, in the time domain the decay envelope shortens such that less useful signal is available to fit before decaying into the noise, and the time domain representations of each metabolite appear increasingly similar over that period.

Kreis (2004) proposes some pragmatic rules for identifying whether a linewidth is acceptable for quantitative analysis:

- The linewidth should be less than 0.07-0.10 ppm (approx. <9Hz at 3 Tesla)
- The linewidths should fall within the normal range of the dataset in question (defined as the mean ± 2 standard deviations)

It is important to note that there are cases where an increased linewidth itself reflects pathology, due to local inhomogeneities arising from changes in the magnetic susceptibility of diseased tissue (e.g. acute haemorrhage).

2.4.2 Eddy currents

Whenever a time varying magnetic field is applied to a conductive object, Faraday's law of induction dictates that a circulating current is generated which opposes the change (i.e. eddy currents). This current, in turn, generates an opposing time varying magnetic field. Such eddy currents arise in conductive materials in the magnet housing/shim coils when magnetic field gradients are applied during a pulse sequence. As discussed in the previous section, spectral quality in MRS is highly dependent on the homogeneity of $\boldsymbol{B_0}$, and the presence of eddy currents can cause substantial distortions in spectra.

Even where an FID has been affected by eddy current distortions, these can be ameliorated during post-processing where a separate water non-suppressed reference FID has been obtained in the same voxel. Perhaps the most popular method is ECC proposed by Klose (1990): performing a point-wise division of the FID phase with the phase of the reference FID. Since the RF centre frequency is coincident with the

water signal resonance, the time dependence of the phase of the reference FID is caused by eddy currents alone, and so the division removes such effects from the FID of interest. An alternative method is QUALITY, proposed by De Graaf et al. (1990): dividing each point in the FID of interest by the reference FID signal envelope, before applying pure Lorentzian damping. Although this provides a spectrum with an analytical lineshape, the pointwise signal division suffers at late times in the FID where the signal has decayed into the noise. For this reason, the two are sometimes combined in 'QUECC' (Bartha et al., 2000), where QUALITY is applied to the first n points and ECC is applied to later points, where n is determined heuristically.

An alternative strategy is to directly model distortions in the spectral peaks during the fitting process, which is necessary where a reference FID is not available. This can be modelled in the frequency domain as a convolution with a flexible array of delta functions (Slotboom et al., 1998, Provencher, 2001), or in the time domain using self-deconvolution (Sima et al., 2009). In the latter case, the FID is denoised and the decay envelope approximated by a flexible model.

2.4.3 Chemical shift displacement

When an RF pulse is applied to a region of interest, spins which have different resonant frequencies will be selected in different spatial regions due to chemical shift effects. The displacement between the regions selected for spins at the RF centre frequency (ω_0) and spins at an offset frequency (ω) is as follows:

$$\Delta x = \frac{\omega - \omega_0}{BW} \mathbb{V}_x \tag{2.19}$$

where $\mathbb{V}_x$ is the voxel dimension in the x direction, and BW is the RF pulse bandwidth. Therefore, a choice of high bandwidth RF pulses minimises this regional spread.

For example, the vendor supplied PRESS sequence on some scanners uses 90° RF pulses with a 2.4kHz bandwidth at 3 Tesla, resulting in a 14.3% voxel displacement between the singlet NAA resonance at 2.01ppm and water. The corresponding 180° refocusing pulses have a reduced bandwidth of 1.4kHz, resulting in a 24.6% voxel displacement. These two resonances would therefore be coincidentally selected by all three pulses in 49% of the prescribed voxel during PRESS, whereas a STEAM sequence using three of the same 90° pulses would have fully selected both resonances in 63% of the prescribed volume. Moreover, choosing a 15x15x15mm^3 voxel will result in selection of a water voxel and NAA(-CH_3) voxel with centres separated by 5.6mm in PRESS, and 3.7mm in STEAM.

On GE systems, high bandwidth 'very selective suppression' (VSS) pulses are applied as default alongside the standard volume selection. This selects a sub-region within the excited volume, and

exploits the high bandwidth of the VSS pulses to minimise chemical shift displacement and create a sharp voxel profile.

2.4.4 Anomalous J modulation

Although the nature of weak homonuclear J coupling can be predicted under the assumption of ideal RF pulses, the chemical shift displacement artefacts associated with typical RF pulses cause additional complications. For example, the A and X spins in an AX_n system will be excited in different regions during a PRESS sequence due to the frequency difference. As a result, the J coupling will be different in each sub-region of the PRESS voxel, and the measured signal from the X doublet is the sum of the resultant signals.

Where TE is chosen as the 'positive in phase' condition (e.g. TE=2/J) so that the X doublet resonances are in phase with each other and with singlet resonances elsewhere in the voxel, the observable magnetisation in each sub-voxel can be expressed as follows (Yablonskiy et al., 1998):

$$M_1{}^{\pm} = m_1 V_1 \tag{2.20a}$$

$$M_2{}^{\pm} = m_2 V_2 e^{\pm i\pi J \cdot TE_1} \tag{2.20b}$$

$$M_3{}^{\pm} = m_3 V_3 e^{\pm i\pi J \cdot TE_1} \tag{2.20c}$$

$$M_4{}^{\pm} = m_4 V_4 \tag{2.20d}$$

where m_k is the relaxation-affected magnetisation density in sub-voxel k, V_k is the volume of sub-voxel k, and $\pm$ denotes each of the split components of the doublet signal. The signal loss due to the phase dispersion in sub-voxels 2 and 3 is therefore minimised with a short TE_1, and maximised with $TE_1 = \frac{1}{J}$.

At other TEs, the cancellation effects between sub-voxels are more substantial, and so signal intensities are more heavily affected by the choice of RF bandwidths and sequence timings.

2.4.5 Unwanted coherences

The signal of interest is generated at the echo time, however there are additional undesired signals arising during the course of an experiment which require suppressing. By repeating an experiment and varying the phase of the RF pulses, the signals can be combined in a way which suppresses unwanted signals. This process is termed phase cycling, with longer sequences of phase cycling improving the suppression of unwanted coherences. In addition to phase cycling, crusher gradients can be applied where necessary during the sequence to cause additional dephasing of undesirable transverse signals, while preserving the desirable signals. Stronger crusher gradients are more effective at suppressing unwanted signals, but may introduce some degree of signal loss due to diffusion effects.

2.4.6 Motion

Once a homogeneous $\boldsymbol{B_0}$ has been achieved in an organ such as the brain, small rotational or translational motions will disturb this and introduce $\boldsymbol{B_0}$ fluctuations. These appear as global frequency and phase shifts across the whole recorded spectrum. In cases where the changes in $\boldsymbol{B_0}$ are comparable to the linewidth of an individual spectrum (i.e. 3-8Hz), blind averaging of FIDs will result in a broadened or distorted lineshape.

More substantial motion will result in the selection of tissue outside of the defined region of interest, leading to a potentially erroneous interpretation of the acquired data.

2.5 ANALYSIS CONSIDERATIONS

After acquiring data from an *in-vivo* MRS experiment, there are various strategies to perform post-processing and analysis in order to effectively quantify the metabolite signals.

2.5.1 Pre-processing

Aside from the lineshape corrections outlined in Section 2.4.2, there are various other strategies for post-processing acquired data.

Residual water removal: any water signals which were not fully removed by water suppression can be modelled as a series of damped exponentials within a narrow frequency range of the water signal, and subtracted from the FID. This reduces interference from the broad tails of the water signal on the lower intensity metabolite signals.

Zero-filling: appending the FID with zeros in the time domain has the effect of interpolation in the frequency domain.

Apodisation: multiplying the FID signal by a decaying function suppresses the noise at the end of the signal while broadening the peaks in the frequency domain.

Frequency/phase registration: individual spectra can be aligned in both frequency and phase to account for the effects of small movements, and any hardware-related drifts in $\boldsymbol{B_0}$ over the course of the experiment.

2.5.2 Analysis

A number of software packages are available for use within the research community, although very few are approved for routine clinical use. These each apply a signal model to the acquired data, and typically employ a least squares optimisation strategy to minimise the difference between the model and the data.

Fitting strategies can be generally grouped into those which use a cost function in the time domain (e.g. AMARES (Vanhamme et al., 1997), QUEST (Ratiney et al., 2004), AQSES (Poullet et al., 2007), TARQUIN (Wilson et al., 2011)) or in the frequency domain (e.g. LCModel (Provencher, 2001), TDFDfit (Slotboom et al., 1998), FITaid (Chong et al., 2011)), with the latter requiring additional discrete Fourier transform operations which increase the computational cost.

Although the classical signal model decomposes the FID into a series of Lorentzian spectral peaks, or equivalently a series of decaying exponentials in the time domain, *in-vivo* signals have lineshapes which deviate from this ideal. Where normally distributed $\boldsymbol{B_0}$ inhomogeneities become appreciable, the peaks display a lineshape and decay function which more closely resembles a Gaussian distribution. In practice, *in-vivo* lineshapes display some characteristics of both distributions, and are best represented by a Voigt distribution (Marshall et al., 1997):

$$\Gamma(t) = \exp(-Lt - Gt^2) \tag{2.21}$$

where L and G represent Lorentzian and Gaussian weighting factors, respectively. These may also be modified by the parametric or analytical lineshape models discussed in Section 2.4.2 to account for eddy current distortions. Fitting strategies such as LCModel or TARQUIN, which employ a basis set of characteristic metabolite signals, apply the lineshape model to the full basis function representing each signal component. The signal model generally takes the following form in the time domain, as a combination of linear and non-linear parameters, plus a baseline term (B):

$$\sum_{k=1}^{N} \mathrm{a}_k s_k(t)\Gamma(t)\exp(j\theta + j\omega_k t) + B(t) \tag{2.22}$$

where θ is a global zeroth order phasing term, and for metabolite k, a_k denotes its signal amplitude, $s_k(t)$ denotes its basis function, and ω_k is its offset from the expected frequency. The most efficient way of fitting this large number of parameters is by implementing the variable projection (VARPRO, Golub and Pereyra (1973)) method, in which the linear terms in the signal model (i.e. the a_k terms) can be neglected during a non-linear least squares fit, reducing the number of free parameters to be estimated at each iteration.

2.5.3 Baseline

The baseline term in Equation 2.22 corresponds to unknown broad signals which affect the measured signal, perhaps arising from residual water and macromolecules. In the absence of an accurate signal model for these components, the baseline is usually estimated non-parametrically.

Both LCModel and AQSES employ a flexible spline model to account for the broad background signals, constraining this with a regularisation term to avoid overfitting of non-background signals. Although the AQSES signal model is in the time domain, it exploits the smoothness in the frequency domain during the fitting process. Various investigators have shown that the degree of imposed smoothness has a non-trivial impact on the accuracy of estimates for the parametric components of the model (Emir et al., 2011, Zhang and Shen, 2014).

Both TARQUIN and QUEST exploit the rapid decay of the baseline signals in the time domain, either by truncating the first n points of the FID in a single stage fit (TARQUIN) or by perform the fitting in two stages: with and without the rapidly decaying components over the initial n points (QUEST). In both cases, the number of points n has a non-trivial impact on the accuracy of estimates for the parametric components of the model.

2.5.4 Uncertainty

The level of uncertainty in metabolite signal intensities is typically expressed in terms of the Cramér-Rao lower bounds (CRLBs, Cramér (1946), Rao (1992)). These are independent of the estimator, and define the theoretical lower limits of the variance for the parameter estimates. Any estimation strategy that has a variance approaching the CRLB is considered optimal. The CRLB can be determined as follows (Ratiney et al., 2004):

$$CRLB = \sqrt{diag(F^{-1})} \tag{2.23a}$$

where the Fisher information matrix F is given by:

$$F = \frac{1}{{\sigma_{noise}}^2} \cdot \mathrm{real}(\mathbb{J}^H \mathbb{J}) \tag{2.23b}$$

where σ_{noise} denotes the standard deviation of the signal noise, $\mathbb{J}$ denotes the Jacobian matrix evaluated at the solution, and H denotes the Hermitian conjugate.

The CRLB is only strictly valid in the situation where the model is correct, and there is no fitting bias. This is almost never the case, and so must be interpreted with caution. Nevertheless, it provides some useful information about the uncertainty surrounding the estimated parameters.

It is important to note that using CRLBs as quality control criteria may bias the results of a study assessing pathological decreases in metabolite concentrations (Kreis, 2016), although they are frequently used for this purpose in the literature.

2.5.5 Concentration determination

A number of correction factors must be applied to the estimated metabolite signals before they can be accurately interpreted as concentrations. Firstly, they must be referenced to another signal with known concentration, typically the water signal arising from the same voxel as the metabolite measurement.

Using metabolite and reference signals arising from the same voxel eliminates the need to correct for coil sensitivity effects and voxel volume, and both measurements are performed with the same amount of coil loading. As a result, the concentration can be estimated as follows (Bottomley, 2016):

$$[met] = \frac{n_w S_{met} R_{T1met} R_{T2met}}{n_{met} S_w R_{T1w} R_{T2w}} \frac{[w]}{\rho} \quad \text{mol/kg wet weight} \tag{2.24}$$

where the subscripts w and met refer to tissue water and metabolites, n denotes the number of magnetically equivalent spins contributing to the peak of interest, S denotes the estimated signal intensity, R_{T1} and R_{T2} refer to the respective T_1 and T_2 relaxation correction terms, ρ denotes the specific gravity of the tissue examined, and $[w]$ is the molarity of pure water (=55.5M). All quantities are either known (n, $[w]$), or can be determined experimentally (S, R_{T1}, R_{T2}, ρ). In instances where this is not practical, R_{T1}, R_{T2}, and ρ are typically assumed from the relevant literature.

3 APPLICATIONS OF MR SPECTROSCOPY TO HYPOXIC ISCHAEMIC ENCEPHALOPATHY

This chapter summarises the relevant literature concerning the application of MRS to the neonatal brain, as well as the impact of HIE, and whole-body hypothermia.

3.1 THE HEALTHY DEVELOPING BRAIN

Fluorodeoxyglucose positron emission tomography studies demonstrated in 1986 that the neonatal thalamus, brainstem, cerebellum and sensorimotor cortex have an exceptionally high glucose metabolic rate compared to all other areas of the brain (Chugani and Phelps, 1986). Over the course of the next three months, this pattern spreads to the basal ganglia and wider cortex, before extending to the frontal regions by eight months of postnatal age.

This glucose metabolism supplies approximately two thirds of the required energy in the brain during the neonatal period (Brekke et al., 2015), with most of the rest supplied by ketone bodies from high fat breastmilk (Vannucci et al., 1994). The former increases steadily with postnatal age to reach the much higher metabolic rate observed in the adult brain, where glucose is essentially the sole energy source (Brekke et al., 2015).

Cady et al. (1983) were able to first demonstrate that MR spectroscopy could give an important window into neonatal metabolism, by examining the concentrations of PCr, inorganic phosphate (Pi) and ATP (the latter making up the majority of the nucleotide triphosphate peaks, Lolley et al. (1961)), as well as the pH of tissue *in vivo*. Although ^{31}P MRS is not widely available on clinical MR scanners, several metabolites have also been investigated using ^{1}H MRS over the neonatal period. In order to coherently review the literature on metabolic development as detected by ^{1}H MRS, the following studies relate specifically to regions of highest metabolic activity in the grey matter, which is relevant for the work presented later.

Kreis et al. (1993) examined 77 subjects with a range of ages from neonates (34.5 weeks gestational age) up to 18 years, who had been referred for MRI without any evidence of primary cerebral pathology. Using a two-point STEAM acquisition at 1.5T (TR pairs of 1.5s & 3.0s or 1.5s & 5.0s) and correcting for metabolite relaxation and partial volume effects, they measured metabolite concentrations in the occipital grey matter with an external trimethylsilylethanol reference. With this method, they found a rapid increase in [NAA] at birth and a more gradual decrease in both [Cho] and [Ins] over the timescale of years, accompanied by minimal changes in the relaxation rates of the respective metabolite singlet peaks.

Cady et al. (1996) used a series of PRESS acquisitions and an internal water reference to measure metabolite concentrations at 2.4T, over the narrower range of 32 to 42 weeks gestational age. They found [tNAA] and [tCr] stable in the thalamus across range of infants with no cerebral pathology (n=19), while [Cho] and [Lac] decreased with age. They also calculated that thalamic tNAA T_2 and Cho T_2 increased rapidly with age, at a rate of 17.3 ms/week and 17.8ms/week respectively.

Kreis et al. (2002) used PRESS acquisitions to investigate metabolite concentrations in infants free from cerebral pathology, over a similar range of 32-43 weeks (n=36) at 1.5T. Although not applying per-subject corrections for incomplete metabolite T_1 relaxation, they used an internal water reference, and corrected for partial volume effects, to estimate concentrations in both the thalamus and occipital grey matter. They observed increases in both [tNAA] and [tCr] at a rate of approximately 1mmol/kg wet weight per month, and decreases in parenchyma water T_2 at a rate of 4.1ms per week (comparing preterm infants at 34 weeks gestation to term infants at 41 weeks gestation). They also noted similar rates of [NAA] increase between the thalamus and regions of occipital grey matter, with the former having consistently higher [NAA].

More recent studies at 1.5T by Blüml et al. (2013) and at 3T by Tomiyasu et al. (2013) have used an internal water reference to estimate metabolite concentrations, but neglect significant relaxation effects of either metabolite or water T_1, or partial volume effects, and so suffer from greater inaccuracies than some of those studies discussed above. However, both agree that [NAA] is increasing in grey matter regions over the neonatal period, with the former study suggesting rates of approximately 0.3 mmol/kg per week in the occipital grey matter and the latter study suggesting a change of approximately 0.2 mM per week in the basal ganglia.

The most consistent finding across these five studies is rapidly increasing [tNAA] in the initial neonatal period, although this was not observed by Cady et al. (1996). Rather, they observed a rapidly decreasing NAA relaxation rate, which could appear as a decreased [tNAA] in the presence of uncorrected T_1 saturation effects. This change, if not caused by fitting bias from overlapping Glx resonances, would imply that the bulk cytosolic viscosity decreases with gestation, reflecting a maturational change in either cell populations or cytosolic water content.

The observations of Kreis et al. (2002) concerning parenchyma water T_2 also reflect important maturational changes in the neonatal brain, as the regions which are actively myelinating at birth (i.e. the posterior limb of the internal capsule) have an increased relaxation rate. Although this might be expected to be solely due to an increased proportion of water trapped in myelin (with a much shorter T_2), a more recent study by Melbourne et al. (2016) has suggested that microstructural organisation plays an additional role. Using microstructural models of MR diffusion and multicomponent T_2

relaxometry imaging, they suggest there is a concomitant increase in the number or calibre of neurites (i.e. axons or dendrites) with gestation. This would also give a plausible explanation for the proposed gestational changes in [tNAA].

3.2 HYPOXIC ISCHAEMIC ENCEPHALOPATHY

Hypoxic ischaemic encephalopathy (HIE) results from the combination of a lack of sufficient cerebral blood flow (ischaemia) and oxygen (hypoxia) at the time of birth. This typically occurs at the end of an otherwise normal pregnancy, and carries a high risk of death or lifelong disability (de Vries and Jongmans, 2010). The neonatal brain undergoes dramatic metabolic changes in response to hypoxic ischaemic insult, with a distinctive time course.

3.2.1 Primary energy failure

Under aerobic conditions, acetyl-CoA undergoes oxidation in the mitochondria via the tricarboxylic acid cycle, and adenosine triphosphate (ATP) is generated through oxidative phosphorylation. When oxygen supply is impaired and oxidative phosphorylation declines, PCr can be used to replenish the depleting ATP by conversion of adenosine diphosphate (ADP), along with the production of Cr and inorganic phosphate (Pi). These acute events outline a primary energy failure in HIE. Upon termination of the insult, Cr can be converted back to PCr via the activity of creatine kinase, and the metabolism rapidly reverts to its baseline state (Volpe, 2008).

Since ^{31}P MRS can be used to monitor PCr, Pi and ATP, it has been employed by various investigators to evaluate the severity of this energy failure. Delpy et al. (1982) first demonstrated that rabbits subjected to carotid artery occlusion had a rapid decrease in PCr, and a more gradual reduction in ATP, while Pi increased. Upon reperfusion, these returned back to baseline levels in most cases, apart from where the injury was particularly profound. These findings were first confirmed in human neonates by Cady et al. (1983). Similar metabolic changes were then observed in lambs by Hope et al. (1984), before Lorek et al. (1994) replicated these in a porcine model, which is more closely representative of the neonatal brain.

3.2.2 Secondary energy failure

The discovery of a delayed 'secondary' energy failure was outlined by Hope et al. (1984), observing that in HIE, neonates would have a normal ratio of PCr/Pi one day after birth, but this would become abnormally low over the coming days. Azzopardi et al. (1989) were able to more clearly associate the delayed nadir in PCr/Pi with the neurodevelopmental outcome one year later, such that a lower minimum ratio predicted more severe motor impairments or death. Lorek et al. (1994), upon reproducing this in their porcine model, were able to demonstrate that secondary energy failure

followed a predictable time-course, and that the minimum PCr/Pi observed in this delayed phase was well predicted by the duration of the acute insult. More recently, Bainbridge et al. (2014) have demonstrated a concomitant biphasic pattern in the oxidation state of cytochrome-c-oxidase (the final protein in the mitochondrial electron transport chain), suggesting secondary energy failure is the consequence of mitochondria entering a persistent dysfunctional state.

Along with the changes observed in ^{31}P MRS, secondary energy failure is associated with important observations in ^{1}H spectra. Groenendaal et al. (1994) found that infants with severe HIE had a prominent lactate peak at one week after birth, which was predictive of a poorer neurodevelopmental outcome, with Cady et al. (1994a) promoting tLac/tNAA as a reliable quantitative measure of this. In their porcine model, Cady et al. (1994b) quantified a persistent increase in the T_2s of tNAA, tCr and Lac, attributing this to impairment of the Na^+/K^+ pump and a subsequent influx of cytosolic water in the neurons. Such pathology has also been associated with the leakage of NAA into the extracellular space (Taylor et al., 1994). In the same porcine model, Penrice et al. (1997) then demonstrated that changes in the ratios of tNAA, tCr and Lac peak areas followed a reproducible time-course in parallel with that observed in ^{31}P MRS. Furthermore, they identified that tLac/tNAA and tLac/tCr were significantly elevated in the period between primary and secondary energy failure, where ^{31}P spectra appeared normal.

3.2.3 Therapeutic hypothermia

In the small window of time after primary energy failure, during which metabolism appears largely normal, investigators turned their attention to novel interventions which might be able to avoid secondary energy failure altogether. After a number of false dawns with putative neuroprotectants, it was Thoresen et al. (1995) who first demonstrated a promising neural rescue strategy with mild hypothermia. In their porcine HIE model they observed the striking recovery of NTP after primary energy failure, recovering to the levels observed in control piglets who underwent no hypoxic-ischaemic insult. Thoresen et al. (1997) also demonstrated that hypothermia specifically prevented the release of excitotoxins, while Edwards et al. (1995) showed its potential in inhibiting apoptosis. After the completion of 11 randomised controlled trials (n=1505), there is evidence that therapeutic hypothermia improves survival without increasing major disability in infants with moderate or severe encephalopathy, if initiated within the first six hours after birth (Jacobs et al., 2013). It has also been adopted as the standard of care for these babies across the NHS (National Institute for Clinical Excellence, 2010).

3.3 PROGNOSTICATION

Even in the initial studies by Cady et al. (1983) there was a strong interest in applying MR spectroscopy as a way of assessing neonatal brain injury, and in the case of HIE, its use as a prognostic tool.

3.3.1 ^{31}P and ^{1}H MRS before the introduction of hypothermia

Azzopardi et al. (1989) demonstrated that the ratios of both PCr/Pi and NTP/EPP (EPP= exchangeable phosphate pool) had high positive predictive values of 93% and 91% respectively, but poor sensitivities of 74% and 47% respectively, for identifying infants who died or had very severe disability one year later. Although this study only evaluated global cerebral metabolism, Moorcraft et al. (1991) found no added benefit in performing spatially localised spectroscopy in deeper regions.

Various studies have assessed the prognostic accuracy of ^{1}H metabolite peak area ratios in HIE, many of which are reviewed by Thayyil et al. (2010) alongside other MR biomarkers, covering various small single-centre studies in the period of time where hypothermia was not yet widely introduced into standard clinical practice. They found tLac/tNAA in the deep grey nuclei had a pooled sensitivity of 82% and a pooled specificity of 95% (across 10 studies) in predicting an abnormal neurodevelopmental outcome at least one year after birth. This outperformed the expert scoring of injury from T_1- and T_2-weighted MRI, as well as measurements of the apparent diffusion coefficient from diffusion MRI. tNAA/Cho (six studies) and tNAA/tCr (nine studies) were less predictive, with respective pooled sensitivities of 72% and 61%, and pooled specificities of 59% and 71%.

Fewer studies have measured absolute metabolite concentrations or relaxation rates in HIE, likely due to the additional acquisition time required. Cady et al. (1996) first demonstrated that thalamic [tNAA], [tCho] and [tCr] were reduced in infants with the most severe outcomes, while [tLac] was increased. In addition, he observed that all of these metabolites had increased T_2s in infants with the worst outcomes. Cheong et al. (2006) performed a comprehensive experiment in 10 control neonates and 17 neonates with HIE (10 with normal/mild outcome at one year, 7 with severe/fatal outcome at one year), evaluating the pathological effects on metabolite peak area ratios, concentrations, concentration ratios, T_1s and T_2s. They found [tNAA] to be the only measure which could distinguish all three groups (control, normal/mild, and severe/fatal), and that none of [tCr], [tCho] or [tLac] could discriminate between HIE neonates with normal/mild outcomes or severe/fatal outcomes. Of the relaxation rates, only the T_2 of tLac discriminated between the two HIE groups, although the T_2 of all metabolites increased with poorer prognosis. Of the metabolite peak area ratios, tLac/tNAA, tLac/tCho, tNAA/tCr, and tNAA/tCho could all discriminate between the two HIE groups. Interestingly, the ratios of metabolite concentrations were less predictive than their straightforward peak area ratios, suggesting

that the latter maximises the effect of pathological changes in both concentrations and T_2s. However, this same effect renders peak area ratios more difficult to interpret.

3.3.2 ^{1}H MRS after hypothermic neuroprotection

Following the widespread introduction of hypothermia as the standard of care for HIE, Chandrasekaran et al. (2011) first evaluated the prognostic accuracy of tLac/tNAA in a cohort of 11 cooled and 17 uncooled infants from a single institution, suggesting that this was unaltered by the intervention. More recent single-centre studies agree that tLac/tNAA is still an effective prognostic tool after hypothermia (e.g. Badii et al. (2014), De Vis et al. (2015)). However, as yet there have been no prospective multi-centre studies to evaluate the real potential clinical impact of tLac/tNAA on a wider scale.

Wisnowski et al. (2016) evaluated the metabolic effects of hypothermia by performing scans in HIE infants both during and after the invention. Although they report estimated concentrations, the acquisition scheme used (PRESS, TR/TE=2000/35ms) results in substantial bias due to partial saturation effects, which is particularly problematic due to the well-characterised effect of temperature on T_1 (Dickinson et al., 1986), and on chemical shifts which biases results even at 7T (Wermter et al., 2017). Furthermore, highly covarying resonances (e.g. Cr and PCr, NAA and NAAG) were analysed independently. It is therefore unsurprising that they observe, for example, significant reductions in NAA by the same amount as the significant increases in NAAG during hypothermia.

3.3.3 Other MR biomarkers in HIE

Apart from ^{1}H MRS, a number of authors have investigated the prognostic accuracy of other MR biomarkers. The most straightforward of these is the interpretation of T_1- and T_2-weighted structural imaging. Rutherford et al. (2010), Cheong et al. (2012) and Shankaran et al. (2015) have all examined these measures as sub-studies of the three major randomised controlled trials of hypothermia, and in each case the images were interpreted and reported by a single expert. Despite this, there was a range of prognostic accuracies reported (sensitivities and specificities between 60-90% and 65-92% respectively), and since this is reliant on subjective image scoring which is only semi-quantitative, it is of limited use in comparison to more objective quantitative biomarkers in multi-centre settings.

Perhaps the second most widely studied MR biomarker is diffusion weighted imaging, which follows a distinct time course over the days immediately following HIE (McKinstry et al., 2002). Initially, the apparent diffusion coefficient decreases by approximately one third over the course of the first few days, before increasing to normal ranges one week after birth. This process is termed pseudonormalisation, and highlights the importance of timing an MR scan to be sensitive to pathology. This is particularly important in the context of clinical trials, as therapeutic hypothermia prolongs this process such that pseudonormalisation occurs over a longer period (Bednarek et al., 2012). As such,

diffusion effects are more difficult to assess in pragmatic multi-centre settings, where different sites may scan at different points in the time course of pathological changes.

An important advantage of diffusion weighted imaging is the ability to generate quantitative maps, which in theory should be comparable across institutions following calibration. However, the lack of a ground truth diffusion phantom which reflects complex cerebral microstructure means that inter-site effects are difficult to assess. Perhaps the most effective method of correcting for inter-centre effects is by including post-hoc covariates, but this is most effective where there are a large number of subjects at each site (Fortin et al., 2017). An alternative approach used by Azzopardi et al. (2016) was to avoid calibration and only include sites where hardware and software specifications were closely matched, to minimise the likelihood of having significant inter-site effects. However, this places restrictions on the possible number of recruited subjects. A direct comparison of conventional MRI, diffusion-weighted MRI and ^{1}H MRS in HIE by Ancora et al. (2013) at a mean age of 8 days after birth indicated that ^{1}H MRS was most effective at predicting clinical outcomes two years later. This agrees with previous meta-analysis by Thayyil et al. (2010), which highlights ^{1}H MRS as the most accurate prognostic MR biomarker.

De Vis et al. (2015) have shown in a retrospective cohort of infants (n=20) that those with abnormal outcomes also have higher perfusion in the basal ganglia and thalamus from arterial spin labelling (ASL) MRI, compared to those with normal outcomes. While this is a promising window into pathology, ASL is not widely available on clinical scanners, and requires extensive post-processing in the case of motion, and several key assumptions about tissue properties (blood and tissue T_1, labelling efficiency) to generate quantitative maps (Sokolska et al., 2015). As more data becomes available on the neonatal brain, and as such techniques become more routinely incorporated into clinical protocols, this may begin to play an important role in prognostication. For the moment, De Vis et al. (2015) indicate that it only provides a small incremental benefit in prognostic accuracy if used in conjunction with simple ^{1}H MRS measures in the basal ganglia.

3.3.4 Amplitude-integrated electroencephalography (aEEG)

Although aEEG is a promising biomarker as it can be easily performed continuously at the bedside, it relies on clinical interpretation and scoring, and as such is only semi-quantitative. Nevertheless, Ancora et al. (2010) directly compared aEEG at 24 hours after birth with ^{1}H MRS at 6-9 days later in a cohort of HIE neonates (n=32), and found it to be a useful prognostic measure with comparable prognostic accuracy. A more recent meta-analysis by Chandrasekaran et al. (2017), examining infants undergoing hypothermia (n=520), demonstrates that aEEG is poorly prognostic at 6 hours (pooled sensitivity 95%, pooled specificity 39%), but reaches maximum prognostic accuracy at 48 hours (pooled sensitivity 85%, pooled specificity 93%). Despite this promise, it is of limited value in multi-centre trials where more

quantitative biomarkers enable the investigation of subtle effects from treatment, or from different pathological phenotypes.

3.3.5 ^{1}H MRS in multicentre trials of HIE

Although ^{1}H MRS biomarkers show clear prognostic potential in HIE, their application has largely been limited to small single-centre studies (Thayyil et al., 2010). Since the incidence of HIE is low, adequately powered trials of prognostic accuracy take prohibitively long, and pooling data from different sites via meta-analysis has limited accuracy due to the important methodological differences between acquisition protocols and quantification.

This is not a problem restricted to the study of HIE. A recent review by the international MRS Consensus Group highlights that the lack of co-ordinated multi-centre trials with harmonised protocols is a major hurdle to bringing MRS into clinical practice in various pathologies (Öz et al., 2014).

The work presented in the following section outlines the largest ever prospective trial examining MR biomarkers in HIE, in which harmonised MRS acquisitions were conducted across multiple institutions in the UK and the USA.

4 THE PROGNOSTIC ACCURACY OF MRS IN HYPOXIC ISCHAEMIC ENCEPHALOPATHY

This chapter outlines a multi-centre prospective observational study in which the prognostic accuracy of MR spectroscopy is assessed in HIE. I am preparing this work for publication as the first author, and will give an oral presentation on the results at the Joint Annual Meeting of the ISMRM-ESMRMB 2018.

4.1 INTRODUCTION

Cerebral magnetic resonance (MR) biomarkers, obtained within days or weeks of the initial injury, are more objectively quantifiable than long term clinical outcome measures, and give insight into different facets of brain injury. T_1- and T_2-weighted MRI provides information about macroscopic tissue properties, MR spectroscopy (MRS) offers insight into metabolism, and diffusion MRI (dMRI) is sensitive to changes in tissue microstructure. Meta-analyses have demonstrated that these measures provide accurate prognoses in various single centre studies (Thayyil et al., 2010), but it is challenging to achieve similar results across a network of institutions with different MR scanner hardware and software implementations.

In this study, the aim was to harmonise the acquisition of MR biomarkers across several tertiary neonatal care centres in the UK and USA, and assess their prognostic accuracy in predicting adverse neurodevelopment in term or near-term infants two years after neonatal encephalopathy.

4.2 MRS PROTOCOL

A protocol was designed that could incorporate the measurement of both metabolite peak area ratios at a long echo time, and the absolute concentration of tNAA. To achieve this, the protocol consisted of several separate PRESS acquisitions in the same 15x15x15mm^3 thalamic voxel, during which care was taken that the voxel, shim, and RF transmit and receive gains were kept constant (Table 4.1). The rationale behind this protocol is explained in the following sub-sections, and outlines of the postprocessing and calculation steps shown in Figure 4.1 and Figure 4.2.

4.2.1 Step 1: Metabolite peak area ratios (long TE)

The first acquisition provided sufficient data for an assessment of the ratio of T_1- and T_2-weighted metabolite peak areas, at an echo time at which tLac undergoes minimal signal loss due to the anomalous J modulation effects outlined in Section 2.4.4. The division of the acquisition into 16 sub-spectra ensured that post-hoc motion correction strategies could be employed (i.e. frequency and phase correction), and any sub-spectra demonstrating excessive sporadic motion could be discarded altogether, prior to coherent averaging.

Step	TE (ms)	TR (ms)	Water suppression	Phase cycling	Sub-spectra	Dummy scans	Centre frequency (ppm)
1	288	2288	On	8	16	2	2.01
2	60	2060	On	8	8	2	2.01
3	60	5000	On	8	8	2	2.01
4	60	10060	Off	8	1	1	4.67
5	124	10124	Off	8	1	1	4.67
6	205	10205	Off	8	1	1	4.67
7	316	10316	Off	8	1	1	4.67
8	495	10495	Off	8	1	1	4.67
9	1000	11000	Off	8	1	1	4.67

Table 4.1: Key acquisition parameters for the full MRS protocol for the MARBLE study.

4.2.2 Steps 1-3: Metabolite T_1 and T_2 Relaxometry

The first and second acquisitions together enable a two-point T_2 relaxometry experiment, as the longitudinal magnetisation recovery time for the PRESS sequence (TD = TR – TE) is held constant, while the echo time is varied.

Similarly, the echo time is held constant between acquisitions 2 and 3, while the longitudinal magnetisation recovery time is increased from 2000ms to 4940ms. This allows a separate two-point T_1 relaxometry experiment.

4.2.3 Steps 4-9: Absolute quantification

With acquisitions 4 to 9, the centre frequency of the RF pulses is shifted to account for chemical shift displacement effects, and a six-point T_2 relaxometry experiment is performed on the unsuppressed water signal in the same voxel location, to correct for partial volume effects (Ernst et al., 1993). Although the long repetition time should ensure that the parenchyma water signals have reached equilibrium magnetisation, the potential presence of saturated free water (with a T_1 and T_2 on the order of seconds) requires that the longitudinal magnetisation recovery time be held constant.

The water T_2 relaxometry experiment allows measurement of parenchymal water signal at equilibrium magnetisation, while the combination of T_1 and T_2 relaxometry experiments provides an equivalent measurement of parenchymal metabolite signals at equilibrium magnetisation. This, in turn, allows calculation of the absolute concentration of metabolites within the voxel.

It is important to note that since the water-suppressed acquisitions are performed with the RF centred on 2.01ppm, only the concentrations derived from the NAA methyl singlet resonance are fully corrected for chemical shift displacement errors.

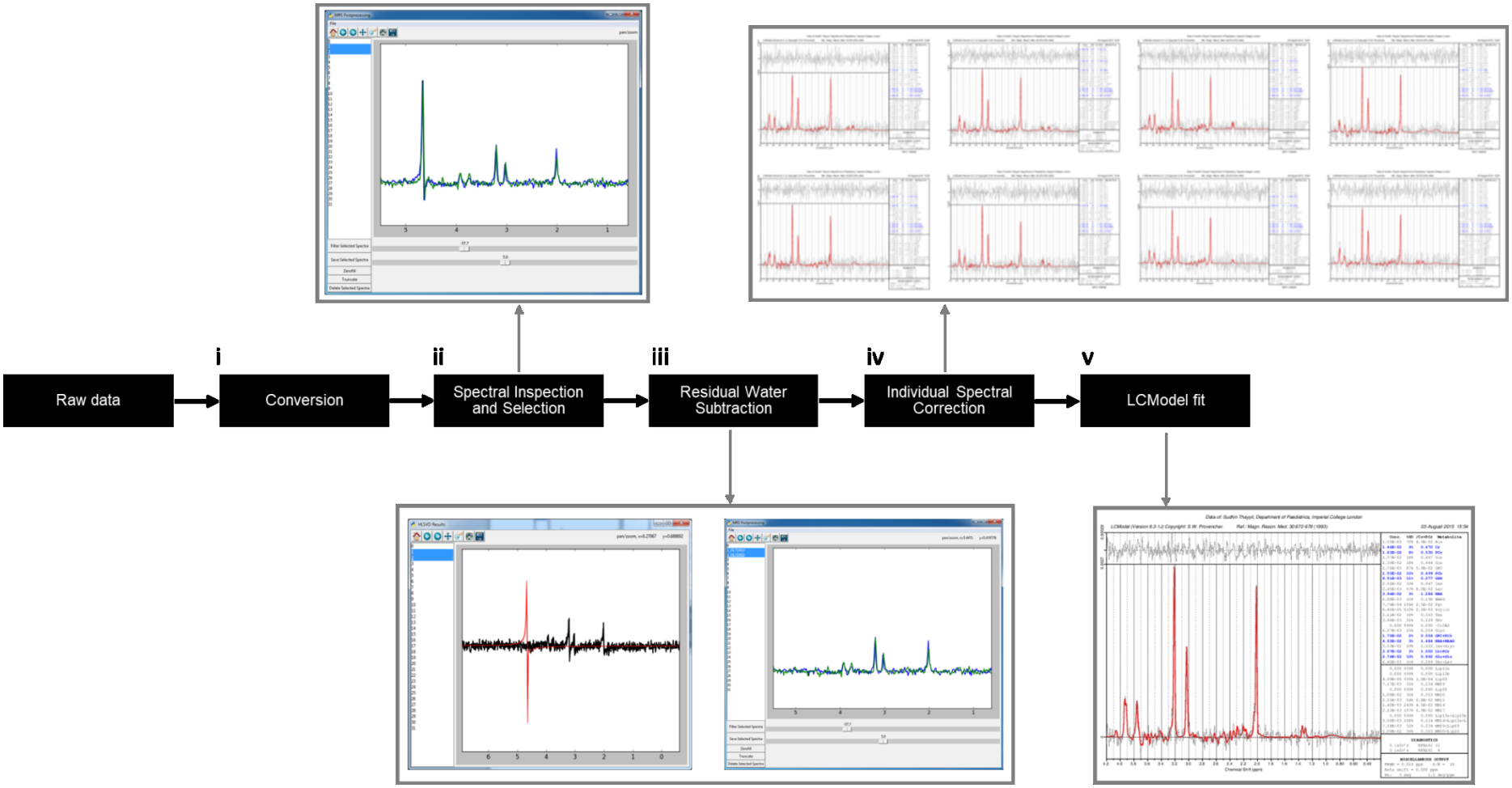

Figure 4.1: Schematic of the post-processing strategy used. I developed a GUI in Python to allow (i) raw data conversion from Philips (DICOM), Siemens (IMA) or GE (P-files) scanners, (ii) manual selection of non-corrupted sub-spectra, and (iii) residual water subtraction with HLSVD. The sub-spectra are then passed to LCModel (iv) for individual fitting (eddy current correction followed by zeroth order phasing and frequency alignment only), and the newly aligned sub-spectra are averaged and entered into a final LCModel fit (v).

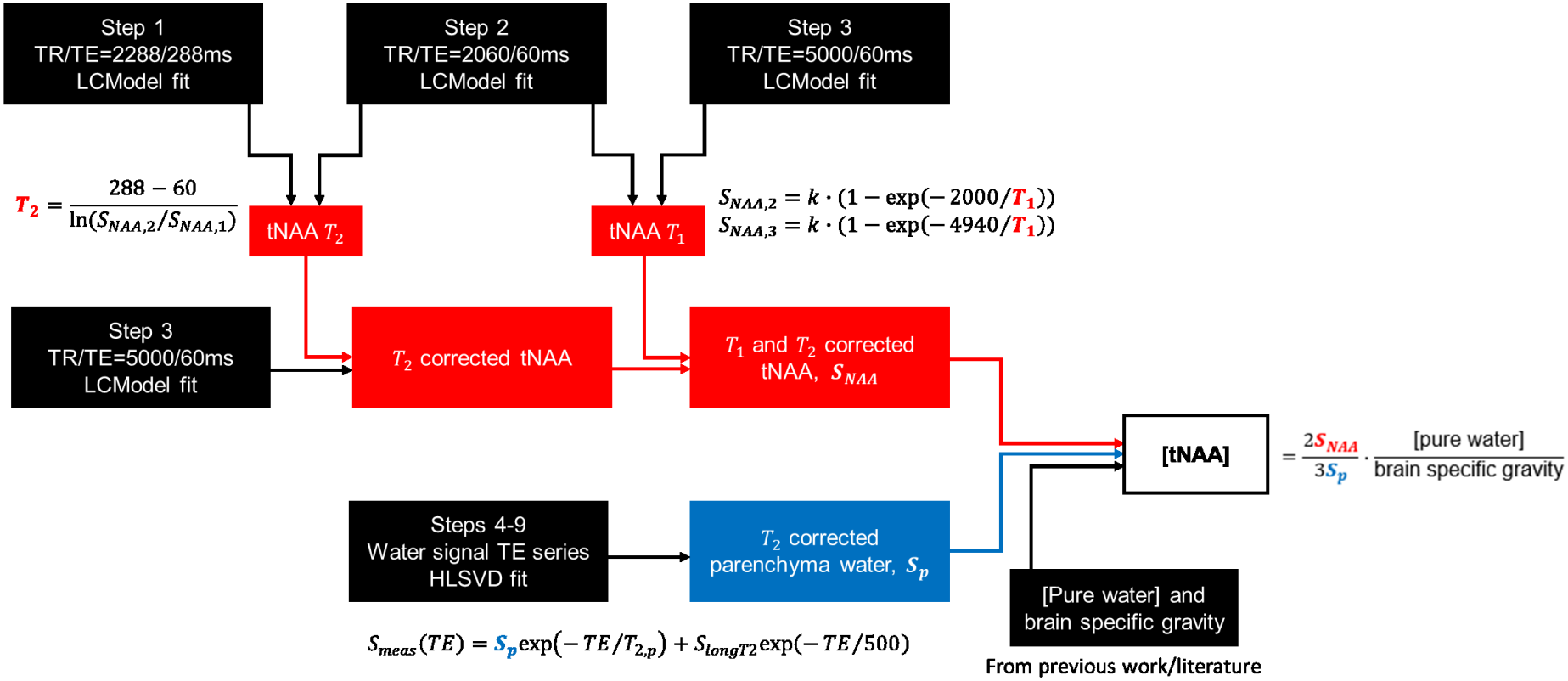

Figure 4.2: Calculation steps to derive thalamic tNAA concentration, [tNAA], from the data collected in the various steps of the protocol. Black boxes show the input data, either obtained from the LCModel/HLSVD fits or reference values from previous work/literature. Red boxes and text denote calculated quantities from the tNAA resonance, and blue boxes and text denote calculated quantities from the water signal. Equations accompany the different calculation steps. $S_{NAA,step}$ denotes the tNAA signal intensity in the corresponding protocol step, with the relaxation corrected intensities for tNAA and parenchyma water given by S_{NAA} and S_p respectively.

4.3 PHANTOM VALIDATION

4.3.1 Methods

One of three spherical phantoms, each containing a buffered solution of 10.0mM NAA and 10.0mM Lac, was scanned at room temperature using the outlined protocol at each study site, allowing quantification of inter-site bias and variance. Since there was only a single site with a Siemens scanner, the phantom was scanned at an additional Siemens site not recruiting to the main study. An identical phantom was prepared at a recruiting site in the USA, due to the difficulty in transportation.

At GE sites, the phase cycling scheme had to be adjusted for the water relaxometry experiment in steps 4-9, as the proposed combination was not permitted by the software. Instead, non-water suppressed signal intensities were measured from the sum of two sub-spectra of 4 phase cycles each.

Water suppressed spectra were analysed using LCModel (v6.3-1J), after applying an eddy current correction based on the method of Klose (1990), zero-filling by a factor of two, and applying 5Hz Lorentzian apodisation. Using VeSPA-Simulate (version 0.9.3, Soher et al. (2011)) metabolite basis sets were simulated with ideal RF pulses for the NAA(-CH_3) singlet and the Lac(-CH_3) and Thr(-CH_3) doublets according to the TE_1,TE_2 combination employed by each vendor in their product PRESS sequence (personal communication with Philips, Siemens and GE clinical science teams).

4.3.2 Results

The results of the phantom measurements are shown at each site in Figure 4.3, for both the measurement of tLac/NAA and [NAA] (ground truth = 10.0mM), grouped by scanner vendor. One site (MWY) had insufficient research privileges to implement the full protocol on their scanner.

A one-way ANOVA test identified a significant inter-site effect on [NAA] measurement (p=0.01), but not on tLac/NAA measurement (p=0.49). Tukey's HSD testing demonstrated significant differences in [NAA] between GE sites and Philips sites (p=0.01) and between GE sites and Siemens sites (p=0.02), but not between Philips and Siemens sites (p=0.93).

Across all sites, the coefficient of variation for tLac/NAA was 4%. The mean [NAA] recorded at the GE sites was 8.7 (SD 0.2) mM, while a mean (SD) of 10.1 (0.4) mM was measured across Philips and Siemens sites.

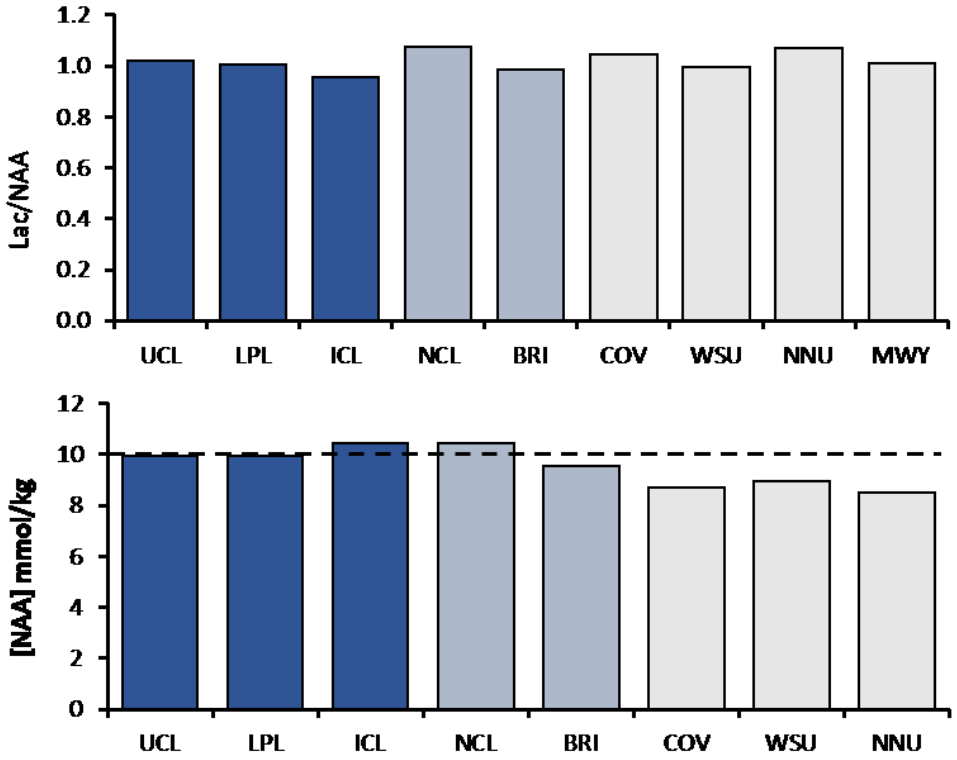

Figure 4.3: Measured Lac/NAA and [NAA] values in the phantom at each site. The dashed line indicates ground truth [NAA]. UCL=University College London (UK), LPL=Alder Hey Children's Hospital (Liverpool, UK), ICL=Imperial College London (UK), BRI=Bristol University (UK), COV=University Hospital Coventry (UK), WSU=Wayne State University (Detroit, MI, USA), NNU= Norfolk and Norwich University Hospital (UK), MWY=Medway Maritime Hospital (Gillingham, UK). Sites are grouped by scanner vendor, from left to right: 3 x Philips (deep blue), 2 x Siemens (light blue), 4 x GE (grey).

4.3.3 Discussion

The tLac/NAA measurements were reproducible across sites (coefficient of variance 4%), despite the variation in sequence implementations (e.g. timings and RF waveforms) and hardware. Due to anomalous J modulation, the T_2 of the tLac resonance could not be accurately measured. However, assuming equal relaxation properties across sites, it appears that losses in tLac signal at TE=288 are similar across vendor-supplied PRESS sequences.

While [NAA] showed a similar coefficient of variance across Philips and Siemens scanners (4%), [NAA] was significantly underestimated when estimated using GE scanners. Inspecting individual sub-spectra, NAA and water signals had reached steady state magnetisation throughout data acquisition (see Figure 4.4), and there were no variations in RF receive/transmit gain or shim adjustments between acquisitions. This, therefore, suggested a consistent bias between NAA and water signal intensities, which was reproducible across GE sites. This could perhaps be explained by the incomplete suppression of water signals outside of the voxel by using 4 step vs 8 step phase cycling, however there were no

spurious echo artifacts in the FIDs which would have suggested the detection of echoes from unintended coherence pathways. Alternatively, a systematic increase in estimated [NAA] may be caused by the use of strong spoiler gradients on Philips and Siemens implementations, as water signals have a higher apparent diffusion coefficient than NAA (Ellegood et al., 2005), and so may be decreased in intensity at the echo. However, the similarity of [NAA] to ground truth values for both Philips and Siemens scanners suggests diffusion losses are negligible, and that in fact the GE implementation underestimates [NAA].

Closer inspection of the fits to the water and metabolite signals failed to reveal systematic errors in estimations (Figure 4.4 and Figure 4.5). In addition, performing wild bootstrapping on the exponential fit to the water signals revealed errors of <1% for the fully relaxed signal on either GE or Philips scanners (Figure 4.6).

The most likely explanation for a systematic difference is the combination of outer volume suppression on GE scanners, and the use of reduced flip angle refocusing pulses (see Appendix II). This could introduce different relative reductions in signal intensity between water and NAA, due to differences in their relaxivities.

Since the measurements of [NAA] had a low coefficient of variation across the three GE sites, a simple linear correction appeared justified. A multiplicative calibration factor of 1.15 was therefore calculated for concentration measurements made on GE scanners.

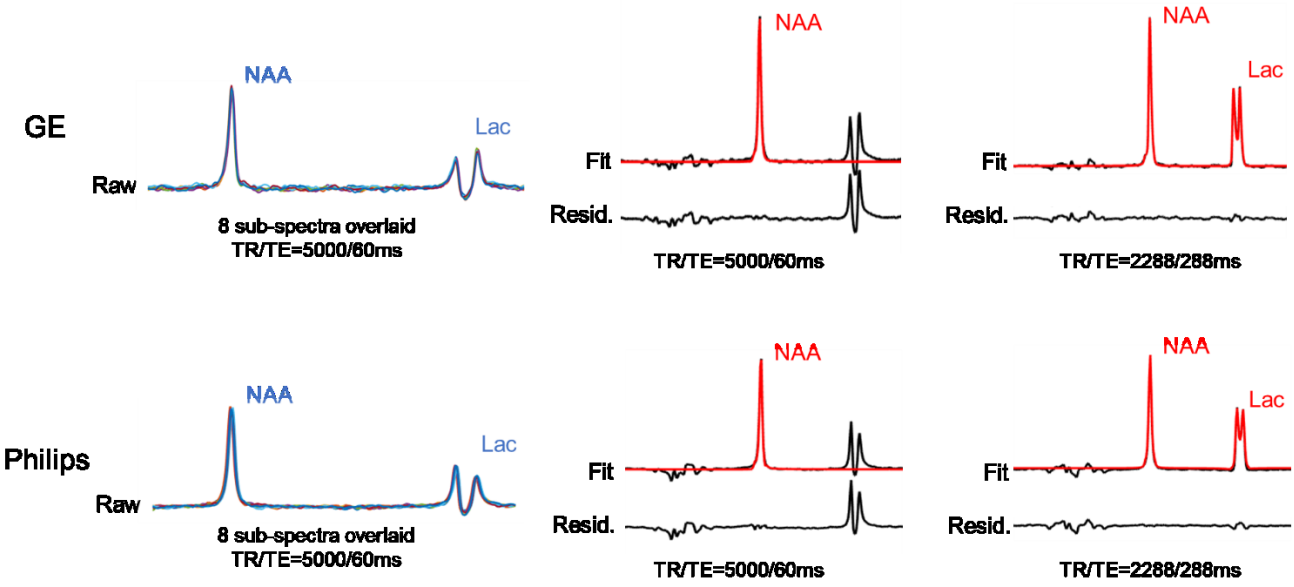

Figure 4.4: Representative phantom spectra and LCModel fits on a GE (top row) and Philips (bottom row) scanner. Left: eight spectra overlaid from the TR/TE=5000/60ms acquisition. Centre: LCModel fit of the NAA signal at TR/TE=5000/60ms. Right: LCModel fit to the NAA and Lac signals at TR/TE=2288/288ms.

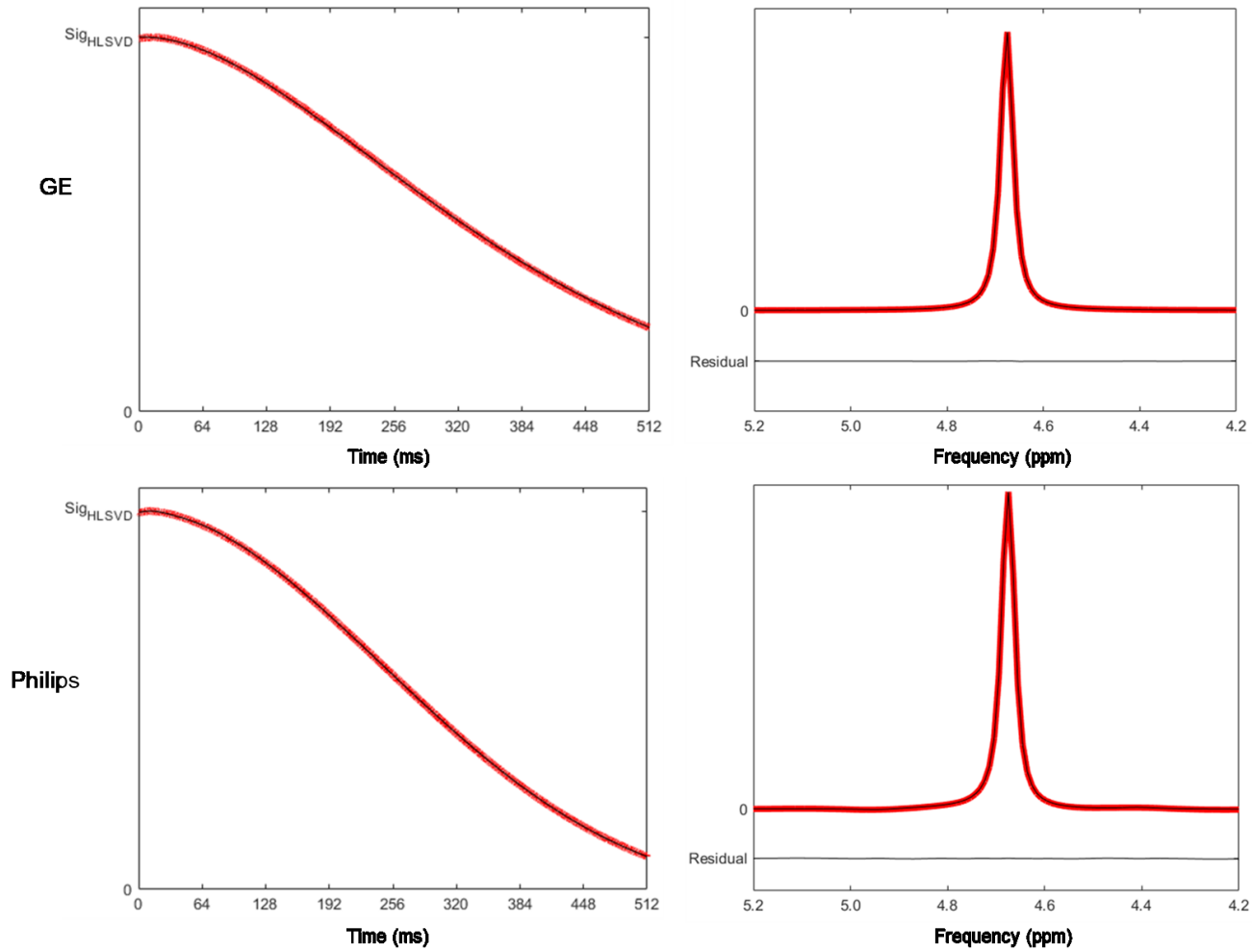

Figure 4.5: Representative phantom water signals (black) and HLSVD fits (red) on a GE (top row) and Philips (bottom row) scanner. Left: Time domain representation. Right: Frequency domain representation, with fitting residual (black).

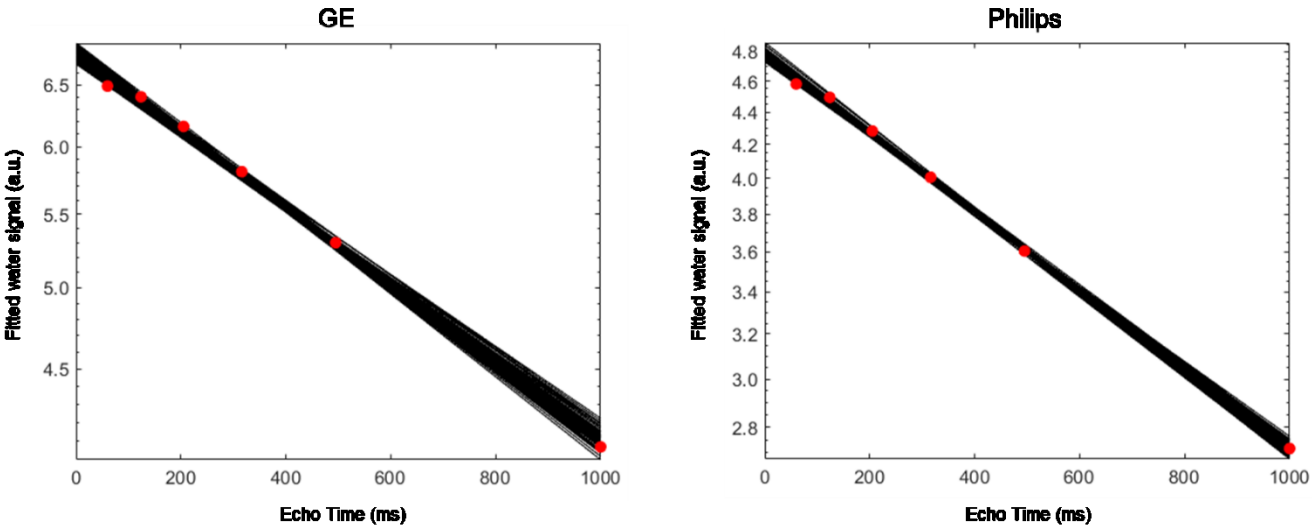

Figure 4.6: Plots of 100 successive wild bootstrap iterations, performing an exponential fit (black lines) to the HLSVD fitted water signals at different TEs (red dots) on each of a GE (left) and Philips (right) scanner.

4.4 ADULT VOLUNTEERS

4.4.1 Outline

To assess the validity of a multiplicative factor on *in-vivo* data derived from GE scanners, healthy adult volunteers (ages 24-39) were each scanned on both a Philips (3T Achieva) and GE (3T Signa HDxt) scanner.

4.4.2 Methods

The same acquisition steps outlined in Section 4.3.1 were performed in five adult volunteers, in a 15x15x15mm^3 voxel centred on the thalamus.

Water suppressed spectra were analysed using LCModel (v6.3-1J), after applying an eddy current correction based on the method of Klose (1990). Using VeSPA-Simulate (version 0.9.3, Soher et al. (2011)) metabolite basis sets were simulated with ideal RF pulses and 1Hz Gaussian broadening, according to the TE_1, TE_2 combination employed by each vendor in their product PRESS sequence. A total of 25 metabolites were included in the analysis: acetate (Act), alanine (Ala), ascorbate (Asc), betaine (Bet), aspartate (Asp), choline (Cho), phosphocholine (PCh), glycerophosphocholine (GPC), creatine (Cr), phosphocreatine (PCr), gamma-aminobutyric acid (GABA), glucose (Glc), glutamate (Glu), glutamine (Gln), glutathione (GSH), glycine (Glyc), lactate (Lac), myo-inositol (Ins), N-acetylaspartate (NAA), N-acetylaspartyl glutamate (NAAG), phosphoethanolamine (PE), propylene glycol (PGC), scyllo-inositol (Scyllo), taurine (Tau), and threonine (Thr). No lipid or macromolecule resonances were included in the basis set. The methyl peaks of NAA, NAAG, Cho, PCr and Cr were separated from other groups in the basis set, and NAA+NAAG were combined (tNAA) for analysis due to their strong covariance.

Results were compared for each adult between scans to determine whether the acquisition on GE scanners underestimated [tNAA] *in-vivo*, by performing a one-sided paired t-test.

4.4.3 Results

In the ten datasets from the five subjects, one was discarded for having a broad linewidth outside of the normal range of the group. As a result, only four subjects had data available for analysis. Due to the low signal-to-noise ratio at the long echo time, T_2 corrections could not be performed on a per-subject basis, and so a single correction factor was applied to the tNAA signal in all subjects (T_2=275ms, reviewing the results of Kirov et al. (2008)).

The thalamic concentrations measured across sites are shown in Figure 4.7, along with the prescribed voxel positions. There was a mean (SD) reduction of 0.4 (1.1) mmol/kg wet weight on the GE scanner (equivalent to 4% of the mean), which was not statistically significant (p=0.25).

4.4.4 Discussion

There was no statistically significant difference detected in the adult subjects on the GE system compared to the Philips system, which was unexpected following the observation of a 15% decrease in the phantom study.

With the loss of one subject, the power of the experiment reduced from 80% to 70% to detect the predicted effect size at 95% significance, increasing the likelihood of a false acceptance of the null hypothesis.

In addition, there are regional variations in [tNAA] *in-vivo* which were not present in the phantom study, and the performance of the shimming system on each scanner is more important due to the heterogeneity of iron deposition in the region of interest. As a result of these effects, there is increased within-subject variance which reduces the ability to detect scanner-specific differences. This may explain why an underlying systematic difference detected in the *in-vitro* study may have gone undetected.

The experiment could have been altered by focusing on a homogeneous region of the adult brain which is easier to shim effectively, and allows an increase in voxel size (e.g. posterior white matter). This would have reduced the linewidth to be closer to the neonatal spectra, and improved the signal-to-noise ratio, therefore increasing the sensitivity to inter-site effects. While less representative of the specific protocol used in the later neonatal study, this would have given some additional insight into systematic differences between scanners.

On the balance of the *in-vitro* and *in-vivo* results, the decision was made to apply the correction factor of 1.15 for GE-derived metabolite concentrations in the neonatal study. However, a sensitivity analysis would be required to determine whether its inclusion had a substantial impact on the results of the main study and their interpretation.

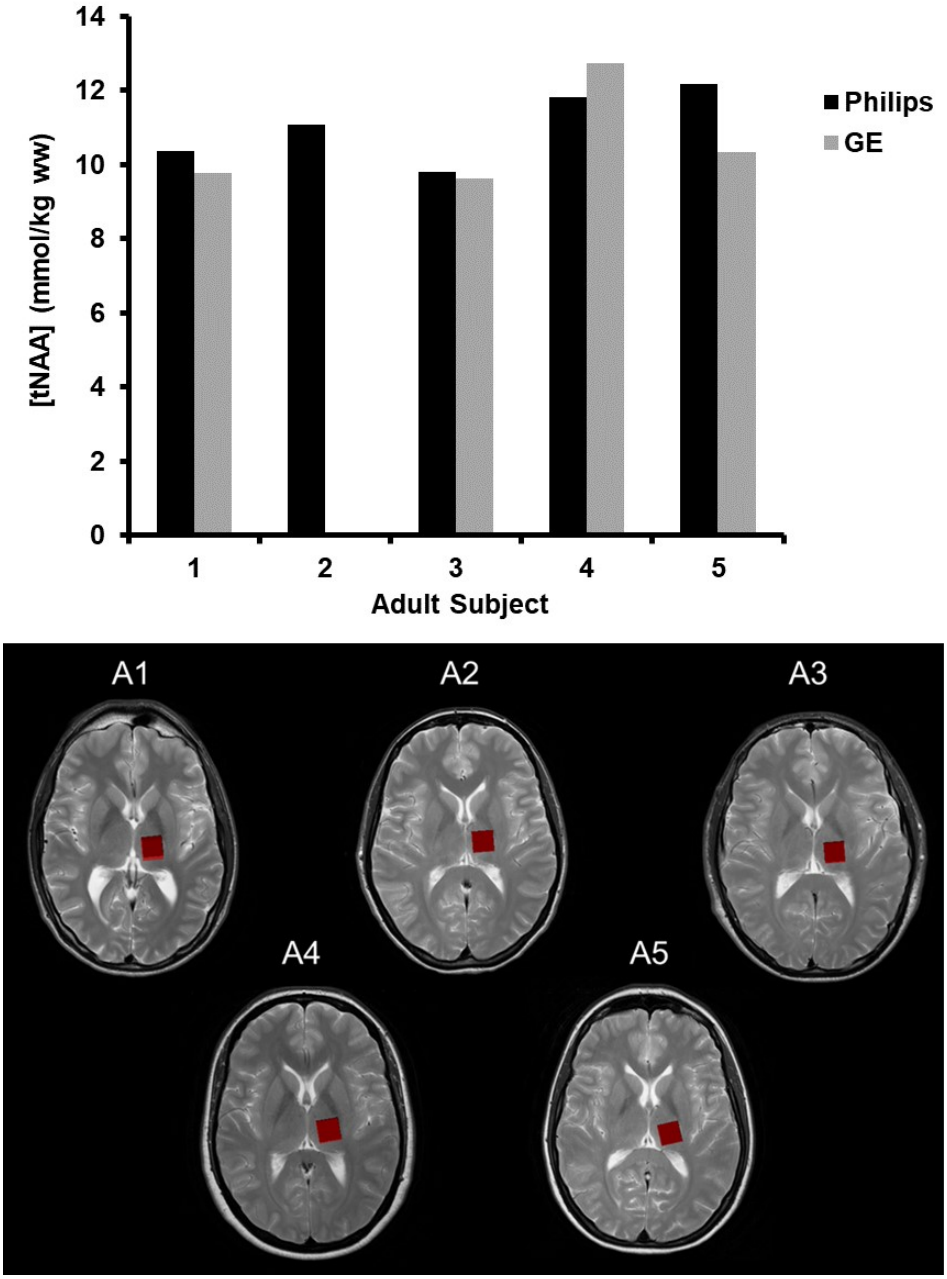

Figure 4.7: Top: thalamic [tNAA] measured in the five healthy adults on each scanner. Bottom: prescribed thalamic voxel positions in the five adults (red), overlaid on an axial T_2-weighted image (Philips).

4.5 CLINICAL STUDY

4.5.1 Study design

Consecutive term or near-term infants (≥36 weeks) with neonatal encephalopathy, and requiring therapeutic hypothermia, were recruited between January 2013 and June 2016, with staggered start dates across several participating sites in the UK and USA. Infants were excluded if they had life-

threatening congenital malformations, syndromic disorders, neurometabolic diseases, or any alternative diagnoses for encephalopathy that were apparent within six hours of birth. In a small number of cases, appropriate MR data had been collected as part of standard clinical care, and so parents were approached for their consent to use these data and participate in later neurodevelopmental follow-up. The study was approved by North London Research Ethics Committee (11/H0717/6) and all local sites, and all parents provided written informed consent.

Upon recruitment at each site, infants were assessed according to the National Institute of Child Health and Human Development (NICHD) neurological examination, a modification of the Sarnat staging system. This yielded an encephalopathy grade (none/mild, moderate, or severe) which reflected the severity of neurological symptoms across six categories.

Neurodevelopmental outcome was assessed using the Bayley Scales of Infant and Toddler Development (3rd Edition, Bayley-III) and a detailed neurological examination by trained and certified examiners, masked to MRS data. Severe disability was defined as any of the following: Bayley-III composite cognitive and language scores <70, Gross Motor Function Classification System (GMFCS) level 3-5, hearing impairment requiring hearing aids, or blindness. Moderate disability was defined as composite cognitive and language scores between 70-84 or any of the following: GMFCS level 2, hearing impairment without the need for amplification, or a persistent seizure disorder. The primary outcome was a composite of death, or in survivors, moderate or severe disability (yes/no). Where infants had severe disabilities that prevented them from undertaking Bayley-III, they were assigned nominal values for the composite scores in each domain (cognitive = 54, language = 46, motor = 46).

4.5.2 MR protocol

For the study, MR scans were performed between 4-14 days after birth on a 3 Tesla scanner (Philips, Siemens or GE). Pseudonymised MR data were transferred to Imperial College London for post-processing and analysis. *In-vivo* quality assurance included the visual assessment of motion from imaging before and after MRS; visual assessment of voxel placement accuracy; assessment of protocol adherence; manual rejection of motion corrupted sub-spectra; automated spectral corrections for both frequency and phase; and the rejection of spectra with linewidths outside the normal distribution of the full dataset (Kreis, 2004). Residual water signals were removed from the raw spectra before applying phase and frequency corrections using LCModel, within an automatic post-processing pathway developed in Python (Figure 4.1). All MR data analyses were conducted while blinded to the clinical outcomes.

All water suppressed spectra were analysed using LCModel (v6.3-1J, Provencher (2001)), with basis sets simulated with ideal RF pulses according to the TE_1,TE_2 combination employed by each vendor using

VeSPA-Simulate (v0.9.3, Soher et al. (2011)). The methyl peaks of NAA, NAAG, Cho, PCr and Cr were separated from other groups in the basis spectra to allow quantification of individual relaxation rates. NAA+NAAG were combined (tNAA), and PCr+Cr were combined (tCr) due to strong covariance. A total of 34 metabolites were included in the analysis: acetate (Act), alanine (Ala), ascorbate (Asc), betaine (Bet), aspartate (Asp), choline (Cho), phosphocholine (PCh), glycerophosphocholine (GPC), creatine (Cr), phosphocreatine (PCr), gamma-aminobutyric acid (GABA), glucose (Glc), glutamate (Glu), glutamine (Gln), glutathione (GSH), glycine (Glyc), lactate (Lac), myo-inositol (Ins), N-acetylaspartate (NAA), N-acetylaspartyl glutamate (NAAG), phosphoethanolamine (PE), propylene glycol (PGC), scyllo-inositol (Scyllo), taurine (Tau), threonine (Thr), and the Lip/MM resonance parameterisation proposed by Provencher (2001).

Metabolite T_1s and T_2s were calculated along with the relaxation corrected signal intensities, as outlined in Figure 4.2. Water unsuppressed signals were quantified using a bandwidth limited five component exponential fit with HLSVD (Pijnappel et al., 1992). The parenchymal water signal (S_p) was quantified by biexponential fit of the water signals measured across the water unsuppressed TE series (via a multi-start lsqcurvefit solver in MATLAB 2017a), with a long T_2 component (fixed at 500ms, similar to Kreis et al. (1993)) accounting for mobile water, according to the following equation:

$$S_{meas}(TE) = S_p \exp\left(-TE/T_{2,p}\right) + S_{longT2} \exp(-TE/500) \quad (4.1)$$

where S_{meas} is the measured water signal at each TE, and $T_{2,p}$ is the parenchymal water T_2 and S_{longT2} is the nuisance signal arising from long T_2 water components such as cerebrospinal fluid, and arterial or venous blood. Due to the immature myelination of the neonatal brain, grey and white matter could not be accurately distinguished through image-based segmentation, and so the parenchymal tissue was assumed to be homogeneous in character.

4.5.3 Statistical analysis

Receiver operating characteristic curves were produced to examine the prognostic accuracy of measures derived from MRS, and to choose optimal cut-off points which maximized sensitivity and specificity. At these points, prognostic indices were obtained with 95% confidence intervals using the exact binomial method.

A hierarchical linear regression model was then used to examine the association between MRS measures and motor composite scores obtained with Bayley-III (as this is least confounded by socioeconomic factors), controlling for gestational age at scan. Due to its skewed distribution, tLac/tNAA was transformed to ln(tLac/tNAA) for analyses.

Finally, a multivariate regression model for predicting two-year motor performance was designed. All available concentrations and relaxation rates for tNAA, tCho, tCr and parenchyma water were added into a lasso (least absolute shrinkage and selection operator) regression analysis (Tibshirani, 1996), and a leave-one-out cross-validation was performed to identify the most parsimonious model which described the variation in motor composite scores (glmnet 2.0-13 Friedman et al. (2010), R v3.4.3 R Core Team (2013)).

4.6 RESULTS

A total of 223 infants were recruited, of whom 190 (85%) had neurodevelopmental outcome assessments for analysis. Of those, 190 (100%) had T_1- and T_2-weighted imaging, 160 (85%) had MRS including thalamic metabolite peak area ratios, and 82 (43%) had MRS including thalamic metabolite concentrations. Of the 190 infants with outcome data, 31 (16%) had an abnormal neurodevelopmental outcome (death or moderate/severe disability).

4.6.1 Clinical characteristics

Across all eligible infants, the mean (SD) birth weight was 3.4 (0.5) kg, at a median (IQR) gestational age of 40 (2) weeks. Median (IQR) Apgar scores at 5 and 10 minutes were 4 (2) and 6 (3) respectively, with 216 undergoing 72 hours of moderate hypothermia, six being rewarmed prior to 72 hours, and one rewarmed at 96 hours due to local clinical decisions. According to assessment against NICHD criteria within the first six hours after birth, encephalopathy was graded severe in 23 (10%) infants, moderate in 163 (73%), and mild in 36 (16%). In one more case, the infant was initially assessed as having no encephalopathy at admission, but this progressed within the first six hours and the decision was made to begin hypothermia. The seizure status on admission was recorded in 209 (94%) babies, of whom 114 (55%) demonstrated seizures.

4.6.2 MR data quality

MR scanning was performed at a median (IQR) age of 7 (5) days. Five of the MRS datasets were lost due to difficulties exporting them from GE scanners (as proprietary P-files), while 24 were rejected due to poor quality: in 12 cases the voxel was positioned incorrectly during acquisition, four datasets demonstrated excessive motion, and eight had a linewidth outside the normal distribution for all subjects (i.e. 95% confidence interval). The scatter plot of signal-to-noise ratio against linewidth (output by LCModel) for different spectra is shown in Figure 4.8. Representative spectra (reflecting the median linewidth across all data) are shown for one subject in Figure 4.9.

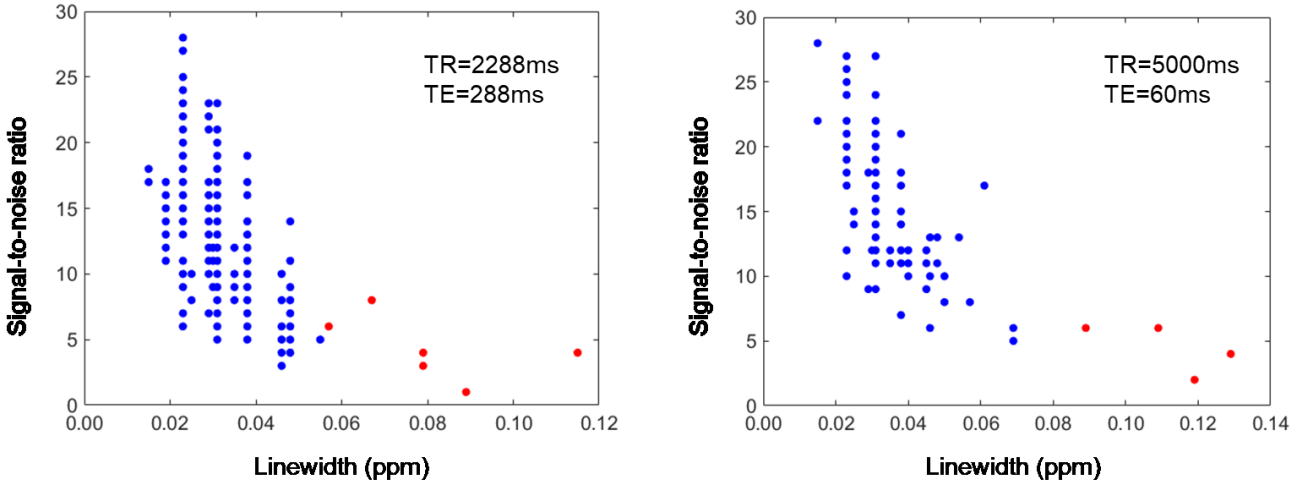

Figure 4.8: Scatter plot of signal-to-noise ratios and linewidths obtained from all analysed spectra with TR/TE=2288/288ms (left) and TR/TE=5000/60ms (right). Blue points indicate datasets included in the final analysis, with red points excluded due to excessive linewidths outside of the normal range for each acquisition.

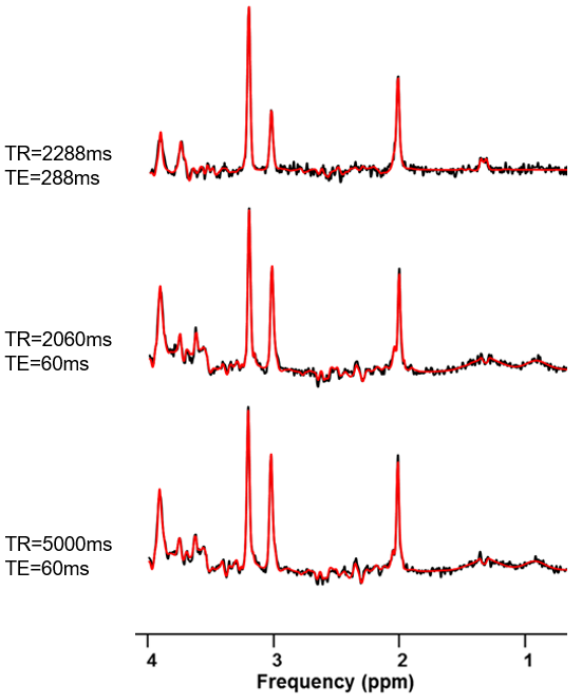

Figure 4.9: Representative PRESS spectra from a single subject (in black), with the LCModel fit overlaid (in red). These spectra have been pre-processed with manual rejection of motion corrupted sub-spectra, followed by automatic frequency and zeroth order phase alignment, prior to averaging.

The mean correlation matrix for fitted metabolite signals across all analysed data is given in Figure 4.12 for the TR/TE=2288/288ms spectra, and in Figure 4.13 for the TR/TE=5000/60ms spectra. Standard deviations are also shown to illustrate the stability of the covariance. High covariance was observed between: Asc, Glu, Gln and GSH; Ins and Glyc; NAA, NAAG and MM20; Lac and Thr; Cr and Cho multiplets; and Lip/MM resonances. Although the mean covariance between Cr and PCr was small, the standard deviation was high, indicating an unstable relationship between the two during fitting.

4.6.3 Prognostic accuracy

The prognostic accuracy of the MR biomarkers in predicting an abnormal outcome is shown in Table 4.2, with the optimal cut-off values and the resulting sensitivity and specificity for metabolite concentrations shown in

Table 4.4, and in the wider context of the MARBLE study in Figure 4.14. Thalamic [tNAA] was the single best prognostic indicator, with an area under the ROC curve of 0.99 (95% CI 0.94 to 1.00, n=82), while tLac/tNAA had an area under the ROC curve of 0.94 (95% CI 0.89 to 0.97, n=160). Median spectra from infants with normal and abnormal outcomes are shown in Figure 4.10 and Figure 4.11.

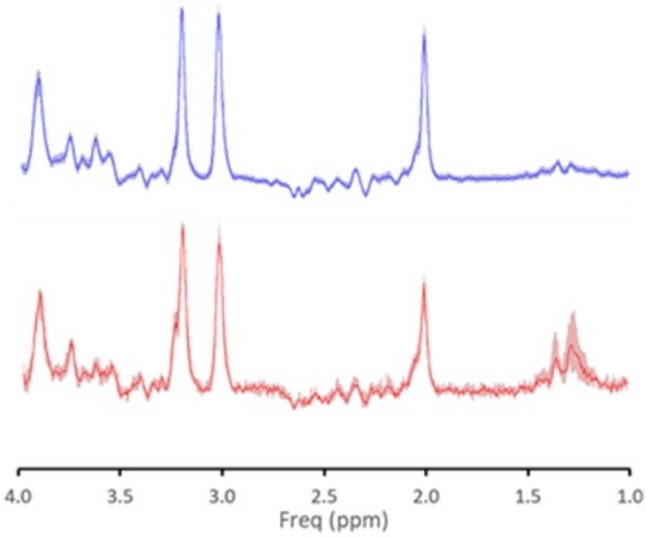

Figure 4.10: Median spectra (TR=5060ms, TE=60ms) from infants with normal (blue) and abnormal (red) outcomes two years later, with shading denoting the interquartile range. Note the decrease in the region of NAA and increase in the region of Lac with abnormal outcomes.

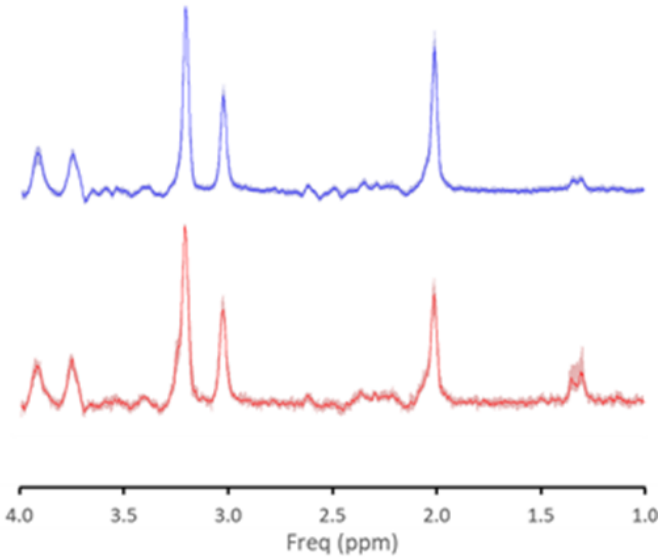

Figure 4.11: Median spectra (TR=2288ms, TE=288ms) from infants with normal (blue) and abnormal (red) outcomes two years later, with shading denoting the interquartile range. Note the decrease in the region of NAA and increase in the region of Lac with abnormal outcomes.

Clinical assessment of the neonatal encephalopathy stage at admission had an AUC of 0.71 (95% CI 0.60 to 0.82, n=190) in predicting an adverse neurodevelopmental outcome, while the stage at discharge had an AUC of 0.61 (95% CI 0.47 to 0.75, n=166).

The mean values of MRS measures in infants, grouped by normal and abnormal developmental outcomes, are shown in Table 4.2. There were significant decreases in [tNAA], [tCr], [tCho], tNAA/tCr, and tNAA/tCho, as well as significant increases in tLac/tNAA, and the T_2 of each of tCr and parenchyma water in infants who subsequently had abnormal neurodevelopmental outcomes, compared to those assessed to have normal neurodevelopment.

The relationship between each of the MRS measures and the subsequent composite motor score, corrected for gestational and postnatal age at scan, is shown in Table 4.3. The motor composite score had a significant positive correlation with [tNAA], [tCr], [tCho], tNAA/tCr, and tNAA/tCho, as well as a significant negative correlation with ln(tLac/tNAA), and the T_2 of each of tCr and parenchyma water.

The prognostic accuracy of MRS measures, in the context of the key prognostic markers in the MARBLE study protocol (Lally et al., 2015), is shown in Figure 4.14. Both tLac/tNAA and [tNAA] had an AUC which was significantly higher than any conventional MRI scores or clinical outcome assessments at discharge (DeLong test, $p<0.05$).

MRS measure	Normal	Abnormal	P-value	AUC (95% CI)
[tNAA]	7.1 (0.8)	4.0 (1.3)	**<0.001**	0.99 (0.94-1.00)
[tCr]	10.8 (1.1)	7.0 (2.4)	**<0.001**	0.98 (0.92-1.00)
[tCho]	2.8 (0.4)	2.0 (0.7)	**<0.001**	0.86 (0.76-0.93)
ln(tLac/tNAA)*	-0.67 (0.96)	-2.00 (0.37)	**<0.001**	0.94 (0.89-0.97)
tNAA/tCho*	0.84 (0.15)	0.71 (0.13)	**<0.001**	0.74 (0.67-0.81)
tNAA/tCr*	1.56 (0.23)	1.25 (0.31)	**<0.001**	0.80 (0.72-0.86)
tNAA(CH_3) $\boldsymbol{T_1}$	1410 (220)	1480 (870)	0.51	0.65 (0.43-0.87)
tNAA(CH_3) $\boldsymbol{T_2}$	480 (100)	500 (120)	0.31	0.57 (0.41-0.73)
tCr(CH_3) $\boldsymbol{T_1}$	1600 (200)	1600 (240)	0.73	0.51 (0.33-0.70)
tCr(CH_3) $\boldsymbol{T_2}$	205 (17)	219 (18)	**0.003**	0.68 (0.54-0.82)
tCho(CH_3) $\boldsymbol{T_1}$	1150 (180)	1260 (200)	0.07	0.72 (0.55-0.89)
tCho(CH_3) $\boldsymbol{T_2}$	430 (70)	430 (60)	0.92	0.51 (0.36-0.66)
Parenchyma water $\boldsymbol{T_2}$	126 (15)	140 (15)	**0.001**	0.80 (0.66-0.93)

*Table 4.2: Differences in cerebral MRS measures between infants with normal and abnormal neurodevelopment at two years of age. P-values are reported for Student t-test. AUC =Area under the receiver operating characteristic curve. *Ratio of metabolite peak areas at TR/TE=2288/288ms only, uncorrected for relaxation.*

MRS measure	R^2 change	P-value	Coefficient
[tNAA]	0.55	**<0.001**	11.9
[tCr]	0.36	**<0.001**	7.1
[tCho]	0.32	**<0.001**	24.1
ln(tLac/tNAA)*	0.44	**<0.001**	-21.2
tNAA/tCho*	0.15	**<0.001**	56.7
tNAA/tCr*	0.27	**<0.001**	41.3
tNAA(CH_3) $\boldsymbol{T_1}$	0.00	0.59	n.s.
tNAA(CH_3) $\boldsymbol{T_2}$	0.00	0.75	n.s.
tCr(CH_3) $\boldsymbol{T_1}$	0.00	0.85	n.s.
tCr(CH_3) $\boldsymbol{T_2}$	0.05	**0.01**	-0.27
tCho(CH_3) $\boldsymbol{T_1}$	0.04	0.08	n.s.
tCho(CH_3) $\boldsymbol{T_2}$	0.00	0.71	n.s.
Parenchyma water $\boldsymbol{T_2}$	0.12	**0.001**	-0.56

*Table 4.3: The relationship between each of the MRS measures and two-year motor composite score on Bayley-III. This is controlled for gestational age at birth and postnatal age at the time of scan. The coefficient represents the change in motor score for a unit change in the MRS measure. Concentrations are in terms of mmol/kg wet weight, while relaxation rates are in milliseconds. *Ratio of metabolite peak areas at TR/TE=2288/288ms only, uncorrected for relaxation.*

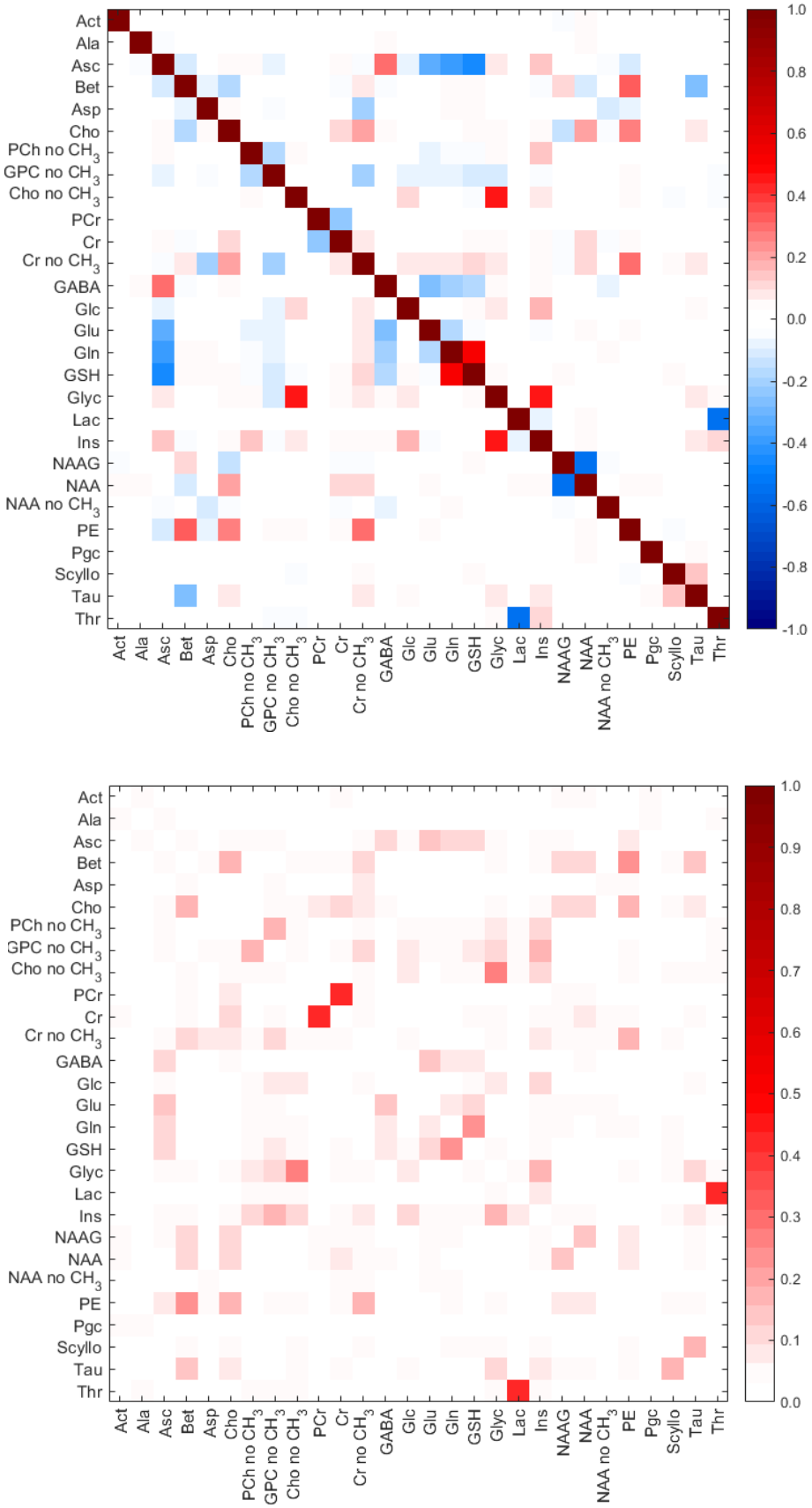

Figure 4.12: Mean (top) and standard deviation (bottom) of the correlation matrices across all of the fitted spectra acquired with TR/TE=2288/288ms in the study. 'no CH_3' indicates multiplet resonances fitted independently from the metabolite's methyl singlet peak.

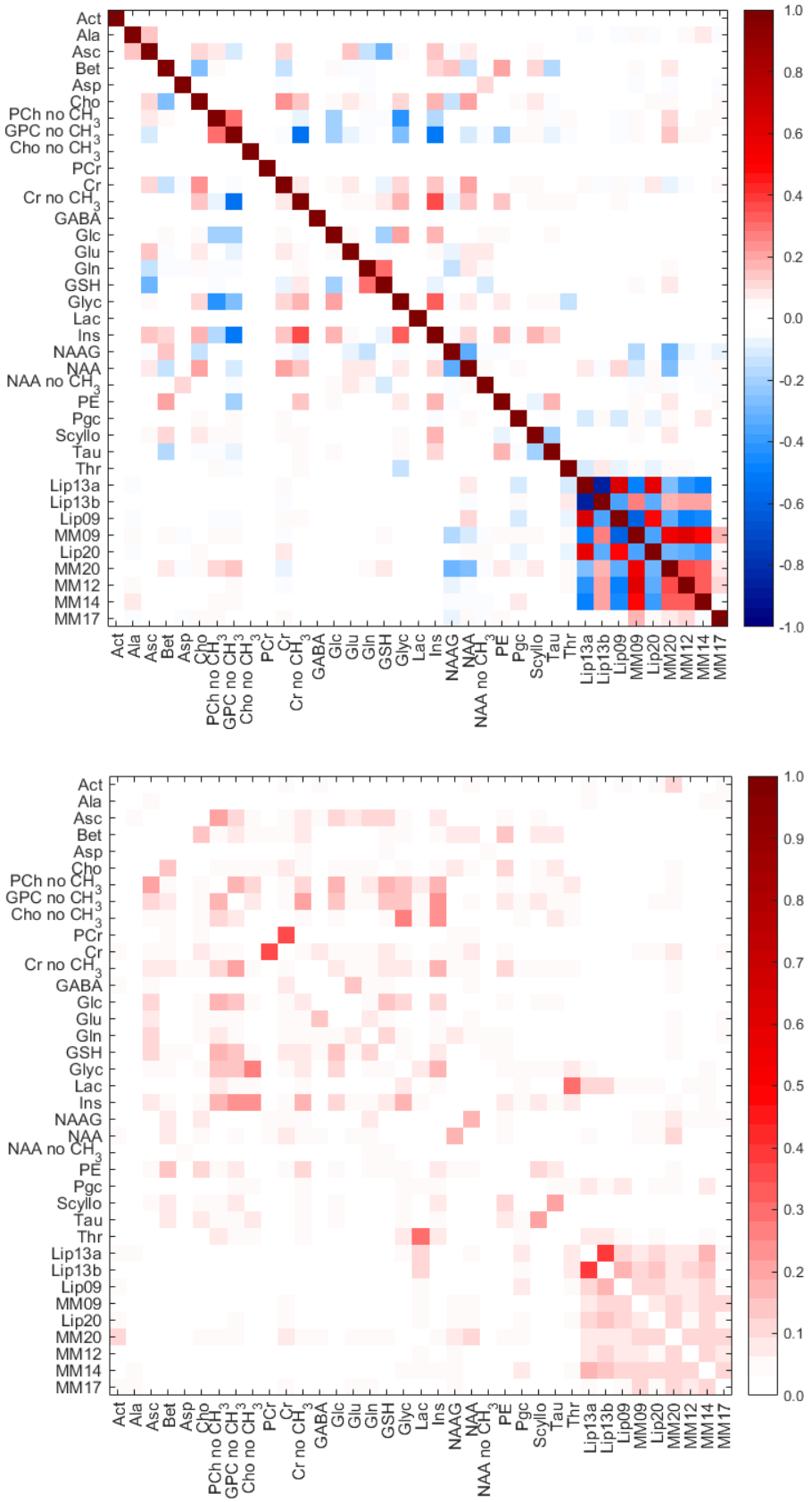

Figure 4.13: Mean (top) and standard deviation (bottom) of the correlation matrices across all of the fitted spectra acquired with TR/TE=5000/60ms in the study. 'no CH_3' indicates multiplet resonances fitted independently from the metabolite's methyl singlet peak..

MR Measure	Cut-off	TP/FP	TN/FN	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)
[tNAA]	≤5.5	12/2	68/0	100 (74-100)	97 (90-100)
[tCr]	≤9.0	11/2	68/1	92 (62-99)	97 (90-100)
[tCho]	≤2.5	10/17	53/2	83 (52-98)	76 (64-85)

Table 4.4: Sensitivity and specificity of each metabolite concentration at the optimal cut-off. TP = true positive; FP = false positive; TN = true negative; FN = false negative.

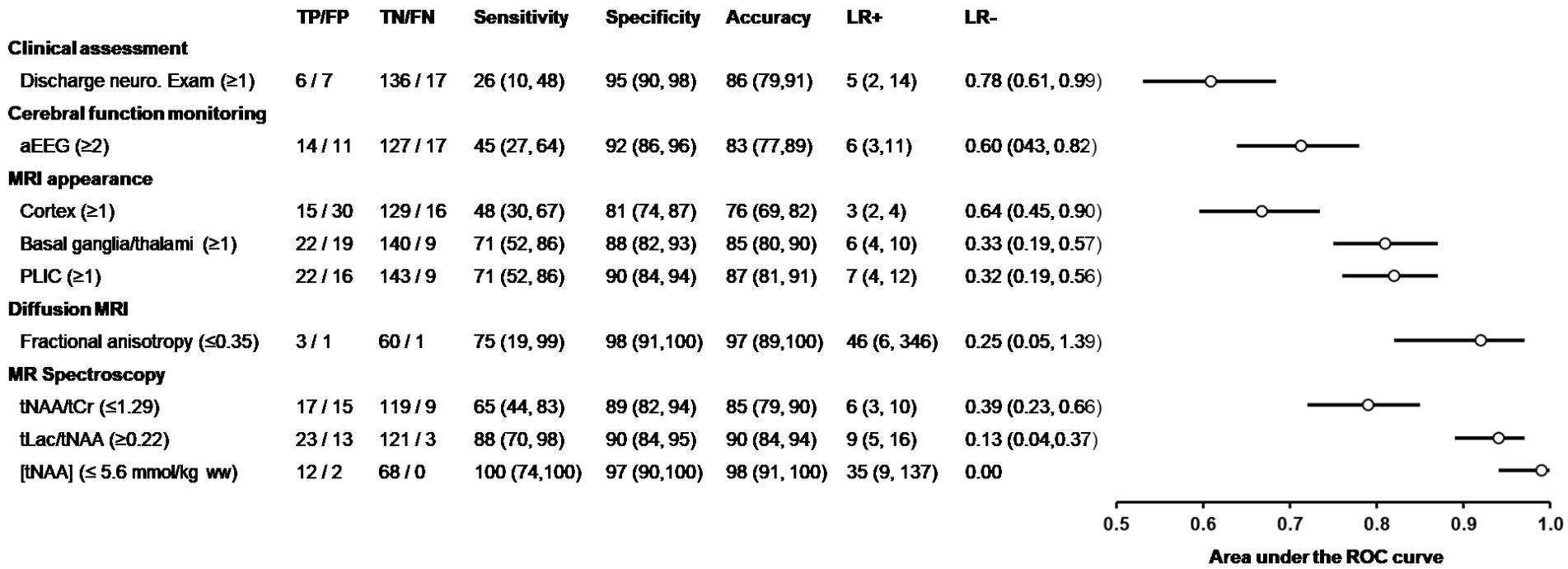

Figure 4.14: Prognostic accuracies of various measures analysed in the MARBLE study. Parentheses and error bars indicate 95% confidence intervals, using the exact binomial method. TP = true positive; FP = false positive; TN = true negative; FN = false negative; LR+ = positive likelihood ratio; LR- = negative likelihood ratio; aEEG = amplitude integrated electroencephalogram.

4.6.4 Linear regression model

The concentrations and relaxation rates in Table 4.3 were then appraised for inclusion in a multivariate linear regression model, along with gestational and postnatal age at scan, in order to predict the later composite motor scores. The most parsimonious model was selected in which the mean-squared cross validation error within one standard error of the minimum. This model contained a single variable, [tNAA], uncorrected for either gestational or postnatal age at scan (Figure 4.15).

The resulting linear model had an intercept of 20.6 (SE 8.2) and a coefficient of 10.8 (SE 1.2) as follows:

$$2\text{y motor composite score} = 20.6 + [\text{tNAA}] \times 10.8 \qquad (4.2)$$

where [tNAA] is provided in mmol/kg wet weight. This model had an adjusted R^2 of 0.51 ($p<0.001$), and did not include an adjustment for gestational at birth or postnatal age at scan.

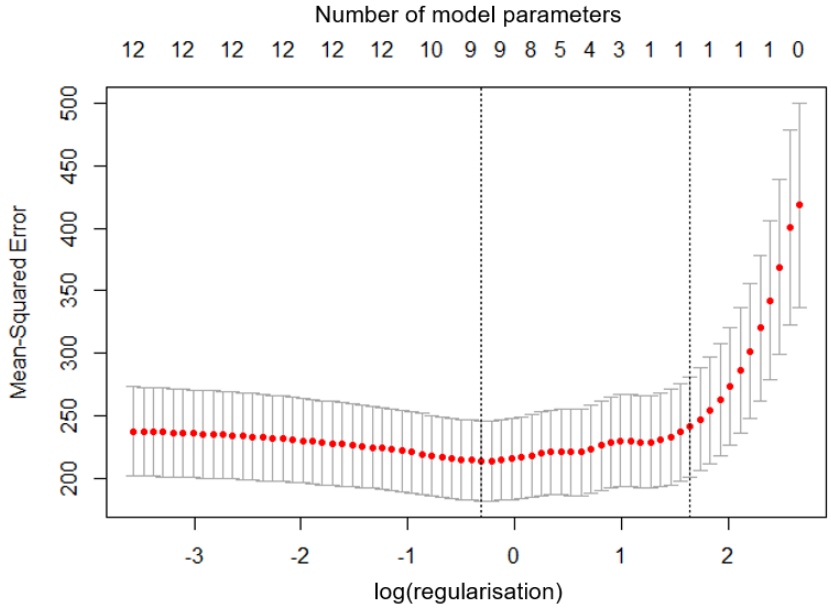

Figure 4.15: Plot of mean-squared cross-validation error as the lasso regularisation term is increased, reducing the number of terms in the linear model. The leftmost vertical dashed line indicates the linear model with the lowest cross-validation error, including nine parameters. The rightmost vertical dashed line indicates the most parsimonious linear model within one standard error of the lowest cross-validation error, which included [tNAA] as the only variable.

4.6.5 Sensitivity analysis

The inclusion of a correction factor for the concentrations derived from GE scanners had minimal impact on the results of the various analyses (see Appendix VI for a comparison).

4.7 Discussion

In this prospective multi-site study of encephalopathic babies who received therapeutic hypothermia, MRS biomarkers measured soon after birth accurately predicted adverse outcomes observed two years

later. In particular, the thalamic concentration of tNAA had the highest accuracy in predicting a binary abnormal/normal neurodevelopmental outcome following neonatal encephalopathy (AUC 0.99, 95% CI 0.94 to 1.00), and was associated with motor performance at two years of age.

tLac/tNAA had an overall AUC of 0.94 (95% CI 0.89 to 0.97) with optimal criteria giving a sensitivity of 89% (95% CI 70% to 98%) and specificity of 91% (95% CI 85% to 95%), confirming a previous meta-analysis. This measure can be obtained in a single short acquisition during MR scanning, although the metabolic information obtained is ambiguous. Not only is it affected by changes in metabolite concentrations, but also their relaxation characteristics, as well as the sensitivity of the acquisition technique to chemical shift artefacts. Multi-centre data can therefore only be directly compared where the protocols are carefully harmonised. In the present study, this was ensured through validation measurements in a spectroscopy phantom with known properties.

Clinical examination at birth alone could not identify babies at risk of adverse neurodevelopment (AUC 0.71, 95% CI 0.60 to 0.82), and a second examination at discharge was even less effective (AUC 0.61, 95% CI 0.47 to 0.75).

4.7.1 Clinical characteristics

Only 31 (16%) of the 190 infants assessed at two years after their initial injury had abnormal neurodevelopment. This reflects the characteristics of the cohort, with only 10% being assessed as severely encephalopathic on admission, and 16% assessed as mild. This may either reflect improvements attributed to the implementation and rapid initiation of therapeutic hypothermia, or more likely, the increased tendency to cool babies with mild injury across many UK hospitals (Oliveira et al., 2017). While it was a local decision as to whether to initiate therapeutic hypothermia, the NICHD neurological examination performed across all sites provides an insight into the wider characteristics of infants offered hypothermia as a part of their standard care. Since hypothermia is only proven to be efficacious in moderate or severe encephalopathy, it is unclear how this practice impacts on the neurodevelopmental outcomes of infants with milder injury.

4.7.2 Interpretation of MRS changes

Each of [tNAA], [tCho] and [tCr] was reduced in infants who went on to have an adverse neurodevelopmental outcome. This together suggests either persistent neuronal dysfunction or cell death within two weeks of the initial hypoxic-ischaemic injury, despite undergoing therapeutic hypothermia. A reduction in [tCr] might suggest injury to both neurons and glial cells, however since the reductions in [tNAA] were most predictive of neurodevelopment, it would appear the brain development is most sensitive to specific neuronal injury.

The T_2 of the tCr peak was increased in infants with later adverse neurodevelopment, suggesting an increase in [Cr] relative to [PCr]. This would suggest that one aspect of the continuing injury is a degree of persistent metabolic dysfunction.

The T_2 of parenchyma water was also increased in infants with later adverse neurodevelopment. In the absence of substantial NAA T_2 changes, which would indicate an alteration in the neuronal cytosol, this change could be explained by an increase in extracellular water mobility (i.e. vasogenic oedema) with cell death. Although this may also indicate the breakdown of myelin or the delay of myelination, the overall decrease in [tCho] (which also decreases with normal myelination) suggests that this mechanism plays a less important role.

Since none of the other concentrations or relaxivities studied provide sufficient additional prognostic information in comparison to [tNAA] alone, this suggests the degree of neuronal dysfunction or death is the most important factor in determining later neurodevelopment. Other factors which describe the microstructural environment (e.g. tNAA T_2) and cellular energetics (e.g. tCr T_2) provided minimal added value to the prediction of motor scores at two years of age.

The [tNAA], [tCr] and [tCho] reported here in each group of infants is generally lower than observed by Cheong et al. (2006), and by Cady et al. (1996), however this may be explained by the existence of broad Lip/MM signals, which may appear as apparent increases in the concentrations of more mobile metabolites if unaccounted for in the fitting model.

Both Cheong et al. (2006), and Cady et al. (1996) suggest tNAA T_2 is elevated in HIE, however this was not observed here. Since those studies examined infants in the period immediately after birth, any initial neuronal swelling subsequently normalises over the coming days, reflecting the characteristic changes observed in diffusion MRI (Bednarek et al., 2012). This has important implications for the clinical use of [tNAA], as this does not require a time-critical measurement at the immediate onset of HIE.

4.7.3 Limitations

Although this is the largest prospective study of MR biomarkers in HIE, there are some important limitations to this work. Firstly, the metabolite concentrations were only assessed in a subset of the population, at five of the eight study sites. The MR protocol for assessing [tNAA] was carefully optimised to minimise errors due to physical phenomena arising from chemical shift effects and relaxation effects (both in metabolites and water), as well as technical phenomena relating to the scanner. At these five sites, there was either sufficient local technical support to implement the full protocol using manual adjustments, or the steps could be automated using the scanner's software interface. Despite this, data

was collected on clinical MR scanners from all three of the most popular vendors (Philips, Siemens, and GE).

An additional difficulty with this measurement is the long acquisition duration (approx. 25min). As a result, all data required retrospective quality assessments to exclude the effects of motion. However, based on assumptions made from the data acquired during this study, the acquisition process could be significantly shortened to make it amenable to widespread use.

Another limitation is related to the interpretation of surrogate biomarkers. While these data demonstrate the clear prognostic potential of MR biomarkers in cooled infants, it is not clear whether a proposed intervention would affect these values, and if so, whether a change in these biomarkers influences an individual subject's outcome. The only way to test this is through phase III interventional trials with the biomarkers alongside the gold standard of careful assessment of the neurodevelopmental outcome.

4.7.4 The potential of [tNAA]

Quantitative MR biomarkers can be validated across institutions and are inherently objective, making them ideal for collaborative research across multiple sites. However, these must be carefully designed to be applicable across various MR scanner hardware and software platforms, so that investigations are not restricted to recruitment at the few centres which fulfil narrow technical specifications. In this study it was stipulated that MR biomarkers should be collected on 3 Tesla scanners, as this facilitated optimal quality MR imaging and spectroscopy data in comparison to that obtained at lower field strengths.

Typically, neonatal neuroprotection trials are powered for a primary outcome derived from clinical assessments years after the intervention. As a result, such trials require large numbers of subjects, and so the smallest early phase studies will still have a duration of years. The collection of [tNAA] could therefore improve the efficiency of neonatal neuroprotective trials in one of two ways. Firstly, it could be adopted as a surrogate outcome measure in a phase II study to give a rapid 'go or no-go' assessment. Such a model was adopted by Azzopardi et al. (2016) who did not find any effect of the intervention (xenon gas) on dMRI or tLac/tNAA from MRS. Unlike the latter measure, [tNAA] is normally distributed, and more clearly separates between normal and abnormal outcomes, while explaining more of the variation in two-year outcome measures across subjects.

Alternatively, [tNAA] could be used to optimise protocol design before proceeding to a phase III trial. For example, there are several putative adjuvant therapies which require careful dose escalation studies to maximise the likelihood of neuroprotective effects and define an optimal dosing range. Such

studies could use [tNAA] as early indicators of treatment effect, so that the most promising dosing strategy can be obtained before proceeding to the next phase.

4.8 CONCLUSIONS

Thalamic ^{1}H magnetic resonance spectroscopy, and particularly [tNAA], accurately predicts neurodevelopment two years after hypoxic ischaemic encephalopathy, and is an invaluable prognostic tool in the first weeks after birth. Measures such as [tNAA] could also be applied to increase the power of neuroprotection trials, while reducing their duration.

5 DESIGNING A RAPID ACQUISITION PROTOCOL

This chapter focuses on designing a more time-efficient MR spectroscopy protocol for the routine clinical assessment of metabolite concentrations in the neonatal brain.

5.1 OUTLINE

From the previous chapter, it is evident that MR spectroscopy accurately predicts long term neurological outcomes after HIE. In particular, [tNAA] accounts for 51% of the variance in motor performance at two years of age. However, the protocol used in the clinical study was lengthy (~25 mins) and not amenable to routine clinical use. Though the method can be shortened considerably, it is important to do so while maintaining accuracy and precision so that it retains its prognostic value.

To effectively shorten the acquisition time, it is therefore imperative to identify key sources of unwanted variance in the measurement and minimise their impact.

5.2 VARIANCE DUE TO METABOLITE RELAXIVITY

5.2.1 Outline

In HIE, there are important developmental and pathological changes which affect the mobility of metabolites, and without correction these can bias concentration measurements. Since each metabolite is distributed in different concentrations across neurons and glial cells, normal and pathological changes in cellular microstructure can have unpredictable effects on metabolite relaxivities.

In infants with HIE, the results from Table 4.2 and Table 4.3 highlight that metabolite relaxivities are not effective predictors of neurodevelopmental outcome, however they nevertheless introduce undesirable variance into MRS measurements. It is therefore important to understand how much of this variance can be explained by normal maturation in the gestational period, and how much can be explained by the evolution of injury over the postnatal period.

5.2.2 Study design

The estimated relaxivities of tNAA, tCho, tCr, and parenchyma water were obtained from the MARBLE study dataset.

5.2.3 Methods

Firstly, a hierarchical linear regression was performed to examine the relationship between relaxivity and birth gestational age, while controlling for the motor outcome score (assuming this has a linear

dependence on the initial severity of injury), and the postnatal age at scan. A final hierarchical linear regression was then performed to examine the relationship between relaxivity and the postnatal age at the time of measurement, while controlling for the motor outcome score, and the gestational age at birth.

5.2.4 Results

In the MARBLE study, 144 neonates had sufficient data for metabolite T_2 estimations (n=114 with motor outcomes) and 93 had sufficient data for metabolite T_1 estimations (n=79 with motor outcomes). The mean T_1s and T_2s across the whole cohort are shown in Table 5.1:

Resonance	Mean (SD) $\boldsymbol{T_1}$ (ms)	Mean (SD) $\boldsymbol{T_2}$ (ms)
tNAA(CH_3)	1411 (361)	448 (92)
tCr(CH_3)	1597 (238)	206 (19)
tCho(CH_3)	1160 (211)	419 (72)
Parenchyma water	n/a	128 (15)

Table 5.1: Mean (SD) T_1 and T_2 values for the methyl singlet peaks and parenchyma water across the MARBLE cohort.

The regression coefficient for each relaxivity with birth gestational age is shown in Table 5.2, corrected for the motor outcome score and postnatal age.

MRS Measure	R^2 change	P-value	Change per week gestational age
tNAA(CH_3) $\boldsymbol{T_2}$	0.01	0.73	n.s.
tNAA(CH_3) $\boldsymbol{T_1}$	0.01	0.34	n.s.
tCr(CH_3) $\boldsymbol{T_2}$	0.00	0.66	n.s.
tCr(CH_3) $\boldsymbol{T_1}$	0.00	0.89	n.s.
Cho(CH_3) $\boldsymbol{T_2}$	0.01	0.29	n.s.
Cho(CH_3) $\boldsymbol{T_1}$	0.00	0.68	n.s.
Parenchyma water $\boldsymbol{T_2}$	0.19	**<0.001**	-4.3 ms

Table 5.2: Relationship between relaxation rates and birth gestational age, controlling for motor outcome score at two years and postnatal age.

The only significant change could be attributed to the T_2 of parenchymal water, which decreased at a rate of 4.3ms per week of gestation when controlling for the degree of injury. The equivalent regression coefficient for each relaxation time with postnatal age is shown in Table 5.3, corrected for the degree of injury and gestational age at birth.

MRS Measure	R^2 change	P-value	Change per week postnatal age
tNAA(CH_3) $\boldsymbol{T_2}$	0.00	0.91	n.s.
tNAA(CH_3) $\boldsymbol{T_1}$	0.00	0.68	n.s.
tCr(CH_3) $\boldsymbol{T_2}$	0.01	0.24	n.s.
tCr(CH_3) $\boldsymbol{T_1}$	0.01	0.41	n.s.
Cho(CH_3) $\boldsymbol{T_2}$	0.00	0.93	n.s.
Cho(CH_3) $\boldsymbol{T_1}$	0.01	0.54	n.s.
Parenchyma water $\boldsymbol{T_2}$	0.00	0.87	n.s.

Table 5.3: Relationship between relaxation rates and postnatal age, controlling for motor outcome score at two years and gestational age at birth.

No significant changes were observed in any of the relaxation rates with postnatal age. The negative relationship observed between the T_2 of parenchyma water and gestational age was not evident in the postnatal period.

5.2.5 Discussion

There was a significant correlation between parenchyma water T_2 and birth gestational age, but not with postnatal age. This may be related to the well characterised time-course of cell swelling which is common across all infants in the postnatal period, and usually reflected by the 'pseudonormalisation' of diffusion MRI (Bednarek et al., 2012). The decrease in parenchyma water T_2 with birth gestational age (4.3ms per week) is similar to that observed by Kreis et al. (2002) in the period between 32-41 weeks (4.1ms per week).

There was no evidence to suggest that the relaxivity of the prominent methyl singlet resonances for tNAA, tCho, and tCr were affected by maturational changes. This, in combination with Table 4.1, indicates that much of the variation in relaxivities is unpredictable.

A common practice is to apply fixed values of these relaxation parameters to avoid the relaxometry experiments altogether, for example taking the mean values in Table 5.1, but the variance across all subjects is sufficiently wide to produce errors in the estimated signal intensity.

Figure 5.1 illustrates the estimated magnetisation relaxation characteristics from the mean T_1s and T_2 for the methyl tNAA resonance (red), along with the 95% confidence intervals (grey). These intervals are then shown normalised to the mean value (right), to illustrate the uncertainty associated with applying a fixed T_1 or T_2 correction. These plots overestimate the true biological variance across the cohort, as this is affected by additional experimental errors in Table 4.2 from the two-point fit.

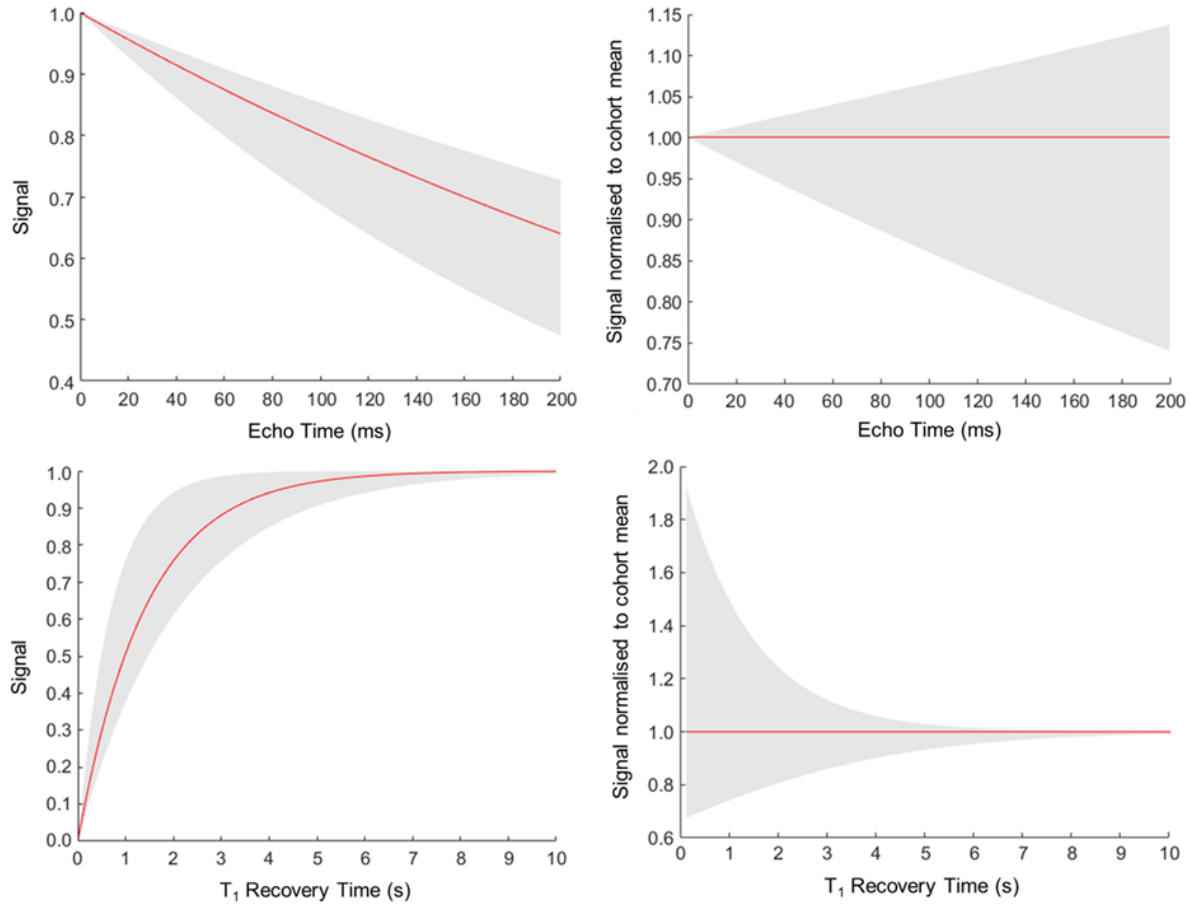

Figure 5.1: Top left) Profile of T_2 relaxation corresponding to the cohort mean for tNAA (red) with 95% confidence intervals (grey). Top right) The same 95% confidence intervals in terms of the cohort mean profile, giving an insight into the uncertainty when applying a fixed T_2 correction. Bottom row) The equivalent plots for T_1 relaxation of tNAA.

This gives some insight into the extent that errors can be reduced by increasing the longitudinal recovery time, and shortening the echo time. The former is associated with a penalty in terms of an increased acquisition time, while the latter is dependent on the technical specifications of the scanner

(i.e. gradient performance) and the sequence implementation (i.e. RF waveforms). It is important to note that most investigators employ a single repetition time of ~2s in the neonatal brain and use this to estimate metabolite concentrations (for example Wisnowski et al. (2016)). In the case of tNAA, this is associated with a substantial error of up to ±20% due to the variations in measured T_1 relaxation alone, which is not associated with prognosis (see Table 4.2 and Table 4.3). While the biological variance in the T_1 of tNAA will be lower than observed in the data here, as there are additional errors introduced by a two-point experiment, there will still be substantial error when extrapolating from a short TR to the fully relaxed state at thermal equilibrium.

5.2.6 Conclusions

There is substantial variation in the estimated relaxivities of metabolites and parenchymal water in the neonatal period after HIE, but this is unpredictable in nature. Since these are not predictive of later neurodevelopment, an optimal measurement of metabolite concentrations should minimise the uncertainty associated with this variation.

5.3 Temporal changes in metabolite concentrations

5.3.1 Outline

While [tNAA] accurately predicts two-year motor outcomes after HIE irrespective of the birth gestational age or postnatal age, it is important to know how much this and other metabolites vary over the timescale of admission to the neonatal unit, as this would necessitate time-dependent prognostic criteria. This would also provide some insight into characteristics of the evolving injury.

5.3.2 Study design

The estimated concentrations of tNAA, tCho, and tCr were obtained from the MARBLE study dataset.

5.3.3 Methods

Firstly, a hierarchical linear regression was performed to examine the relationship between metabolite concentrations and the birth gestational age, while controlling for the eventual motor outcome score, and the postnatal age at scan.

Secondly, a hierarchical linear regression was performed to examine the relationship between metabolite concentrations and the postnatal age at the time of measurement, while controlling for the eventual motor outcome score, and the gestational age at birth.

5.3.4 Results

The regression coefficients for each concentration with birth gestational age is shown in Table 5.4, corrected for the eventual motor outcome score (i.e. the assumed injury severity) and postnatal age.

Resonance	R^2 change	P-value	Change per week gestational age
[tNAA]	0.08	**<0.001**	+0.25 mmol/kg ww
[tCho]	0.00	0.89	n.s.
[tCr]	0.03	0.06	n.s.

Table 5.4: Relationship between metabolite concentration and birth gestational age, controlling for postnatal age at scan and two-year motor score on Bayley-III.

The only significant change could be attributed to [tNAA], which increased at a rate of 0.25 mmol/kg wet weight per week of gestation. The regression coefficients for each concentration with birth gestational age is shown in Table 5.5, corrected for the eventual motor outcome score and postnatal age.

Resonance	R^2 change	P-value	Change per week postnatal age
[tNAA]	0.02	0.06	n.s.
[tCho]	0.08	**0.004**	+0.23 mmol/kg ww
[tCr]	0.03	0.06	n.s.

Table 5.5: Relationship between metabolite concentration and postnatal age at scan, controlling for birth gestational age and two-year motor score on Bayley-III.

The only significant change could be attributed to [tCho], which increased at a rate of 0.23 mmol/kg wet weight per week in the postnatal period. The positive relationship between [tNAA] and gestational age was not evident in the postnatal period.

5.3.5 Discussion

There are important changes in metabolite concentrations which are not associated with adverse outcomes, but which affect the results of a measurement over the neonatal period. The increased [tNAA] with birth gestational age agrees with the increases observed during normal development by Kreis et al. (2002), and this relationship was not observed in the postnatal period. This suggests that the normal trajectory of increasing [tNAA] is interrupted by HIE, and flattens out in the initial period after

birth. However, this overlooks the possibility that a more mature brain to be more resilient to hypoxic-ischaemia. In any case, these results indicate that [tNAA] is a stable prognostic measure over the first two weeks after birth, provided birth gestational age is taken into account.

The significant and rapid increase in [tCho] with postnatal age is surprising, as this would be expected to decrease with normal myelination. Rather than suggesting a delay in normal myelination, this would indicate continuing inflammation in a tertiary phase of injury, as outlined by Fleiss and Gressens (2012), which occurs in the months and years after the initial injury .

5.3.6 Conclusions

There are important variations in metabolite concentrations with gestational and postnatal age which reflect underlying developmental and inflammatory processes. These would suggest that time-dependent prognostic criteria are appropriate for routine clinical use.

5.4 VARIANCE DUE TO PARENCHYMAL WATER CONCENTRATION

The water concentration of cerebral tissue varies with gestational age (Lentner, 1981, Kreis et al., 1993), although the effect of such changes is small until months after birth. For example, from term to one month of age, water content (as derived from data published by Lentner (1981)) varies by <1%, which would be below the noise floor of most imaging experiments which examine parenchymal water concentration (e.g. multiple flip angle spoiled gradient echo).

On the other hand, it is important to consider the effects of oedema on the observed water concentration in the neonatal brain, as this is commonly observed with hypoxic ischaemic encephalopathy. In the case of cytotoxic oedema (which is typically evident over the first 10 days after birth), the concentration is unaffected as water is redistributed from the extracellular to the intracellular space. At later times, vasogenic oedema is more prevalent, in which case the water concentration will increase due to the filling of extracellular space. In this case, it is important to accurately measure changes to the parenchyma water compartments, for example using T_2 relaxometry of the water signal via MRS or MRI.

5.5 PROPOSED SOLUTION: TWO-POINT SATURATION RECOVERY STEAM

5.5.1 Acquisition details

The following protocol comprises 13 separate STEAM acquisitions in the same 15x15x15mm^3 thalamic voxel, with varying parameters as shown in Table 5.6, and processed according to the steps outlined in

Figure 5.2. Although there are multiple steps, the duration is significantly shorter than in the PRESS protocol for the previous study, at approximately 10 minutes in total.

Step	TR/TM/TE (ms)	Water suppression	Phase cycling	Sub-spectra	Dummy scans	Cent. freq. (ppm)
1	1500/20/20	On	4	24	1	2.01
2	1500/20/20	Off	4	1	1	4.67
3	3500/20/20	On	4	20	1	2.01
4	3500/20/20	Off	4	1	1	4.67
5	9030/20/20	Off	1	1	1	4.67
6	9040/20/40	Off	1	1	1	4.67
7	9060/20/80	Off	1	1	1	4.67
8	9090/20/140	Off	1	1	1	4.67
9	9130/20/220	Off	1	1	1	4.67
10	9170/20/300	Off	1	1	1	4.67
11	9270/20/500	Off	1	1	1	4.67
12	9420/20/800	Off	1	1	1	4.67
13	9770/20/1500	Off	1	1	1	4.67

Table 5.6: Key acquisition parameters for the full STEAM MRS protocol. Each dummy scan represents a single acquisition of all phase cycling steps.

5.5.2 Choice of acquisition scheme

Although STEAM acquisitions provide a lower inherent signal-to-noise ratio than PRESS, they are less susceptible to chemical shift displacement artefacts, and can achieve a shorter TE to minimise variation due to the variance T_2 relaxation effects outlined in Section 5.2.

At the time of writing, on GE scanners, the RF centre frequency can only be tuned with research privileges provided by the vendor. Even then, the centre frequency of each pulse in the sequence must be manually entered, in turn, at the start of each acquisition. This is both cumbersome and prone to operator errors, and so STEAM mitigates these effects so that chemical shift artefacts are minimised (see Section 2.4.3).

Also, at the time of writing, Siemens scanners do not allow TE>300ms for STEAM acquisitions without research privileges from the vendor. This limits the sensitivity of water T_2 relaxometry to the segmentation of mobile water, and so a fixed long T_2 component must be included in the fitting process.

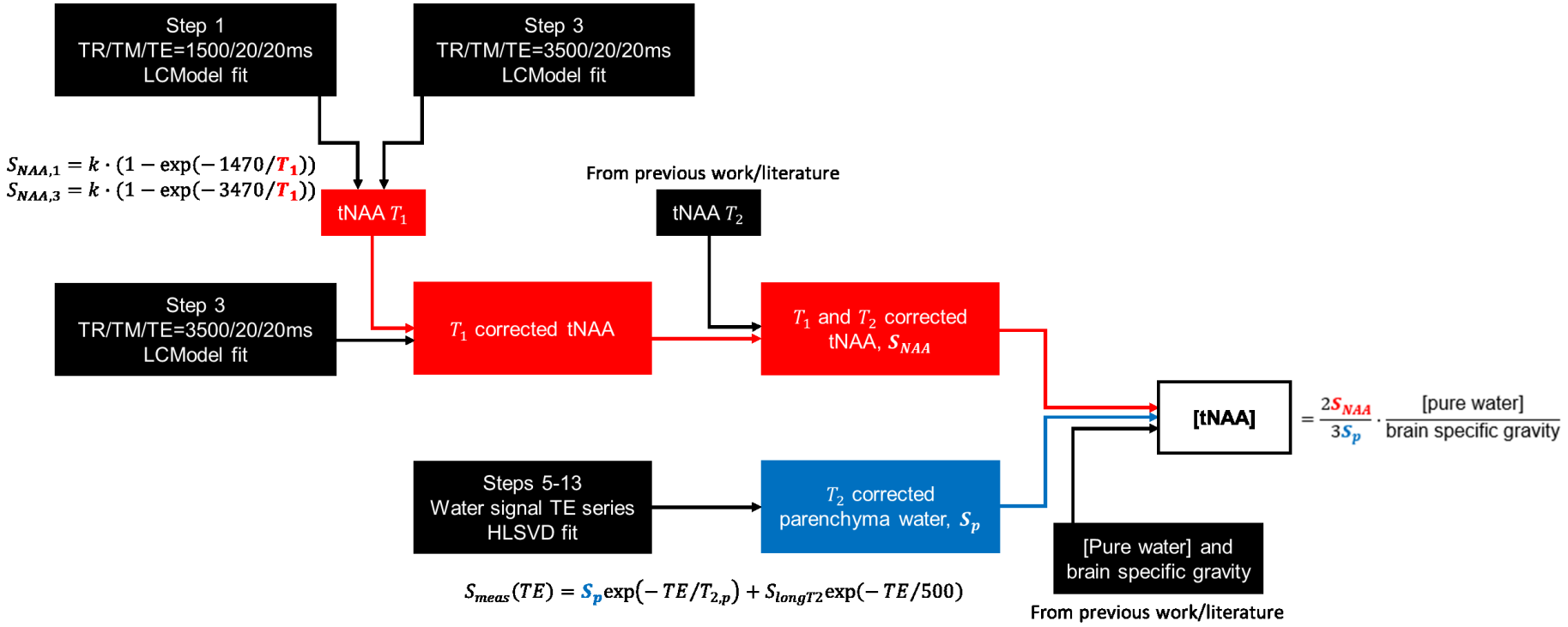

Figure 5.2: Calculation steps to derive thalamic tNAA concentration, [tNAA], from the data collected in the various steps of the protocol. Black boxes show the input data, either obtained from the LCModel/HLSVD fits or reference values from previous work/literature. Red boxes and text denote calculated quantities from the tNAA resonance, and blue boxes and text denote calculated quantities from the water signal. Equations accompany the different calculation steps. $S_{NAA,step}$ denotes the tNAA signal intensity in the corresponding protocol step, with the relaxation corrected intensities for tNAA and parenchyma water given by S_{NAA} and S_p respectively.

5.5.3 Steps 1-4: two-point saturation recovery STEAM

Steps 1 and 3 comprise a two-point T_1 relaxometry experiment for each metabolite, allowing the calculation of fully T_1 relaxed signal intensities. The short TE ensures that differential T_2 decay effects between metabolites and water are minimised.

Steps 2 and 4 allow eddy current correction of steps 1 and 3, and can be combined to determine the fully T_1 relaxed brain water signal under the same phase cycling scheme as steps 1 and 3.

5.5.4 Steps 5-12: brain water T_2 relaxometry

In the remaining steps, the TR is incremented along with the TE to enable T_2-based segmentation of mobile extra-parenchymal water. These steps could also be replaced by an imaging-based segmentation technique.

5.5.5 Lipids and macromolecules

The use of a short TE complicates the fitting of key metabolites since they overlap with broader resonances from lipids and macromolecules. However, their short T_1 ensures that they are fully relaxed throughout the experiment. This information could later be exploited to improve the fitting process.

5.6 Phantom Validation

5.6.1 Outline

In proposing an alternative acquisition scheme, it is important to verify that it yields equivalent measurements to the previous scheme. The first step in this process is to perform an experiment in a phantom of fixed metabolite concentrations.

5.6.2 Methods

The MRS protocols for the PRESS and STEAM experiments have been outlined in Table 4.1 and Table 5.6 respectively. Each protocol was used to quantify the concentration of acetate in a spectroscopy phantom, also containing ethanol. Acetate comprises a singlet peak at 1.92ppm due to 3 methyl protons, and so provides a reasonable estimate of the singlet peak of N-acetylaspartate (singlet peak at 2.01ppm due to 3 methyl protons). Corrections for T_1 and T_2 of both acetate and water were performed by fitting the signals to the ideal signal equations, and an acetate concentration calculated assuming a water concentration of 55.5M. Due to the simplicity of the spectrum, all acetate and water peaks were fitted in jMRUI using the HLSVD algorithm (Pijnappel et al., 1992) in the region of the acetate singlet.

5.6.3 Results

Example plots of the raw data from water suppressed STEAM and PRESS acquisitions are given in Figure 5.3. These STEAM spectra shown have an inherently lower signal-to-noise ratio than the PRESS spectra since they use:

- lower flip angle refocusing pulses (90° vs 180°)
- half the number of phase cycles per spectrum (i.e. number of excitations)
- a shorter longitudinal magnetisation recovery time (3470ms vs 4940ms)

To compensate for these losses, the protocol includes a larger number of signal averages.

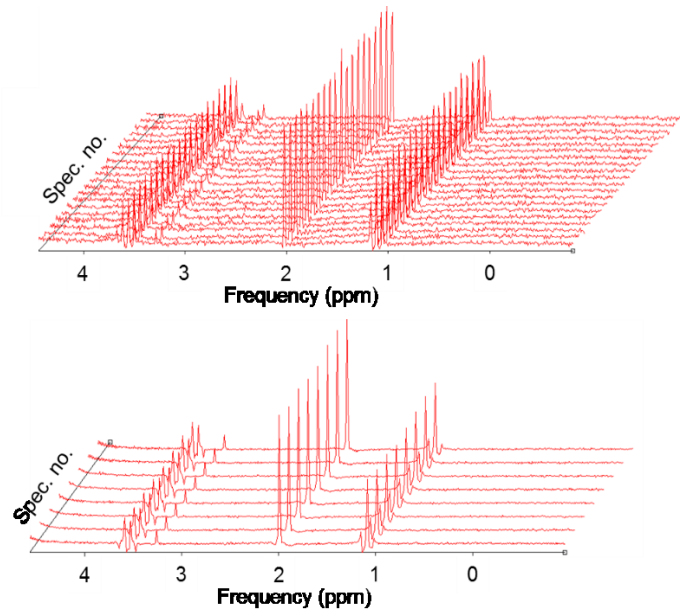

Figure 5.3: Spectra obtained using the STEAM (top, TR/TM/TE=3500/20/20ms) and PRESS (bottom, TR/TE=5000/60ms) protocols

The recorded signal intensities from acetate and water are shown in Figure 5.4 for both the STEAM and PRESS methods. The acquisition time for the PRESS acquisitions was approximately 25min, while the STEAM protocol had a duration of approximately 10min. The fitted relaxation times, peak areas and concentrations are given in Table 5.7, along with standard errors calculated from error propagation (starting with the standard deviations of signal intensities between sub-spectra).

The estimates of the concentration and T_1 of acetate were in agreement between protocols, but the estimated water T_2 was shorter in the STEAM protocol.

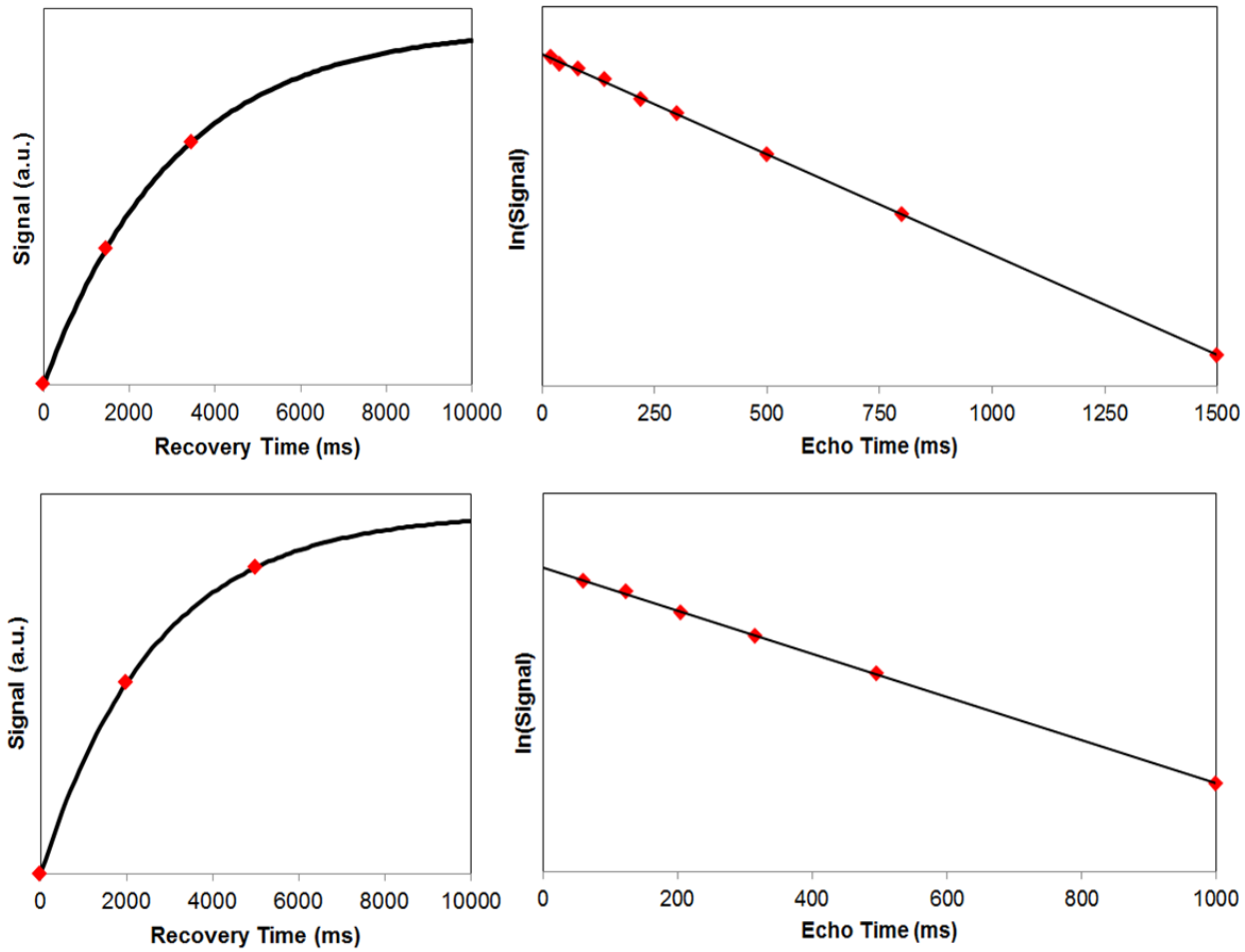

Figure 5.4: Left) Fitted acetate signal from the STEAM protocol (top) and PRESS protocol (bottom) along with fitted T_1 recovery curve. Right) Fitted water signal from the STEAM protocol (top) and PRESS protocol (bottom) along with the fitted T_2 decay.

Protocol	Water T_2 (ms)	Water T_1 (ms)	Acetate T_2 (ms)	Acetate T_1 (ms)	[Acetate] (mM)
PRESS	1250 (70)	---	2500 (500)	2700 (400)	71 (5)
STEAM	1050 (20)	1400 (300)	---	3100 (600)	73 (9)

Table 5.7: Estimates of relaxivities and concentrations (with standard errors in parentheses) derived from the phantom experiment.

5.6.4 Discussion

The resulting *in-vitro* acetate concentrations from PRESS and STEAM were within 3% agreement, providing proof of principle in this phantom model. Since the relaxation times in the phantom are much longer than in-vivo, and since there is only a single homogeneous compartment, the next step in validation was to measure agreement between PRESS and STEAM *in-vivo*.

5.7 OPTIMISING TWO-POINT SATURATION RECOVERY

5.7.1 Outline

Prior to implementing an *in-vivo* protocol, it is important to optimise the measurement based on the characteristics of the neonatal brain. For a given acquisition two-point saturation recovery experiment, a TR pair must be chosen which minimises the error on the concentration measurement.

5.7.2 Methods

A Monte Carlo simulation (Metropolis and Ulam, 1949) of the two-point saturation recovery STEAM experiment was performed (MATLAB 2017a), with a nominal T_1 of 1500ms, adding normally distributed noise to signal estimates at each TR (with an equal coefficient of variation on each). Specifically, for each run, these simulated signals were fitted to an ideal T_1 recovery curve (see Section 2.2.4) using the lsqcurvefit function in MATLAB (2017a). This process was repeated 3000 times to generate estimates of the mean bias and standard deviation in the measurement of the fully relaxed signal and metabolite T_1, by comparing with the nominal values. An outline of the experiment is shown in Figure 5.5.

For all experiments, the first TR in the two-point experiment was fixed at 1500ms, as this could give a reasonable signal-to-noise ratio (SNR) for metabolites in as short an acquisition time as possible, as illustrated in Figure 5.6.

This process was then repeated, varying the relative errors on the signal at each TR in order to determine their influence on the estimates.

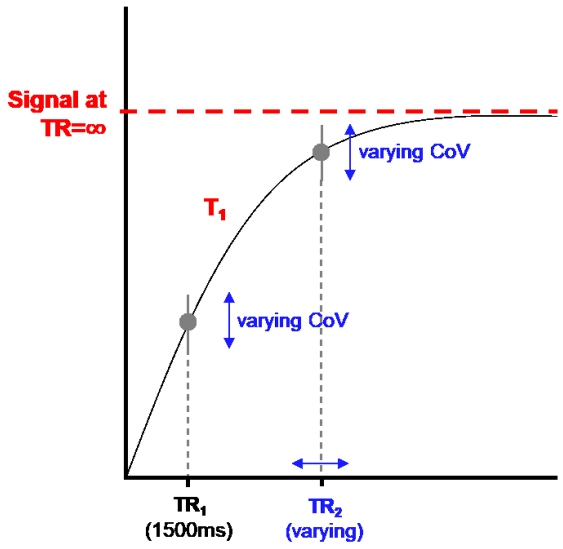

Figure 5.5: Illustration of the Monte Carlo experiment to quantify the uncertainty in the relaxation corrected signal and T_1 (red) as experimental conditions are varied (blue) for two-point saturation recovery.

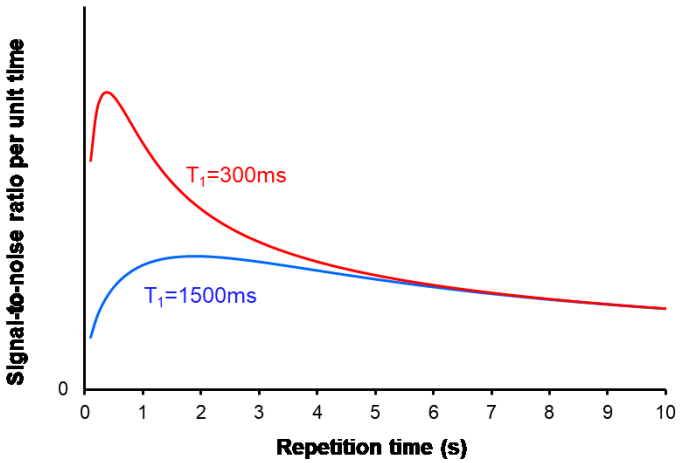

Figure 5.6: Theoretical signal-to-noise ratio per unit time plotted against repetition time. Plots are shown to represent the behaviour of short T_1 macromolecules (red) and longer T_1 metabolites (blue).

5.7.3 Results

The mean bias and standard deviation for the fully T_1 relaxed signal are shown in Figure 5.7, where the second TR in the experiment was incremented from 2s to 10s, and the coefficient of variance (CV) on the measurement at both TRs was incremented from 1% to 10%. The mean and standard deviations for the estimates of the metabolite T_1s (with a nominal value of 1500ms) in the same experiment are also shown. By increasing the second TR from 3500ms to 5000ms, the bias on the signal estimate is reduced by a factor of approximately 3.

With a longer second TRs and smaller CVs, the estimated T_1 consistently showed a small bias of approximately +5%. This increased with a shorter second TRs and greater coefficients of variation.

The effect of unequal CVs between the first and second spectra in the experiment are shown in Figure 5.8 (top four panels) for the fully relaxed signals and T_1s, with the second TR fixed at 3500ms as suggested in Section 5.5.1. The bottom four panels in Figure 5.8 then show the same plots with the second TR extended to 5000ms. The bias and variance is reduced in both the fully relaxed signal and the estimated T_1 when employing this longer TR.

A longer TR permits fewer signal averages in the same time, and so employing a TR of 5000ms would reduce the signal-to-noise ratio by a factor of ($\sqrt{3.5/5} =$) 0.84 compared to that with 3500ms. However, the additional signal recovery of the metabolites at 5000ms compensates for some of these losses, resulting in an overall factor of 0.89 (assuming a T_1 of 1500ms). The cost of this small reduction in signal-to-noise ratio is offset by the much-reduced bias and variance in the signal estimate (see Figure 5.8).

The number of signal averages is then a balance between achieving sufficient signal-to-noise ratio in the 1500ms TR acquisition to allow an adequate fit, and spending sufficient time at the longer second TR to minimise the variance due to partial saturation effects (as shown in Figure 5.8).

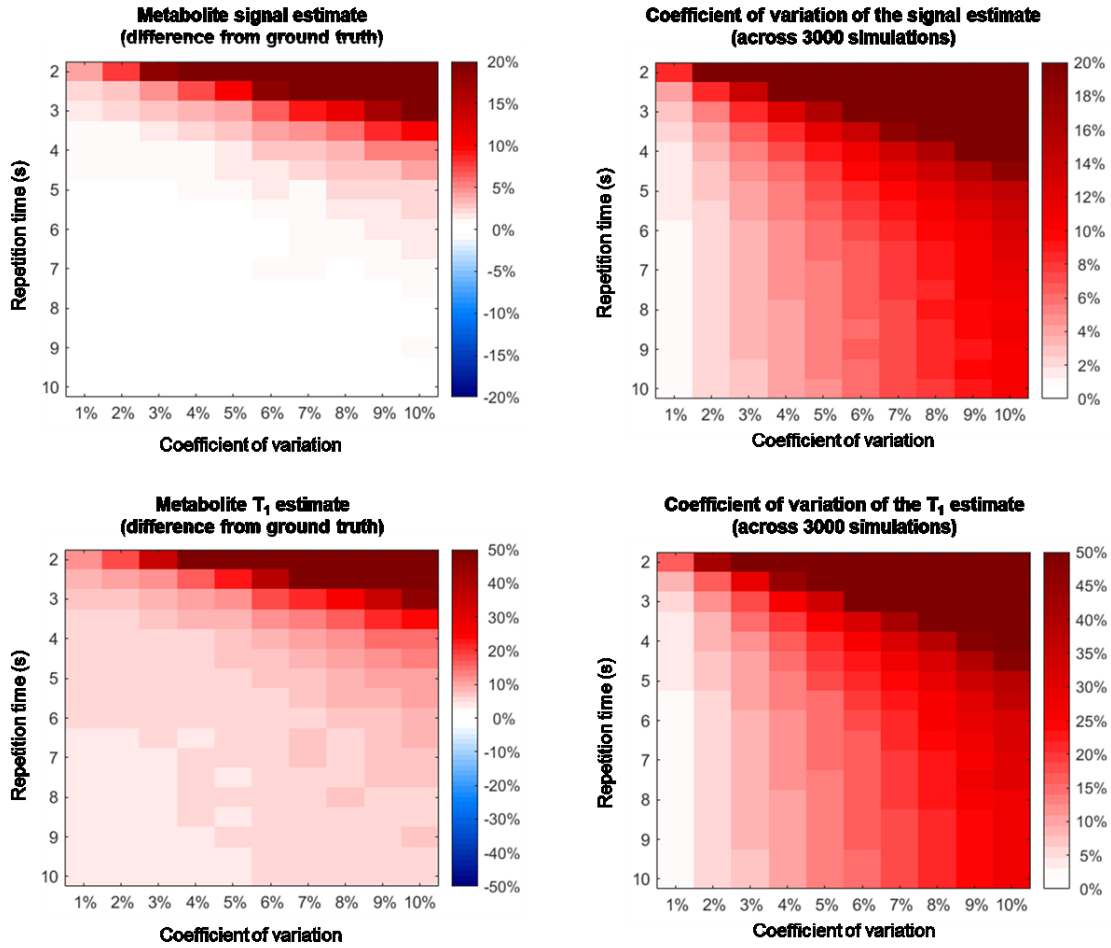

Figure 5.7: Left) Mean estimate of the fully relaxed metabolite signal (top) and metabolite T_1 (bottom), as a percentage difference from the ground truth. Right) Coefficient of variation on the mean estimate of the fully relaxed metabolite signal (top), and metabolite T_1. Both are plotted as a function of the second TR in the experiment (y axis) and the coefficient of variation on the signal at both TRs (x axis).

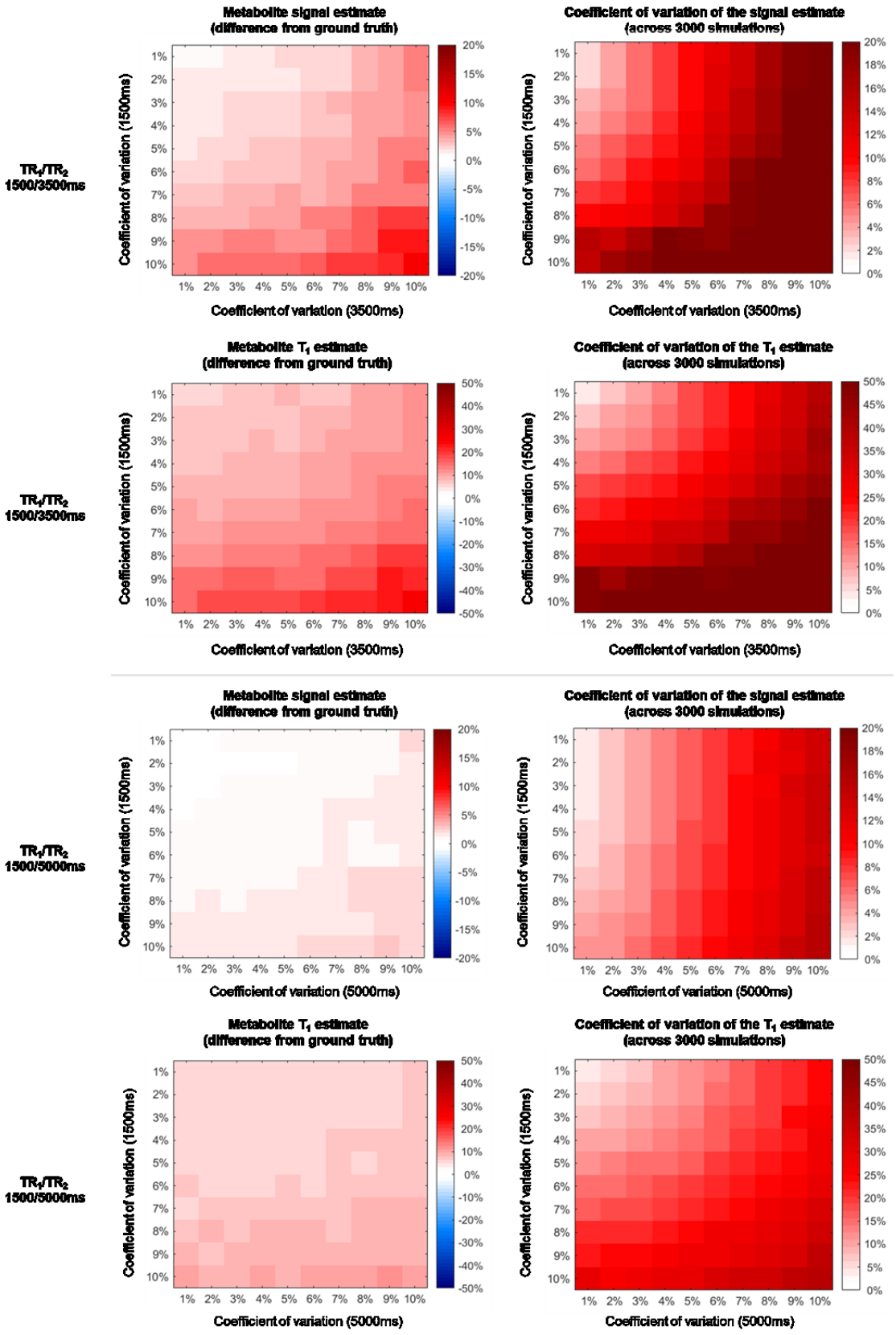

Figure 5.8: Left) Mean estimate of the fully relaxed signal (first and third row) and T_1 (second and fourth rows), as a percentage difference from the ground truth. Right) Coefficient of variation on the mean estimate of the fully relaxed signal (first and third row), and T_1 (second and fourth row). The top four panels relate to TR_1/TR_2=1500/3500ms pair, the bottom four to TR_1/TR_2=1500/5000ms.

5.7.4 Discussion

By increasing the TR of the second acquisition in the two-point saturation recovery experiment, the estimates of fully relaxed signals and T_1s are less biased and more precise.

In the case where the longer TR is at 3500ms, estimates are sensitive to the signal to noise ratio of both spectra in the experiment. If the longer TR is instead at 5000ms, fully T_1 relaxed signal intensities are largely independent of the error at the shorter TR. Fully T_1 relaxed signal estimates in both cases are more precise than the estimates of T_1s themselves, which is a desirable property for robust estimation of metabolite concentrations. The poorer precision in T_1 estimates from a two-point experiment may explain some of the variance observed in Section 5.2.

In the light of the imprecision in T_1 estimates, there is an argument to instead use a single TR of >5T_1 in which case all metabolite signals will be fully relaxed. However, the common baseline signals obtained across the two-point experiment could instead be exploited to improve the separability of mobile metabolites from Lip/MM resonances (see Section 7).

The extra time spent on the extended TR (=5000ms) is also better spent than equivalent increases in the number of signal averages at TR=3500ms. Comparing the plots in the top and bottom sections of Figure 5.8, it is evident that such a marginal increase in spectral SNR at a TR of 3500ms is far outweighed by the additional instability in the fitting of the two-point saturation recovery curve. However, extending the TR too far results in poorer efficiency as shown in Figure 5.6, and fewer sub-spectra to permit post-hoc motion correction on collected data. A TR of 5000ms is then a reasonable balance of signal-to-noise ratio efficiency and overall fitting accuracy, while also allowing effective post-hoc motion correction on the collected sub-spectra.

5.7.5 Conclusion

For *in-vivo* experiments, it is important to optimise the two-point saturation recovery experiment by choosing an appropriate pair of TRs. This optimises the technique for accurate and precise routine measurement of metabolite concentrations.

6 VALIDATING THE CLINICAL PROTOCOL *IN-VIVO*

This chapter focuses on the validation of the proposed *in-vivo* STEAM protocol against the lengthier PRESS protocol used previously (Section 4.2).

6.1.1 Outline

While the STEAM technique may be accurate and precise *in-vitro*, there are important differences in applying this strategy *in-vivo*. Firstly, there is a greater number of metabolites which contribute to the detected signal, and their resonances overlap to make the resulting inverse problem more difficult. Secondly, the relaxation times of metabolites are generally shorter (unless an appropriate relaxivity agent is introduced in the *in-vitro* case), which broadens spectral linewidths to further complicate the inverse problem. Finally, the metabolism between *in-vivo* subjects will vary, both in the normally developing neonatal brain and in pathology.

It is therefore important to validate the optimised protocol *in-vivo*, investigating the robustness of the technique to normal and pathological variations.

6.1.2 Study design

Eleven term neonates were recruited between March and October 2016 at a single site (Imperial College London). The study was approved by City and East Research Ethics Committee (15/LO/1091), North London Research Ethics Committee (11/H0717/6), and all parents provided written informed consent.

6.1.3 Methods

These were a heterogeneous population to reflect the variation in spectra observed clinically, with the severity of injury assessed at 6h (NICHD scoring) as follows: two had severe injury, six had moderate, two had mild, and one had no injury. Nine infants were scanned within the first week after birth (range 1-7 days), while one neonate with each of moderate and severe injury was scanned at one month (33 days and 36 days after birth respectively).

Each neonate underwent the PRESS protocol outlined in Section 4.2 and the proposed shortened STEAM protocol outlined in Section 5.5 (with the second TR extended to 5000ms), on a single 3T Philips Achieva scanner. The acquisition time for the STEAM protocol was 12 minutes.

MR spectroscopy data was first analysed in LCModel (Provencher, 2001) as described in Section 4.5 for the PRESS protocol, with the same basis set, optimised for the timings of the Philips implementation and with ideal RF pulses. This was then repeated in LCModel with the STEAM protocol, generating a basis set with the same 1Hz Gaussian lineshape in VeSPA-Simulate (Soher et al., 2011), and including

the same list of metabolites (with NAA, NAAG, Cho, Cr and PCr methyl peaks separated from the multiplet resonances in the same manner as the PRESS basis set). From the STEAM series, the same relaxivity corrections were applied to the metabolite signals (updating the T_1 recovery term with that outlined in Equation 2.17a-b), and water signals (again with a long T_2 component fixed at 500ms).

The level of agreement between PRESS and STEAM protocols was assessed using a Bland-Altman plot, and examined for systematic differences.

6.1.4 Results

The [tNAA] obtained in consecutive subjects is shown in Figure 6.1, with a Bland-Altman plot shown in Figure 6.2. There was a trend towards systematic overestimation of [tNAA] by 0.49 (SD 0.54) mmol/kg wet weight when using the STEAM protocol in comparison to that obtained with PRESS. Closer examination of Figure 6.2 shows a single neonate (Subject 9) in whom the difference between techniques is outside of the normal range, with STEAM overestimating [tNAA] by 47%. This appeared to be a consequence of pathologically increased Lip/MM signals in the region of 2ppm in STEAM spectra, which contaminated the tNAA resonance. Example fits to data from three representative subjects are shown in Figure 6.3, with the Lip/MM fitted components shown separately to NAA.

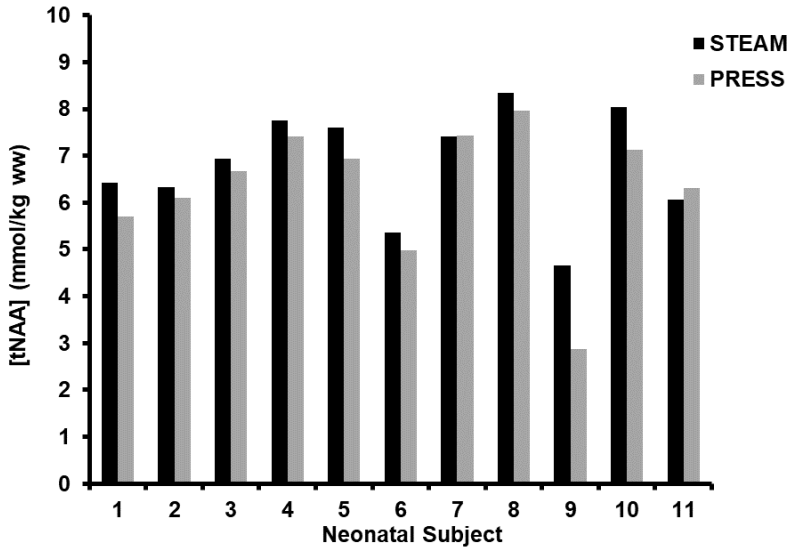

Figure 6.1: Estimated [tNAA] with both PRESS and STEAM protocols in 11 neonatal subjects, analysed with LCModel.

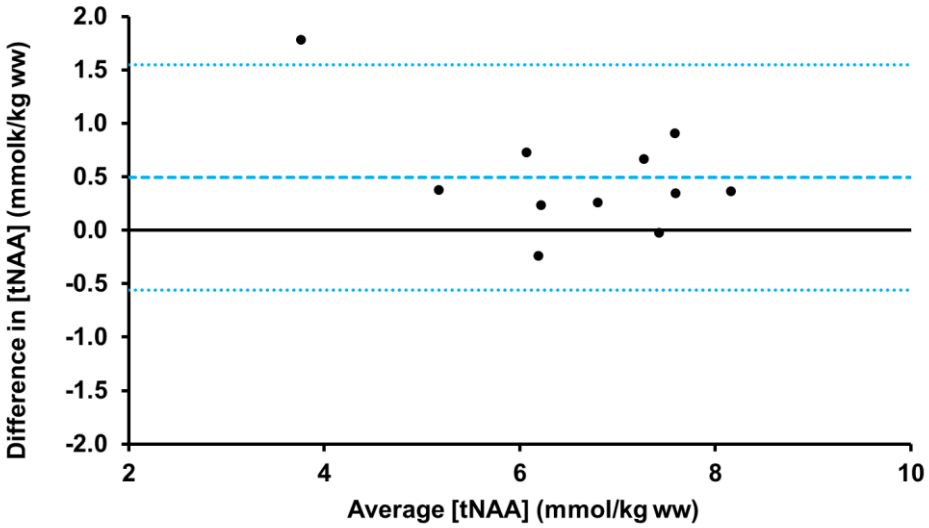

Figure 6.2: Bland-Altman plot of PRESS and STEAM [tNAA] results. Positive differences indicate STEAM > PRESS.

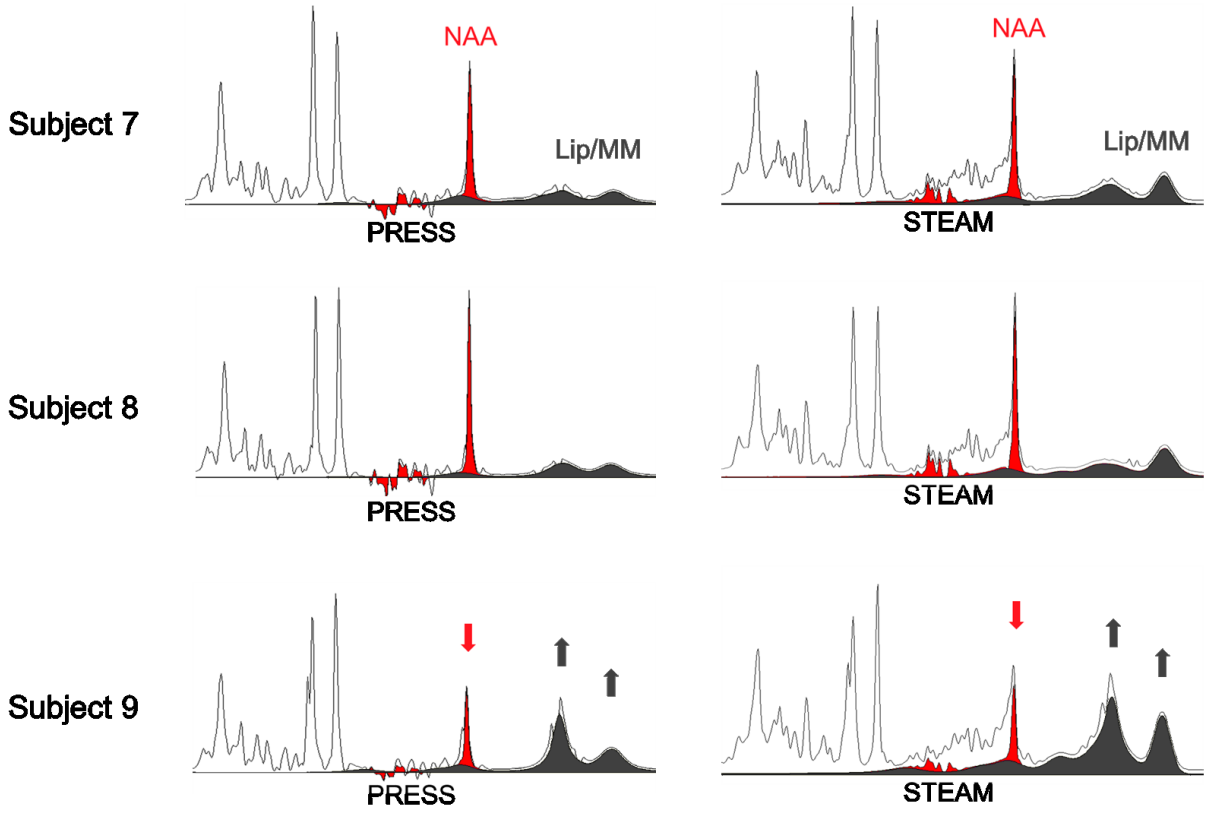

Figure 6.3: Representative spectral fits from three subjects. NAA fit components (red) are shown separate from Lip/MM fit components (grey). Arrows indicate apparent pathological changes in Subject 9, which appear to impact on NAA quantitation accuracy.

6.1.5 Discussion

Within normal variation and mild pathology, the PRESS and STEAM protocols are comparable in their measurements of [tNAA] in the neonatal brain, although STEAM tends to overestimate this.

In the case of substantial pathological changes (and the elevation of Lip/MM resonances), the STEAM protocol is susceptible to inaccuracies. Since various resonances have short T_2s, including those of Lip/MM, their contamination in PRESS spectra at TE=60ms is reduced in comparison to STEAM spectra at TE=20ms. Figure 6.3 demonstrates the simpler spectrum obtained when employing the former, particularly in the region of the tNAA resonances around 2ppm.

Due to the simplicity of the spectra obtained when using PRESS with TE=60ms, T_1 and T_2 relaxometry of the combined tNAA resonance is straightforward in the absence of intense overlapping peaks. In the STEAM protocol with TE=20ms, it is more difficult to separate tNAA from other contributions and so the two-point experiment is affected by additional systematic errors at each TR (effectively moving the experiment further down and to the right of the plots in Figure 5.8).

LCModel employs soft constraints on the relative amplitudes of Lip/MM resonances to better stabilise the fitting of spectra obtained at short TE, however these are likely inaccurate in a range of pathologies, and even healthy volunteers. For example, Marjańska et al. (2018) demonstrate substantial variation in Lip/MM signals at 1.0-2.2ppm between healthy younger adults (19-31 years of age) and healthy older adults (67-88 years of age), in the absence of changes in the 0.9ppm Lip/MM resonance.

In the proposed STEAM protocol, the Lip/MM resonances are fully relaxed due to T_1 effects throughout the experiment, as shown in Figure 6.4. The optimal analysis strategy would exploit this shared information, whereas LCModel can only treat these spectra as independent datasets.

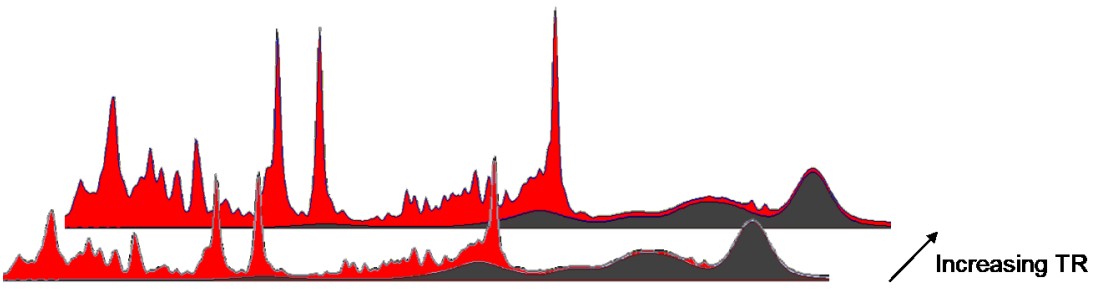

Figure 6.4: Illustration of the changes in spectral appearance with increasing TR. The metabolite signals (red) increase in intensity with TR, while the Lip/MM contributions are fully relaxed (grey) due to their short T_1 and so have equal intensity throughout. These differing behaviours could be exploited to improve the fit.

Rather than imposing (often inaccurate) soft constraints on amplitude ratios, weak constraints could instead be introduced by the experimental design (i.e. Lip/MM signals must be fully relaxed across both TRs and metabolite signals must increase with TR). This would improve the fitting process in two ways: 1) the two spectra are no longer considered independent measurements, and instead form part of a related dataset, avoiding unfeasible solutions (e.g. negative T_1s); 2) Lip/MM resonances can be more easily separated from short T_1 metabolites based on their known differences in relaxation behaviour, rather than on flawed assumptions about their relative abundance. The latter effect would then reduce bias in cases such as Subject 9 (Figure 6.3), where there are clear pathological changes in Lip/MM resonances.

This analysis framework has the potential to give significant advantages over the alternative strategies which are currently available, and will be explored in greater detail in the following sections.

6.1.6 Conclusion

The PRESS and STEAM protocols provide comparable results for [tNAA] in the neonatal brain during HIE, in the absence of severe pathology. However, in the case of prominent Lip/MM resonances, LCModel analysis with two independent fits (for the spectrum obtained at each TR) has poor accuracy. A more effective fitting strategy would exploit the information shared between the two spectra.

7 DEVELOPING AN OPTIMAL ANALYSIS STRATEGY

This chapter focuses on the development of an optimal analysis strategy for two-point saturation recovery datasets, comparisons with existing state-of-the-art alternatives, and characterisation of its inherent performance. Aspects of this work were presented at the ISMRM Annual Meeting 2017, and the Postgraduate Symposium of the British Chapter of ISMRM 2017.

7.1 DEMONSTRATING PROOF OF PRINCIPLE

7.1.1 Outline

Lip/MM resonances compromise the accuracy of short TE spectroscopy, particularly in the case of severe pathology (see Section 6), but also between healthy adults at different ages f and in different brain regions (Hofmann et al., 2001). While it is possible to obtain 'metabolite-nulled' spectra containing subject-specific Lip/MM contributions via various methods (Kassem and Bartha, 2003, Kunz et al., 2010, Penner et al., 2010) to improve the accuracy of short TE spectroscopy, these techniques are both time consuming and impractical for routine use on clinical MR scanners.

One such approach, employed by Hofmann et al. (2001) and Kreis et al. (2005), relies on a saturation recovery experiment to separate Lip/MM components from overlying metabolites based on the distinct differences in their T_1 s. Following excitation, Lip/MM resonances quickly reach their thermal equilibrium state due to their short T_1s, while metabolites are still partially saturated due to their longer T_1s. Using 16 different incremented TRs, these authors were able to demonstrate accurate separation between Lip/MM and metabolites based on the differences in their T_1 recovery curves, allowing accurate quantification. However, such an extensive protocol is impractical for routine clinical use.

Instead of performing such a protocol, a simpler saturation recovery experiment could be performed with just a pair of TRs. Provided both TRs are sufficiently long to ensure near fully recovered Lip/MM signals (i.e. TR≥1500ms), the biggest difference between the two spectra is then due to the T_1 recovery of the metabolites of interest, while Lip/MM resonances are near constant in intensity throughout (as illustrated in Figure 7.1). Importantly, no assumptions need to be made about the relative intensities of the individual Lip/MM resonances, as is usually the case with the analysis of MRS experiments at a single TR (e.g. LCModel soft constraints on intensity ratios). Meanwhile, other spectral features can be assumed constant throughout (i.e. chemical shift offsets, linewidths) to improve the stability of the fit. This could then simplify the calculation of metabolite concentrations as outlined in Figure 7.3.

Our hypothesis was that acquiring a simple pair of acquisitions, incremented in TR, could improve the accuracy of short TE spectroscopy in comparison to an acquisition made at a single TR. To test our

hypothesis, we implemented a method to simultaneously analyse two spectra, fitting the metabolite signal components to an ideal T_1 recovery curve. We then compared its performance to alternative strategies, either using independent fits to both spectra (see Figure 7.2), or a single spectrum at a fixed TR. The testing dataset comprised simulated spectra with varying signal-to-noise ratios and linewidths so that estimates could be compared to ground truth values.

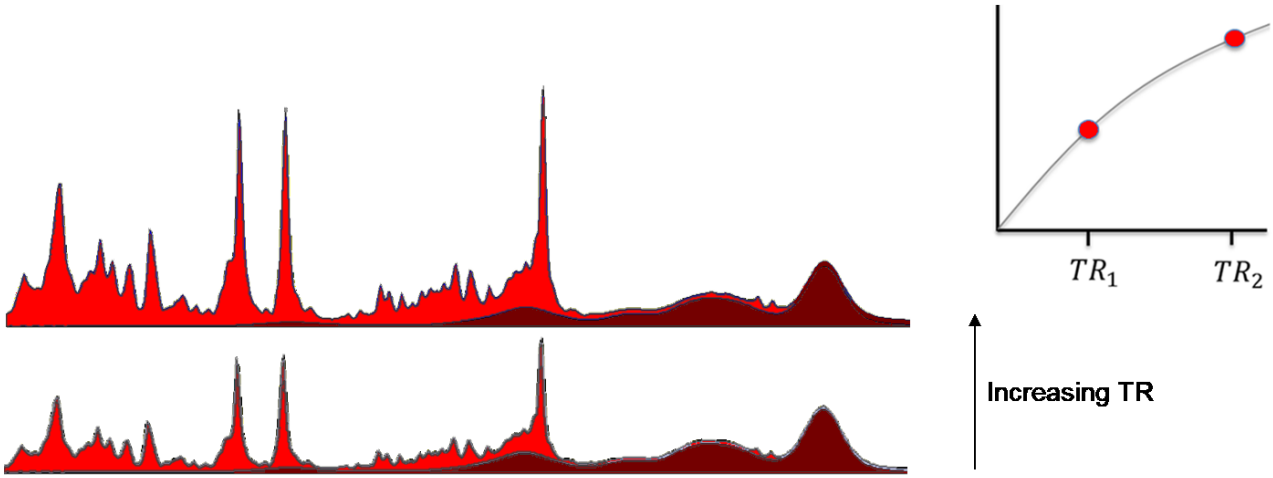

Figure 7.1: Illustration of the proposed combined fitting process. The model assumes that Lip/MM components (dark red) have a fixed contribution throughout while the metabolites of interest (light red) undergo recovery with T_1. Both spectra are fitted simultaneously according to the ideal relaxation recovery curve (top right) for each metabolite. This is the process used for [tNAA] calculation as outlined in Figure 7.3.

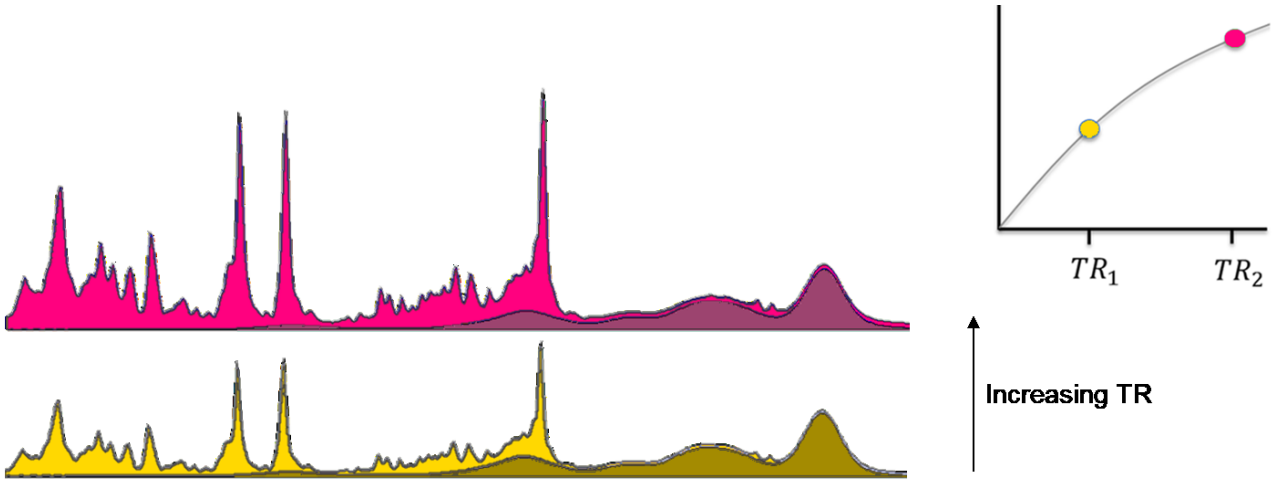

Figure 7.2: Illustration of the standard fitting process, where the two spectra are analysed independently. No fitting information is shared between spectra, and so separate estimates are made for Lip/MM components (dark yellow/magenta) and metabolites of interest (light yellow/magenta) in each spectrum. This strategy may give unfeasible estimates (e.g. T_1<0) when the signals for each metabolite are then fitted to an ideal recovery curve (top right). This is the process used to determine tNAA T_1 in the calculation outlined in Figure 5.2.

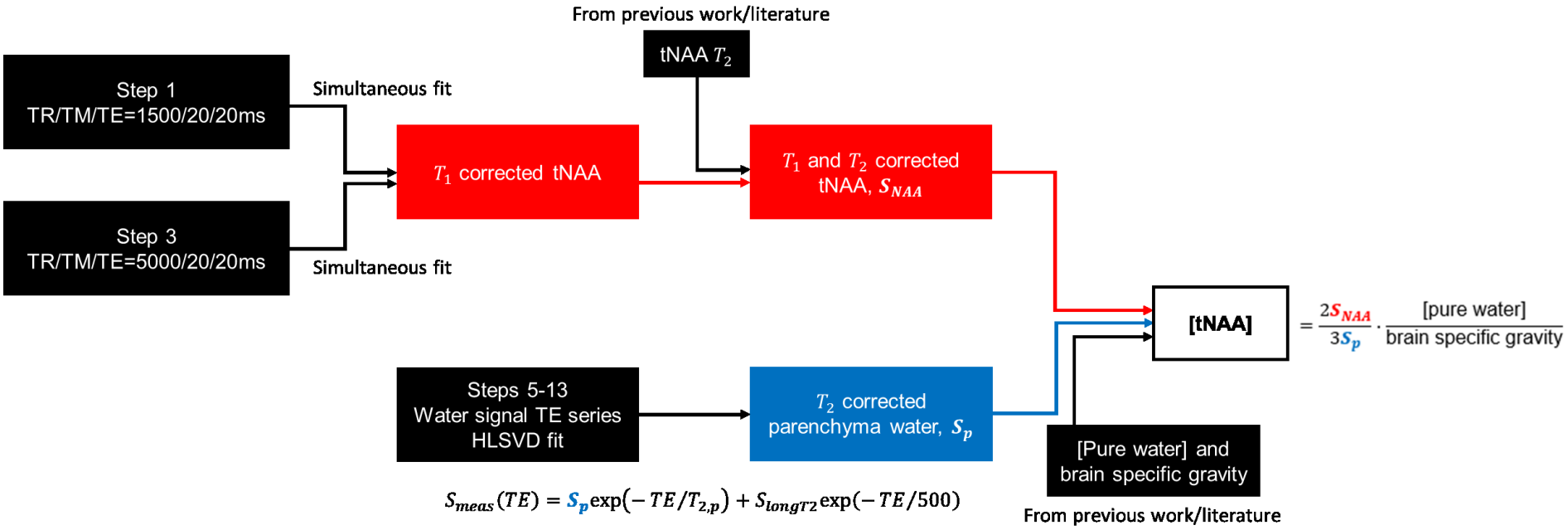

Figure 7.3: Calculation steps to derive thalamic tNAA concentration, [tNAA], using a simultaneous fit to STEAM spectra recorded in the saturation recovery experiment. Black boxes show the input data, either as raw data, HLSVD fits or reference values from previous work/literature. Red boxes and text denote calculated quantities from the tNAA resonance, and blue boxes and text denote calculated quantities from the water signal. Equations accompany the different calculation steps. The relaxation corrected intensities for tNAA and parenchyma water given by S_{NAA} and S_p respectively.

7.1.2 The TARQUIN signal model for a single spectrum

Spectral fitting in the time domain can be summarised as a least squares optimisation problem as follows:

$$\underset{\mathbf{a},\boldsymbol{x}}{\operatorname{argmin}} \sum_{t} \left(\boldsymbol{y}(t) - \hat{\boldsymbol{y}}(\mathbf{a}, \boldsymbol{x}, t)\right)^2 \tag{7.1}$$

where $\boldsymbol{y}(t)$ is a column vector of the measured time domain signal, and $\hat{\boldsymbol{y}}(\mathbf{a}, \boldsymbol{x}, t)$ is a column vector of the signal estimate generated with a combination of linear parameters, $\mathbf{a}$, and non-linear model parameters, $\boldsymbol{x}$. In TARQUIN (Wilson et al., 2011), the signal model is as follows:

$$\hat{\boldsymbol{y}}(t) = \sum_{k=1}^{N} \mathrm{a}_k \boldsymbol{s}_k(\boldsymbol{\theta}, t) \exp\left((j\omega_k - \alpha_k - \beta t)t\right) \tag{7.2}$$

$$\mathbf{a} = \{\mathrm{a}_1, \ldots, \mathrm{a}_k, \ldots, \mathrm{a}_N\}, \qquad \boldsymbol{x} = \{\boldsymbol{\theta}, \omega_1, \ldots, \omega_N, \alpha_1, \ldots, \alpha_N, \beta\},$$

where a_k is the amplitude of the signal from the kth of N metabolites, represented by a basis vector in the time domain $\boldsymbol{s}_k(t)$ which is phased by $\boldsymbol{\theta}$, offset by a frequency shift ω_k, and damped by a combination of Lorentzian (α_k) and Gaussian (β) functions to form a Voigt lineshape. The Gaussian component β is applied equally to all metabolite basis vectors to account for the common effect of $\boldsymbol{B_0}$ inhomogeneity, while the metabolite-specific Lorentzian term α_k accounts for variation in individual T_2s. The phasing term is also applied equally to each basis vector, and accounts for both zeroth and first order phase adjustments.

Equation 7.2 can be expressed more succinctly by grouping the non-linear parameters into a single column vector $\boldsymbol{\psi}_k$ as follows:

$$\hat{\boldsymbol{y}}(t) = \sum_{k=1}^{N} \mathrm{a}_k \, \boldsymbol{\psi}_k(t) = \boldsymbol{\psi}\, \mathbf{a} \tag{7.3}$$

where $\boldsymbol{\psi}$ is a matrix with N columns and the same number of rows as timepoints in $\boldsymbol{y}$. The separability of $\boldsymbol{\psi}$ and $\mathbf{a}$ in the signal model can be exploited using the variable projection method (VARPRO, Golub and Pereyra (1973)). Briefly, this makes use of the useful property that $\mathbf{a}$ can be expressed in terms of $\boldsymbol{\psi}$ as follows:

$$\mathbf{a} = \boldsymbol{\psi}^{\dagger}\, \boldsymbol{y} \tag{7.4}$$

where † denotes the matrix pseudoinverse. Due to this property, the optimisation can be performed over the non-linear parameters in $\boldsymbol{\psi}$ only, with the corresponding values of $\mathbf{a}$ calculated after each iteration. In TARQUIN, this optimisation is performed using the Levenberg-Marquardt algorithm.

7.1.3 Hard and soft constraints in TARQUIN

Since many of the resonances overlap in an *in-vivo* spectrum, particularly at short echo times, TARQUIN employs soft constraints which encourage likely combinations of signal amplitudes and increase fitting stability. For example, NAAG is expected to be observed with 0.15 of the signal of NAA, and ratio of the MM signals at 2.0ppm and 0.9ppm is expected to be 1.5.

Hard constraints are then applied to the non-linear parameters so that they can only be adjusted within a specific range (e.g. zeroth order phase differences are between -π and π).

7.1.4 An adapted TARQUIN model

We modified TARQUIN (v4.3.9, Wilson et al. (2011)) source code so that it accepted two related spectra as inputs, and analysed them simultaneously by fitting metabolite signals to an ideal saturation recovery curve. Briefly, the time signals at each TR were appended into a single column vector, $\begin{bmatrix} \boldsymbol{y}_{TR1}(t) \\ \boldsymbol{y}_{TR2}(t) \end{bmatrix}$, which was approximated by extending each column of the TARQUIN signal model, $\begin{bmatrix} \cdots & \boldsymbol{\psi}_{TR1,k}(t) & \cdots \\ \cdots & \boldsymbol{\psi}_{TR2,k}(t) & \cdots \end{bmatrix}$, as follows:

$$\begin{bmatrix} \boldsymbol{y}_{TR1}(t) \\ \boldsymbol{y}_{TR2}(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\psi}_{TR1,1}(t) & \cdots & \boldsymbol{\psi}_{TR1,k}(t) & \cdots & \boldsymbol{\psi}_{TR1,N}(t) \\ \boldsymbol{\psi}_{TR2,1}(t) & \cdots & \boldsymbol{\psi}_{TR2,k}(t) & \cdots & \boldsymbol{\psi}_{TR2,N}(t) \end{bmatrix} \mathbf{a} \tag{7.5}$$

where $\boldsymbol{\psi}_{TR,k}(t)$ is a column vector representing the scaled basis signal for the kth metabolite at a repetition time of TR, modified by damping, frequency shifting and phasing terms.

The scaled basis vectors at each TR were related by an ideal STEAM saturation recovery curve:

$$\boldsymbol{\psi}_{TR,k}(t) = \left(1 - \exp\left(-\frac{TR - TM - \mathrm{TE}/2}{T_{1,k}}\right)\right) \cdot \boldsymbol{\psi}_{TR\infty,k} \tag{7.6}$$

where $T_{1,k}$ is the T_1 of the kth metabolite which is estimated during the fitting, and $\boldsymbol{\psi}_{TR\infty,k}$ is the scaled basis vector for the kth metabolite recorded with $TR \gg T_1$. This equation assumes that the spectra are recorded in the steady state, using a perfect 90° excitation pulse, and that $TM \ll TR$.

In this adapted model, the signal amplitude for each metabolite at the two TRs is constrained to lie on the ideal recovery curve (Equation 7.6). During the fitting process, the allowable range of T_1 for each resonance is subjected to the following hard constraints to ensure acceptable relaxation behaviour:

$$\begin{cases} 300\text{ms} \leq T_1 \leq 4000\text{ms} & \text{for metabolite signals} \\ T_1 = 300\text{ms} & \text{for Lip/MM signals} \end{cases}$$

This has the effect of ensuring all fitted Lip/MM signals are fully relaxed in both spectra (provided $TR1, TR2 \geq$ 1500ms) while other metabolites were free to display varying degrees of partial saturation (as illustrated in Figure 7.1).

During fitting, the two spectra $\boldsymbol{y}_{TR1}(t)$ and $\boldsymbol{y}_{TR2}(t)$ are aligned with a global frequency shift, and other fitting parameters from the original TARQUIN model (i.e. individual frequency shifts, individual phase shifts, Voigt lineshape terms) are applied equally to $\boldsymbol{\psi}_{TR1}$ and $\boldsymbol{\psi}_{TR2}$. All soft constraints between relative signal amplitudes are relaxed.

7.1.5 Validation study

To compare the performance of the proposed simultaneous analysis strategy with the alternatives, a set of simulated datasets was generated and fitted. Metabolite signal estimates could then be compared to known ground truth values under a range of conditions (i.e. various linewidths and noise levels).

7.1.6 Methods: Simulated acquisition

A metabolite basis set was simulated in VeSPA-Simulate (Soher et al., 2011), for a 3T STEAM acquisition with TR=∞, TM=20ms, TE=20ms and ideal RF pulses (see Appendix IV, Figure A4.1).

The datasets for the validation study were derived from the signals in this basis set, with the addition of varying degrees of Gaussian noise (signal-to-noise ratio of 14,24, or 68) and the application of varying degrees of Voigt line broadening (3Hz, 5Hz or 7Hz).

From these broadened and noisy versions of the basis set, an acquisition at a given TR was simulated by applying separate partial saturation factors to the metabolite signals (all having an equal, long T_1) and the Lip/MM signals (all having T_1=300ms). In this validation experiment, two TRs were fixed at 1500ms and 5000ms, and the metabolite T_1s uniformly incremented from 400ms to 3000ms.

The validation dataset therefore comprised simulated spectra at TR=1500ms and TR=5000ms across three different linewidths, three different signal-to-noise ratios, and 27 different metabolite T_1s. Each of these spectra was generated 20 times in a Monte Carlo simulation. The ground truth abundances of the metabolites and Lip/MMs were chosen to approximate typical neonatal brain spectra, in ratios which agreed with the soft constraints employed in both LCModel (Provencher, 2001) and TARQUIN (Wilson et al., 2011). Representative simulated spectra are shown in Figure 7.4.

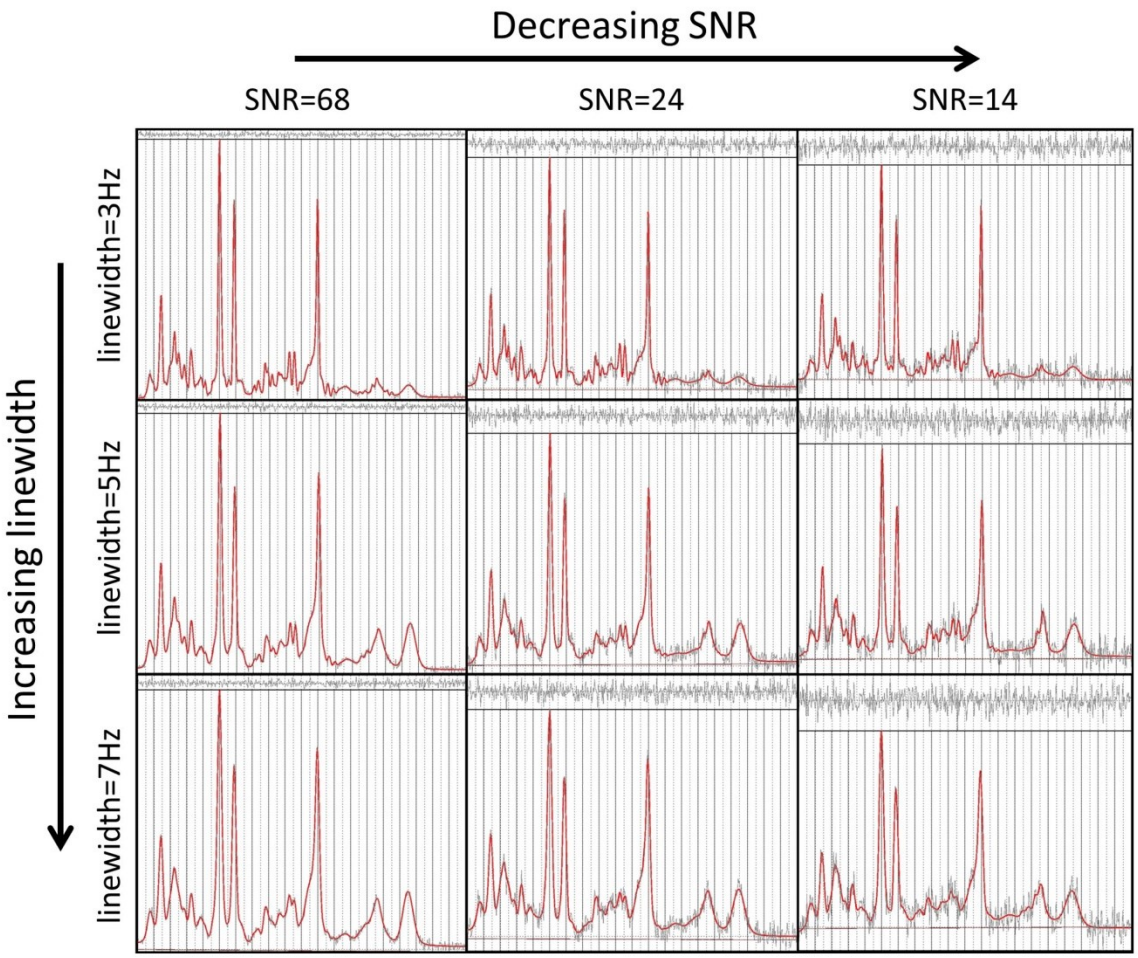

Figure 7.4: Typical STEAM spectra (TR=5000ms) simulated for the experiment, with LCModel fits in red.

7.1.7 Methods: Analysis

Analysis was performed using each of the following strategies, all using the same basis set for the fitting as was used to simulate the acquisitions:

1. Each simulated TR=5000ms spectrum was fitted using LCModel, and a fixed T_1 correction was applied equally to all metabolites (T_1=1700ms assumed throughout).
2. For each simulated TR pair, the two spectra were fitted independently using LCModel (with its default soft constraints on relative signal amplitudes), and the signal amplitudes for each metabolite fitted post-hoc to the ideal saturation recovery curve.
3. For each simulated TR pair, the two spectra were fitted independently using TARQUIN (with its default soft constraints on relative signal amplitudes), and the signal amplitudes for each metabolite fitted post-hoc to the ideal saturation recovery curve.
4. For each simulated TR pair, both spectra were analysed simultaneously according to the adapted TARQUIN method described in Section 7.1.4.

In steps 2 and 3, highly covarying metabolites were grouped together (e.g. tNAA=NAA+NAAG) before fitting the sum to the saturation recovery curve. In step 4 this was not necessary, as each component is assigned its own T_1 during the fitting process.

7.1.8 Results

The fitted metabolite signals, normalised to the ground truth values, are shown plotted against the true simulated metabolite T_1s in Figure 7.5 for different signal-to-noise ratios (with a fixed linewidth of 3Hz). The equivalent results for varying linewidths are shown in Figure 7.7 (at the highest simulated signal-to-noise ratio). Results derived from TARQUIN signal models are shown in blue, while those derived from LCModel are shown in red.

For each metabolite, simultaneously fitting both spectra with the adapted TARQUIN model gave estimates closer to the ground truth values than all other methods (i.e. consistently clustered around unity in Figure 7.5). Fitting a spectrum from a single TR of 5000ms gave a characteristic bias, depending on the difference between the actual T_1 of the metabolite and the assumed T_1 for the relaxation correction factor. This was consistent across the signal-to-noise ratios and linewidths studied.

Performing independent fits of the two spectra with either LCModel or TARQUIN, followed by a calculated T_1 correction, reduced this T_1-dependent bias. However, both overestimated the relaxation corrected tNAA signal by approximately 5-10%. This corresponded to an underestimation of the underlying Lip/MM signals at 2.0ppm by 5-10% (Figure 7.6).

Performing a simultaneous fit to both spectra with the adapted TARQUIN model, all estimates of metabolite signals were closest to their ground truth values in the range of 800ms $< T_1 <$ 2200ms. At lower T_1s (i.e. T_1 < 800ms), the metabolites are difficult to discriminate from the Lip/MMs and so the estimates become more biased.

As noise was increased, the fit to a single spectrum in LCModel is largely unaffected (left column in Figure 7.5), but the variance increases when using independent fits to the two spectra (two middle columns in Figure 7.5). This is likely due to increased variance in the post-hoc T_1 estimate, as the variance increases at long T_1s where the correction factor is largest, and therefore has the largest impact on the final signal estimate. This is particularly troublesome for Glx, which is most heavily affected by increased noise.

Provided the signal-to-noise ratio is sufficiently high, an increased spectral linewidth has a minimal impact on the estimated signal intensities (see Figure 7.7).

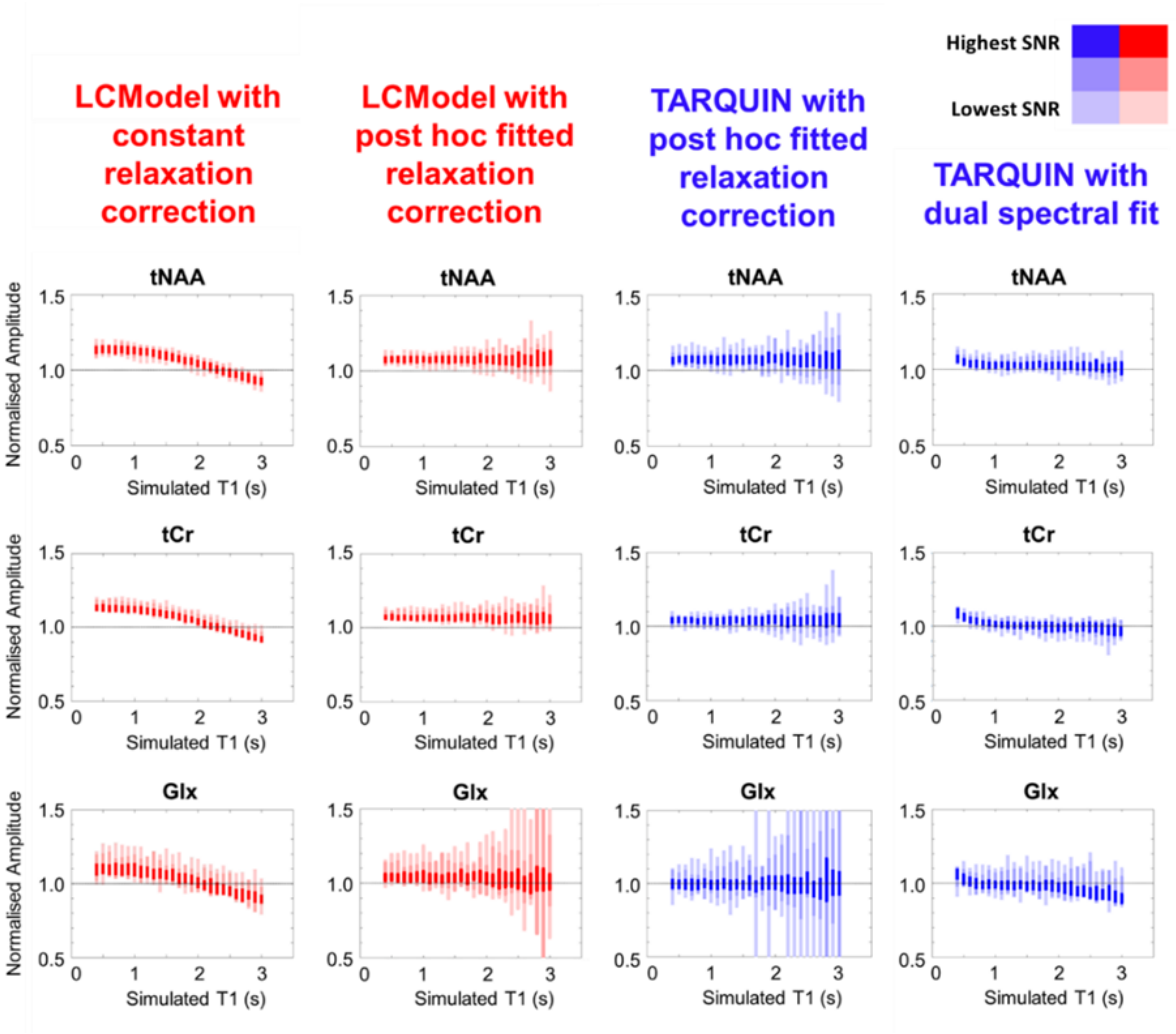

Figure 7.5: Mean normalised amplitude from 20 simulations +/- 1 standard deviation. Deeper colours correspond to higher SNR at a fixed linewidth of 3Hz (top row of Figure 7.4). LCModel derived results are shown in red, TARQUIN derived results are shown in blue.

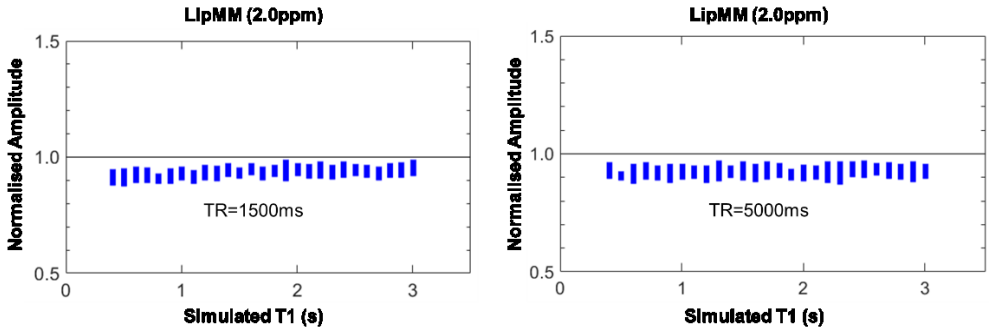

Figure 7.6: Mean normalised amplitude of the lipid and macromolecule resonances (at 2.0ppm) from 20 simulations +/- 1 standard deviation. The results are for individual TARQUIN fits to the spectrum at each TR (SNR=68).

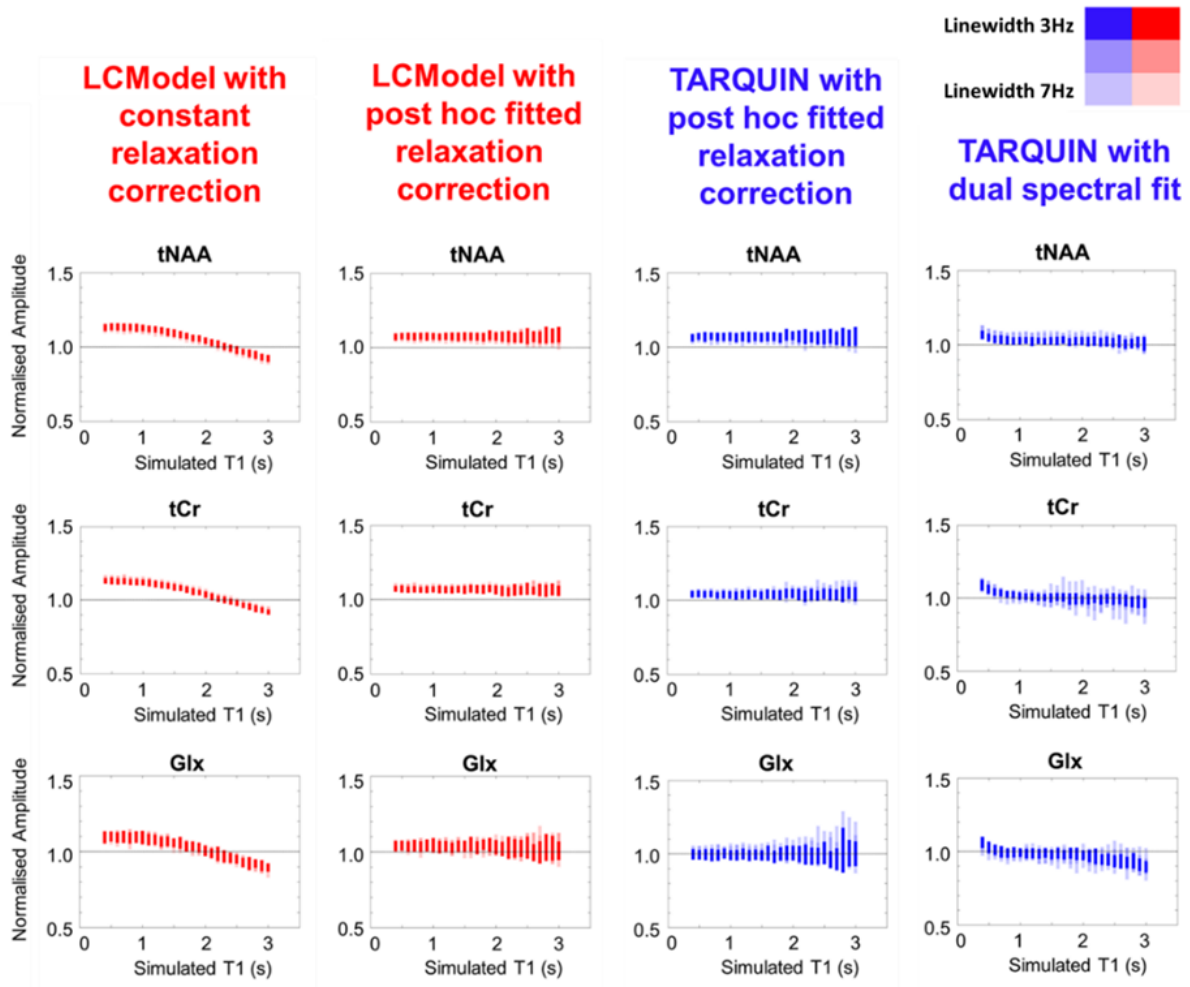

Figure 7.7: Mean normalised amplitude from 20 simulations +/- 1 standard deviation. Deeper colours correspond to sharper linewidth at the highest SNR (left column of Figure 7.4). LCModel derived results are shown in red, TARQUIN derived results are shown in blue.

7.1.9 Discussion

Fitting a single spectrum results in a characteristic bias in estimated metabolite signal amplitudes which depends on the discrepancy between the estimated and true T_1s. However, even where the assumed T_1 matches the true T_1 (1700ms), tNAA, tCr and Glx are all overestimated by approximately 10%, with Glx becoming more biased with decreasing signal-to-noise ratio. These biases are present despite the metabolite signals being almost fully recovered at TR=5000ms, with only a 5% loss due to partial saturation. These effects will therefore be accentuated where a single TR of <5000ms is employed (as is almost always the case in *in-vivo* experiments).

Even with a basis set that completely matches the data, and with the correct T_1 correction term applied, there is a bias in the estimated metabolite concentrations when using an experiment at a single TR. The accompanying underestimation of Lip/MM signals at 2.0ppm suggests that this is due to imperfect separation of the metabolites of interest from broader underlying signals.

Applying a post-hoc T_1 correction using LCModel or TARQUIN (central two columns in Figure 7.5 and Figure 7.7) removes the characteristic T_1 bias evident in the single TR strategy. However, there is a consistent bias best illustrated by the overestimation of tNAA across all linewidths and signal-to-noise ratios, which is present when using both LCModel and TARQUIN. This corresponds to an underestimation of the underlying Lip/MM signals. Although these estimates are equally precise in the presence of linewidth variation, lower signal-to-noise ratios introduce more variance. This decrease in precision is most pronounced for Glx rather than tNAA or tCr, most likely due the low intensity multiplet patterns of Glu and Gln.

Using the adapted TARQUIN model and analysing both spectra simultaneously, the systematic biases are reduced to near zero over a physiologically plausible T_1 range (800ms $< T_1 <$ 2200ms) for all of tNAA, tCr and Glx. At shorter T_1s (i.e. $<$ 800ms) it is difficult to separate the metabolite components from Lip/MM components, and so the former is likely contaminated by the latter, resulting in overestimation of the signal amplitudes. As the signal-to-noise ratio is reduced and the linewidth is increased, the variance in the estimated signals increases, but the mean remains near zero. This increased variance is still reduced compared to the post-hoc LCModel and TARQUIN strategies, particularly for Glx.

7.1.10 Conclusion

When estimating metabolite concentrations, using a two-point saturation recovery experiment reduces the bias associated with using a single long TR. In addition, fitting the two spectra simultaneously improves the robustness of the fit compared to fitting them independently.

7.2 Developing an optimal analysis strategy

Although the quantification of metabolite signals can be improved by using a simultaneous fitting approach, some aspects of the adapted TARQUIN model are not optimal for this type of problem.

7.2.1 TARQUIN macromolecular baseline

Broad baseline signals are not explicitly modelled in TARQUIN, but their influence on the fitting can be adjusted by varying the analysis window of the time domain signal. User defined starting points (t_s) and end points (t_e) define the time window included in the least-squares optimisation as follows:

$$\underset{\mathsf{a},\boldsymbol{x}}{\operatorname{argmin}} \sum_{t_s \le t \le t_e} \left(\boldsymbol{y}(t) - \hat{\boldsymbol{y}}(\mathsf{a}, \boldsymbol{x}, t)\right)^2 \tag{7.7}$$

Extending t_s excludes more of the broad baseline terms which decay rapidly in the time domain. Lowering t_e also allows the user to reduce the influence on the fit from later points in the signal, where

noise dominates. While a heuristic approach can be taken to optimise the start and end points, the results are inevitably biased by this choice.

If t_s is chosen too early in the FID, the signal used for metabolite quantification will be contaminated by residual lipid and macromolecule signals and broad distortions. For example, in the case of residual broad resonances in the region of 2.0ppm, these are ascribed to metabolites such as NAA and NAAG.

Alternatively, if t_s is moved to a later point in the FID, the signal used for metabolite quantification must be back-extrapolated further to estimate the signal intensity (at t=0), resulting in greater uncertainty. This is especially problematic in the case of non-Voigt lineshapes (due to e.g. eddy currents), which are not well approximated by a classical FID intensity-time profile, and where back-extrapolation will be biased.

Lipid and macromolecule T_2s are of the order of tens of milliseconds (Opstad et al., 2008b) and so $T_2{}^*$ decay will occur over a shorter timescale. Inspecting truncated lipid and macromolecule dominated spectra, broad resonances decay into the noise by approximately 50ms due to $T_2{}^*$, depending on the quality of shimming. Any cut-off before this point will therefore be contaminated by the broad Lip/MM signals, while cut-off points beyond this point will discard a substantial amount of useful data from the initial points of the FID.

7.2.2 Alternative: AQSES macromolecular baseline

An alternative is to model the broad baseline in the frequency domain with a non-parametric model, such as a spline function (as employed by the AQSES (Automated Quantification of Short TE MRS) fitting strategy), in combination with a parametric model which describes distinct resonances. In comparison to the TARQUIN baseline, no data is lost from the initial points of the FID in the time domain. However, a method must also be applied to determine the optimal degree of smoothness for the splines. Insufficient smoothness will cause overfitting, while strongly constrained splines will fail to explain sufficient variation in the spectrum.

7.2.3 Time domain vs frequency domain fitting

TARQUIN performs its fitting in the time domain, which allows it to exploit the fact that the FID decays into the noise at later times, performing a 'weighted' fit on the most informative region of the decaying signal. However, it is sensitive to distortions from residual unsuppressed water signals, and the weighted fit must be adjusted to account for the degree of inhomogeneity in the voxel (and hence the rate of signal decay).

Frequency domain fitting, on the other hand, exploits the limited frequency range covered by metabolites of interest. This ensures that fitting residuals only correspond to the region of interest, and

the most prominent distortions from residual water are ignored (although there will still be broad residual baseline effects).

The two representations of the signal are linked by a simple discrete Fourier transformation, and so the most effective post-processing and analysis strategies make the most of both.

7.2.4 Efficiency

Few analysis strategies are designed to work with arrays of interrelated spectra. While, for example FiTAID (Chong et al., 2011) can perform this task, it does not allow incorporation of the efficient VARPRO method, and can only characterise the baseline/Lip/MM components via a combination of Voigt lineshapes. Therefore, a novel solution was designed and built to enable efficient optimisation, best suited to the specific features of two-point saturation recovery experiments.

7.3 PROPOSED FITTING STRATEGY

The proposed fitting strategy was developed in MATLAB (2017a), with the key features outlined in the following sections.

7.3.1 Signal model

The proposed signal model is best descried in comparison with the original TARQUIN model, i.e.:

$$\hat{\boldsymbol{y}}(\mathbf{a}, \boldsymbol{x}, t) = \sum_{k=1}^{N} \mathrm{a}_k \boldsymbol{s}_k(\boldsymbol{\theta}, t) \exp\big((j\omega_k - \alpha_k - \beta t)t\big) \tag{7.8}$$

$$\mathbf{a} = \{\mathrm{a}_1, \dots, \mathrm{a}_k, \dots, \mathrm{a}_N\}, \qquad \boldsymbol{x} = \{\boldsymbol{\theta}, \omega_1, \dots, \omega_N, \alpha_1, \dots, \alpha_N, \beta\},$$

where a_k is the amplitude of the signal from the kth of N metabolites, represented by a basis vector in the time domain $\boldsymbol{s}_k(t)$ which is phased by $\boldsymbol{\theta}$, offset by a frequency shift ω_k, and damped by a combination of Lorentzian (α_k) and Gaussian (β) functions to form a Voigt lineshape.

In the proposed model, the phasing of the basis vectors is performed directly in the time domain as follows:

$$\boldsymbol{s}_k(\boldsymbol{\theta}, t) = \boldsymbol{s}_k(0, t') \exp(j\theta) \tag{7.9a}$$

where

$$t' = t + d \tag{7.9b}$$

and d is a delay term for the start of the FID. This delay causes a first order phase shift, while the Euler term in Equation 7.9a causes a zeroth order phase shift. This basis vector is then scaled according to the ideal partial saturation factor for the STEAM sequence:

$$\boldsymbol{s}_{k,TR}(\boldsymbol{\theta}, t, T_{1,k}) = \left(1 - \exp\left(-\frac{TR - TM - \mathrm{TE}/2}{T_{1,k}}\right)\right) \cdot \boldsymbol{s}_k(\boldsymbol{\theta}, t) \tag{7.10}$$

where TR, TM and TE are the repetition time, mixing time and echo time respectively, and $T_{1,k}$ is the T_1 of the kth metabolite.

Finally, the modelled signal is Fourier transformed ($\mathcal{F}$) so that the least squares optimisation occurs in the frequency domain:

$$\hat{\boldsymbol{y}}_{TR}(\mathsf{a}, \boldsymbol{x}, f) = \mathcal{F}\left(\sum_{k=1}^{N} \mathsf{a}_k \boldsymbol{s}_{k,TR}(\boldsymbol{\theta}, t, T_{1,k}) \exp\big((j\omega_k - \alpha_k - \beta t')t'\big)\right) \tag{7.11}$$

$$\mathsf{a} = \{\mathsf{a}_1, \dots, \mathsf{a}_N\}, \qquad \boldsymbol{x} = \{\theta, T_{1,1}, \dots, T_{1,N}, \omega_1, \dots, \omega_N, \alpha_1, \dots, \alpha_N, \beta, d\}.$$

The optimisation problem can then be performed for spectra from two different TRs (TR1 and TR2) over a discrete frequency range:

$$\underset{\mathsf{a},\boldsymbol{x}}{\operatorname{argmin}} \sum_{f_L \le f \le f_R} \left(\begin{bmatrix} \boldsymbol{y}_{TR1}(f) \\ \boldsymbol{y}_{TR2}(f) \end{bmatrix} - \begin{bmatrix} \hat{\boldsymbol{y}}_{TR1}(\mathsf{a}, \boldsymbol{x}, f) \\ \hat{\boldsymbol{y}}_{TR2}(\mathsf{a}, \boldsymbol{x}, f) \end{bmatrix}\right)^2 \tag{7.12}$$

where f_L and f_R are the left and right frequency limits (e.g. from f_L=4.2ppm to f_R=0.0ppm). Here the column vectors representing the observed signal ($\boldsymbol{y}_{TR}$) and the model signal ($\hat{\boldsymbol{y}}_{TR}$) at each TR have been appended to make single column vectors containing the signal from both TRs.

7.3.2 Spline baseline

A baseline is added to the signal model in Equation 7.12, and is implemented as proposed by Sima and Van Huffel (2007), using regularised B-splines. For these splines, a penalty term $\boldsymbol{p}$ increases in value as the spline deviates from a smooth function. By adding this penalty term to the cost function for the overall fit, the fitting favours smoother baselines with a smaller $\boldsymbol{p}$:

$$\underset{\mathsf{a},\boldsymbol{c},\boldsymbol{x}}{\operatorname{argmin}} \quad \frac{1}{2n}\left(\sum_{f_L \le f \le f_R} \left(\begin{bmatrix} \boldsymbol{y}_{TR1} \\ \boldsymbol{y}_{TR2} \end{bmatrix} - \begin{bmatrix} \hat{\boldsymbol{y}}_{TR1}(\mathsf{a}, \boldsymbol{c}, \boldsymbol{x}) \\ \hat{\boldsymbol{y}}_{TR2}(\mathsf{a}, \boldsymbol{c}, \boldsymbol{x}) \end{bmatrix}\right)^2\right) + 2p(\boldsymbol{c}) \tag{7.13}$$

where $\boldsymbol{n}$ indicates the number of data points in the frequency range $f_L \le f \le f_R$ for each spectrum, $\boldsymbol{c}$ are the (complex) B-spline coefficients, and the signal model at each TR $\hat{\boldsymbol{y}}_{TR}$ now incorporates the B-spline baseline. Assuming an equal baseline at both TRs, the same spline is used for both spectra.

The penalty term $\boldsymbol{p}$ is calculated as follows:

$$p(\boldsymbol{c}) = \lambda^2 \boldsymbol{c}^H \boldsymbol{D}^H \boldsymbol{D} \boldsymbol{c} \tag{7.14}$$

where H denotes Hermitian conjugation, λ is a fixed user-defined term which determines the strength of the penalty, and $\boldsymbol{D}$ is a second order differential operator:

$$\boldsymbol{D} = \begin{bmatrix} -1 & 2 & -1 & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & -1 & 2 & -1 \end{bmatrix} \tag{7.15}$$

As λ is increased, the penalty term in Equation 7.14 increases, favouring the use of smoother splines as illustrated in Figure 7.8. This strategy is employed to model the unknown smooth baseline in various popular fitting tools, including LCModel (Provencher, 2001) and AQSES (Poullet et al., 2007).

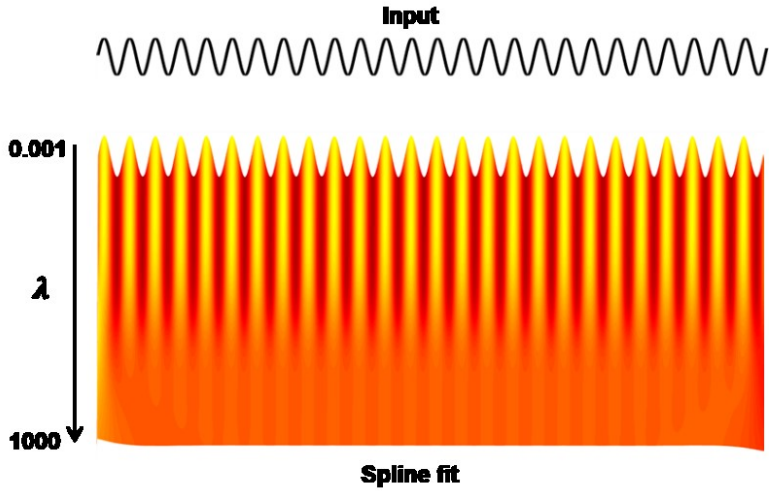

Figure 7.8: 2D surface plot displaying the effect of varying the regularisation parameter, λ, on the spline fit to an ideal sinusoidal function. The y axis is on a log_{10} scale, and red indicates lower values while yellow indicates higher values. Higher regularisation increases the smoothness of the fit.

7.3.3 Formulation of the fitting problem

The fitting problem is a least squares minimisation of the difference between the observed data series, and the non-linear signal model as follows:

$$\underset{\boldsymbol{a},\boldsymbol{x},\boldsymbol{c}}{\operatorname{argmin}} \quad \frac{1}{2n}\left(\sum_{f_L \le f \le f_R}\left(\begin{bmatrix}\boldsymbol{y}_{TR1}\\ \boldsymbol{y}_{TR2}\end{bmatrix} - \begin{bmatrix}\boldsymbol{\psi}_{TR1}(\boldsymbol{x})\\ \boldsymbol{\psi}_{TR2}(\boldsymbol{x})\end{bmatrix}\mathbf{a} - \begin{bmatrix}\mathcal{B}\boldsymbol{c}\\ \mathcal{B}\boldsymbol{c}\end{bmatrix}\right)^2\right) + 2p(\boldsymbol{c}) \tag{7.16}$$

where $\mathcal{B}$ denotes the B-spline matrix, and we have separated the model into linear ($\mathbf{a}$,$\boldsymbol{c}$) and non-linear ($\boldsymbol{x}$) terms:

$$\begin{bmatrix}\hat{\boldsymbol{y}}_{TR1}(\mathbf{a},\boldsymbol{x},\boldsymbol{c})\\ \hat{\boldsymbol{y}}_{TR2}(\mathbf{a},\boldsymbol{x},\boldsymbol{c})\end{bmatrix} \equiv \begin{bmatrix}\boldsymbol{\psi}_{TR1}(\boldsymbol{x})\\ \boldsymbol{\psi}_{TR2}(\boldsymbol{x})\end{bmatrix}\mathbf{a} + \begin{bmatrix}\mathcal{B}\boldsymbol{c}\\ \mathcal{B}\boldsymbol{c}\end{bmatrix} \tag{7.17}$$

$$\mathbf{a} = \{\mathrm{a}_1, \ldots, \mathrm{a}_N\}, \boldsymbol{c} = \{c_1, \ldots, c_M\}, \boldsymbol{x} = \{\theta, d, T_{1,1}, \ldots, T_{1,N}, \omega_1, \ldots, \omega_N, \alpha_1, \ldots, \alpha_N, \beta\}.$$

As all of the parameters in $\boldsymbol{a}$ are $\boldsymbol{c}$ are linear, they can be expressed in terms of just the non-linear parameters $\boldsymbol{x}$, as proposed by Sima and Van Huffel (2007). This then becomes a non-negative least squares problem, which can be solved with the $\boldsymbol{x}$ generated at each iteration of an outer non-linear least squares problem, via the variable projection algorithm (VARPRO, Golub and Pereyra (1973)), similar to that outlined in Section 7.1.2. The cost function and Jacobian for the non-linear least squares problem are then approximated as described by Sima and Van Huffel (2007).

7.3.4 Soft constraints

Soft constraints are imposed on the non-negative least squares minimisation to encourage convergence with metabolite amplitude ratios close to those expected, via the 'weighting method' proposed by Wilson et al. (2011). The minimisation problem is extended to include additional terms as follows:

$$\begin{bmatrix} \boldsymbol{\Psi}_1 & \cdots & \boldsymbol{\Psi}_u & \boldsymbol{\Psi}_v & \cdots & \boldsymbol{\Psi}_N \\ 0 & \cdots & \epsilon\|\boldsymbol{\Psi}_u\|_1{}^{(k)} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \epsilon\|\boldsymbol{\Psi}_v\|_1{}^{(k)} & \cdots & 0 \end{bmatrix} \boldsymbol{a} \approx \begin{bmatrix} \boldsymbol{y} \\ \frac{r_u\left(a_u{}^{(k-1)} + a_v{}^{(k-1)}\right)}{r_u + r_v}\epsilon\|\boldsymbol{\Psi}_v\|_1{}^{(k)} \\ \frac{r_v\left(a_u{}^{(k-1)} + a_v{}^{(k-1)}\right)}{r_u + r_v}\epsilon\|\boldsymbol{\Psi}_v\|_1{}^{(k)} \end{bmatrix} \quad (7.18)$$

where ϵ determines the strength of the soft constraint, $a_u{}^{(k)}$ is the amplitude estimate of metabolite u in the kth iteration, $\frac{r_u}{r_v}$ is the expected ratio of metabolite signal amplitudes, $\boldsymbol{y}$ is a vector of the combined measured spectra, and $\boldsymbol{\Psi}_u$ the full signal model (including baseline terms) for metabolite u.

7.3.5 Hard constraints

Hard constraints are imposed on the non-linear least squares minimisation (outer loop) in the form of box constraints, and the non-negative least squares minimisation (inner loop) ensures that all signal amplitudes are greater than or equal to zero.

7.3.6 Variance estimation

Estimates of the minimum variance (assuming an unbiased estimation) are expressed as Cramér Rao Lower Bounds (CRLBs). As suggested by Zhang and Shen (2014), the Jacobian for the full signal model (including the baseline terms) is computed by augmenting the analytical terms (describing the parametric components of the metabolite signal model) with a numerical estimate (describing the baseline effect; 'jacobianest' v1.6, John D'Errico, MATLAB File Exchange):

$$\mathbb{J}_{TR} = \frac{\partial}{\partial \boldsymbol{z}}\boldsymbol{\psi}_{TR} + \frac{\partial}{\partial \boldsymbol{z}}\mathcal{B}\boldsymbol{c} \quad (7.19)$$

where $\boldsymbol{z}$ describes all linear and non-linear parameters in the signal model. This numerical calculation accounts for an increase in computation time, but is necessary in order to fully account for the interaction between the baseline and parametric terms. The CRLB can then be calculated using Equations 2.23a-b.

7.3.7 Implementation details

The proposed fitting strategy was developed in MATLAB (2017a), with the MATLAB implementation of the 'levmar' Levenberg-Marquardt non-linear least squares algorithm which handles user-defined box constraints (http://users.ics.forth.gr/~lourakis/levmar/). The inner non-negative least squares algorithm was performed with the MATLAB (2017a) lsqnonneg algorithm, which was modified to include the additional terms applying soft constraints (as in Equation 7.18).

To ensure that the problem is well scaled, parameters entering the non-linear least squared problem are transformed as follows:

$$x_{scaled} = \frac{(x_{unscaled} - \bar{x})}{\sigma_x} + 1 \tag{7.20}$$

where $\bar{x}$ is the expected value of the parameter and σ_x is its assumed standard deviation (both defined by the user *a priori*).

Following convergence, the fitted parameters are then rescaled to their original units by inverting this transformation. In addition, experimental data are normalised in the time domain prior to fitting, according to the maximum absolute signal across the two spectra, and rescaled following the fitting process to return parameters in their correctly scaled units.

Frequency adjustments are performed in a pre-processing step, based on the autocorrelation of each spectrum in the frequency domain with a reference signal (comprising an ideal NAA peak), and then the autocorrelation of the two experimental spectra. The separable linear and non-linear parameters in the least squares problem are then estimated using the VARPRO algorithm outlined previously.

7.3.8 Hyperparameters

In this signal model, a key hyperparameter to optimise is the regularisation term λ, which controls baseline smoothness. In other contexts, the smoothness of the baseline has been shown to significantly impact on the estimated spectral linewidth, and in turn, the accuracy of quantitation (Tkac et al., 2002, Kreis and Boesch, 2003, Mangia et al., 2006, Bartha, 2007). In following sections, I investigate the consequences of baseline smoothness on the characteristics of the fit in the proposed signal model. In addition, I examine the impact of ϵ, which determines the strength of soft constraints on the metabolite signal ratios.

7.4 *IN-VIVO* VALIDATION

7.4.1 Outline

It is important to determine whether the proposed strategy is accurate *in-vivo*, and results in robust quantification in the presence of normal and pathological variations in the neonatal brain. The STEAM data from Section 6 were therefore re-fit using the technique outlined in Section 7.3, and compared to the results from the PRESS spectra. Since the latter were largely uncontaminated by broad baseline signals, these were considered as ground truth [tNAA] values.

7.4.2 Methods

The STEAM spectra from the same 11 subjects in Section 6 were fitted with the proposed signal model, with the same basis set, and estimates of the [tNAA] compared to the assumed ground truth.

To investigate the impact of λ on the spectral fit, this was incremented from 10^{-2} to 10^{3} in 20 steps on a $\log_{10}$ scale. To investigate the importance of ϵ values, these were first set to the default values used in TARQUIN (v4.2.1, Wilson et al. (2011)) for the first set of analyses. The soft constraints used are shown in Table 7.1. These were originally developed to suit the adult brain, and so these may not be representative of neonatal spectra which are subject to both pathological and developmental changes. Therefore the analyses were then repeated with all ϵ set to zero (i.e. no soft constraints).

Ratio	Expected value	ϵ
NAAG/NAA	0.15	0.05
Lip09/Lip13	0.267	0.5
Lip20/Lip13	0.15	0.3
MM20/MM09	1.5	0.2
MM12/MM09	0.3	0.9
MM14/MM09	0.75	0.3
MM17/MM09	0.375	0.8

Table 7.1: Amplitude ratio soft constraints implemented in TARQUIN.

Following this, a more extensive parameterisation of the MM resonances was included, splitting them into 9 distinct Gaussian peaks as defined by Považan et al. (2018):

Resonance	Frequency (ppm)	Linewidth (ppm)
MM01	0.91	0.14
MM02	1.21	0.15
MM03	1.43	0.17
MM04	1.67	0.15
MM05	2.04	0.15
MM06	2.26	0.20
MM07	2.99	0.15
MM08	3.21	0.25
MM09	3.77	0.39

Table 7.2: Parameterisation of the macromolecule peaks with nine Gaussian components.

Linewidths were determined by comparison with Provencher's LCModel implementation (Provencher, 2001), or by inspecting Považan's metabolite-nulled spectra. All Lip resonances were kept the same as Provencher's LCModel implementation:

Resonance	Frequency (ppm)	Linewidth (ppm)
Lip09	0.89	0.14
Lip13a	1.28	0.15
Lip13b	1.28	0.89
Lip20a	2.04	0.15
Lip20b	2.25	0.15
Lip20c	2.80	0.20

Table 7.3: Lipid parameterisation employed in LCModel.

Due to the constraints imposed by the experimental conditions, the additional freedom in MM could be allowed in the absence of any soft constraints (i.e. all ϵ equal to zero). This would otherwise lead to unacceptable instability in the fit to a dataset acquired at a single TR.

7.4.3 Results

The impact of λ on fully relaxed metabolite signal estimates is shown in Figure 7.9, using Provencher's MM parameterisation either with or without soft constraints. As λ is reduced in the absence of soft constraints, the additional flexibility results in substantial variations in fitted signal intensity.

There is substantial variation in the fitted tNAA amplitude as λ is varied and with soft constraints applied to relative signal amplitudes (top left panel of Figure 7.9), which is not predictable in nature. Whereas a more flexible baseline gives larger tNAA in some cases, in others this results in lower tNAA than when λ is strongly constrained. Removing soft constraints (top right panel of Figure 7.9) causes substantial differences in the signal attributed to tNAA, with these results being less dependent on the strength of λ used. The imposition of soft constraints also had a substantial effect on the signals attributed to tCr, tCho and Glx (Figure 7.9).

The amplitudes of the fitted metabolite signals all show variation as λ is varied. However, the effect is more pronounced on low intensity multiplet resonances, such as those of Glx (Figure 7.9). The consistently lower amplitude at low λ can be explained by mixing with the more flexible baseline.

In the case of a heavily smoothed baseline (λ=1000) where the results reach an asymptote, the [tNAA] estimates from each model are shown compared to the assumed ground truth values in Figure 7.10. The best agreement was observed when applying Provencher's MM parameterisation to the STEAM data, in the absence of any soft constraints on the relative metabolite signal amplitudes. The accompanying Bland-Altman plot for this case is shown in Figure 7.11. Importantly, the fitting is also robust in the neonate (Subject 9) with pathologically low [tNAA] and high Lip/MM content.

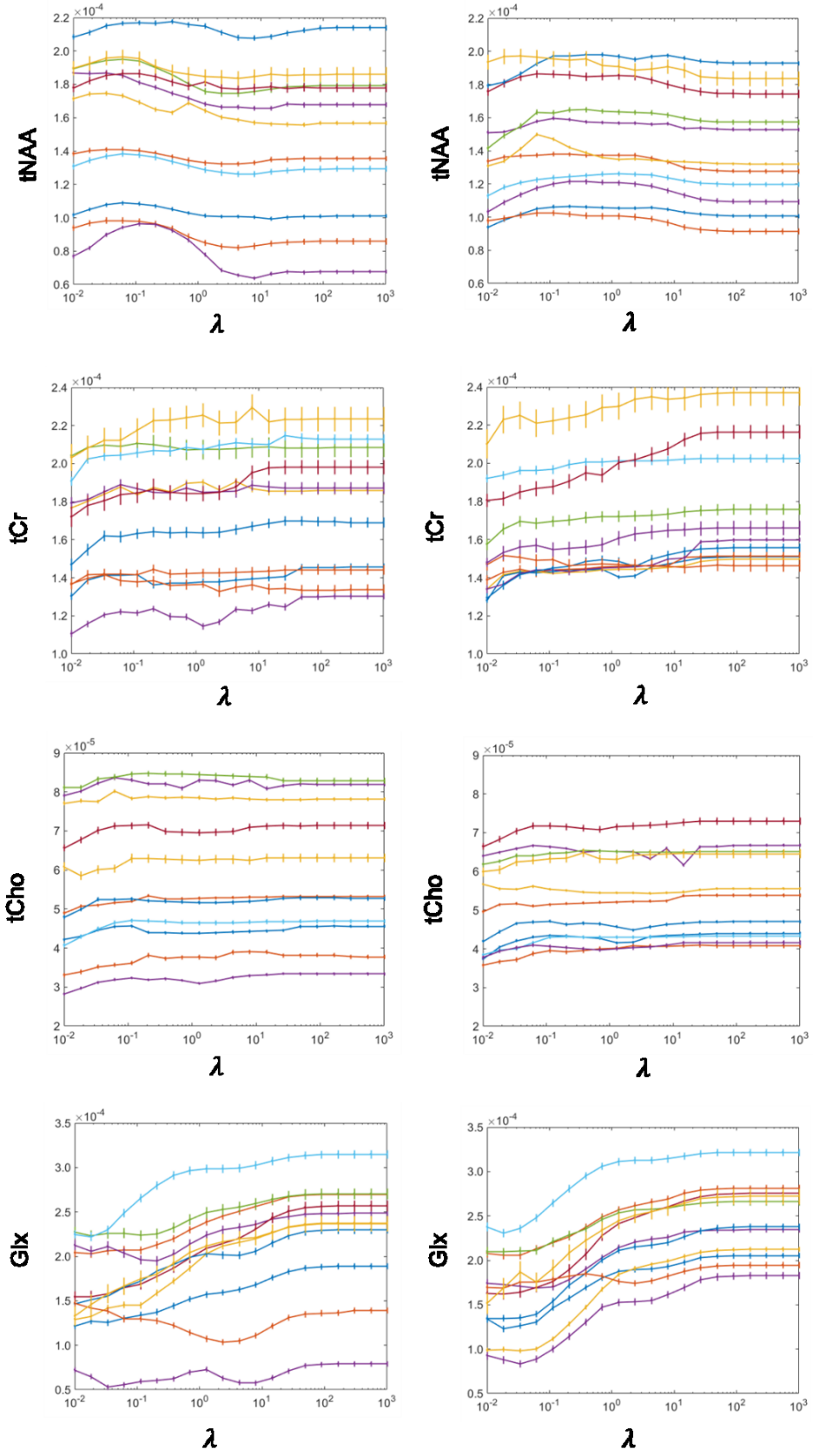

Figure 7.9: Fitted fully relaxed signals for (from top to bottom row) tNAA, tCr, Cho and Glx, when using Provencher's MM parameterisation either with soft constraints on amplitude ratios (left column) or without the soft constraints (right column). Results are shown on the same scale, with the signal on the y-axis in arbitrary units, along with CRLBs (error bars).

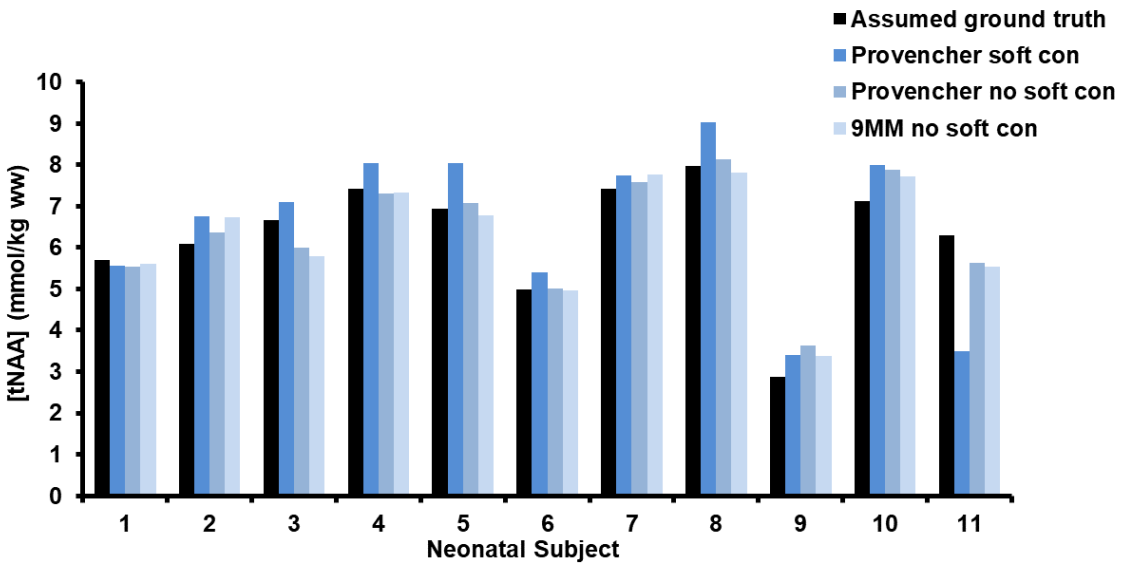

Figure 7.10: Estimated [tNAA] using the STEAM protocol data and proposed fitting method in comparison to the assumed ground truth (black) values from the PRESS acquisition in Section 6. Provencher (no) soft con = Provencher's MM model with (without) soft constraints on signal amplitude ratios; 9MM no soft con = nine MM model without soft constraints on signal amplitude ratios.

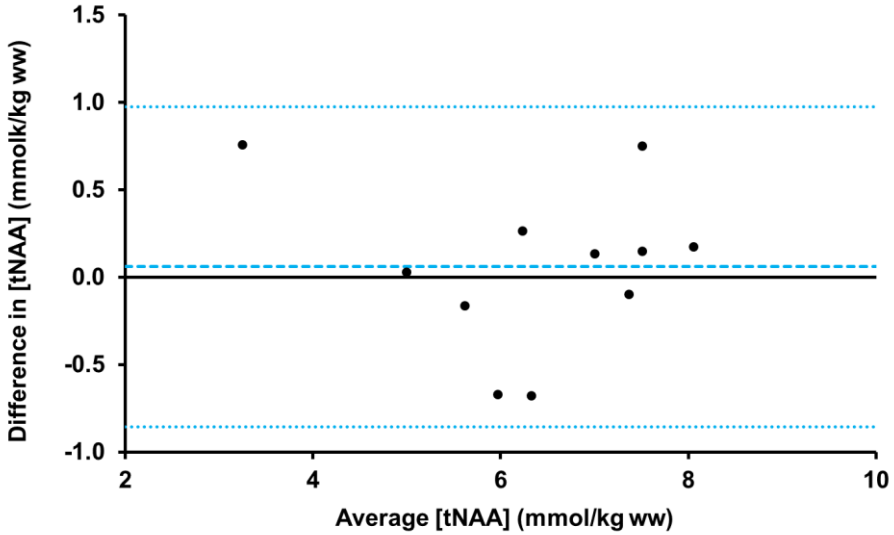

Figure 7.11: Bland-Altman plot illustrating the difference between estimated [tNAA] and the assumed ground truth values from the PRESS acquisition in Section 6. This corresponds to the Provencher MM parameterisation, without soft constraints.

The mean correlation matrix across all 11 subjects is shown in Figure 7.12 along with its standard deviation, with λ fixed at 1000 and using Provencher's MM parameterisation. The covariance between Lip/MM components of the fit was reduced in comparison to the that shown in Figure 4.13, and there was less interaction between these signals and tNAA.

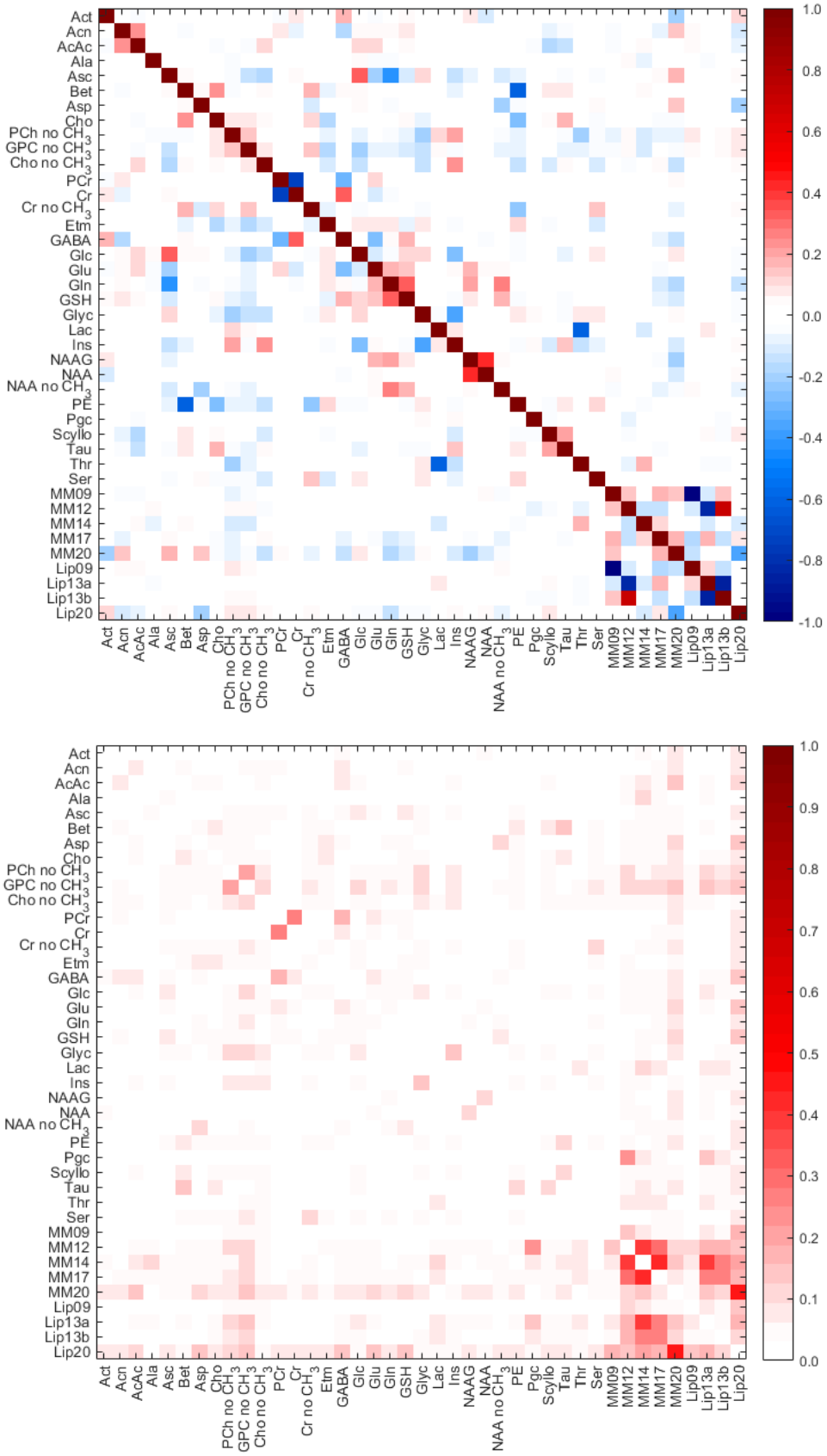

Figure 7.12: Mean (top) and standard deviation (bottom) of the correlation matrices across all 11 subjects, using Provencher's MM parameterisation and λ=1000, with no soft constraints on the signal amplitudes. 'no CH_3' indicates multiplet resonances fitted independently from the metabolite's singlet peak.

Initial results indicate that, without supplying any prior information about expected Lip/MM resonances, this component can be reliably extracted during the fitting process to further improve the method (a patent application is in preparation). A representative fit is shown in Figure 7.13. With this method applied to the data from all 11 subjects, the Bland-Altman plot in Figure 7.14 indicates less variance in comparison with any of the parameterised Lip/MM strategies.

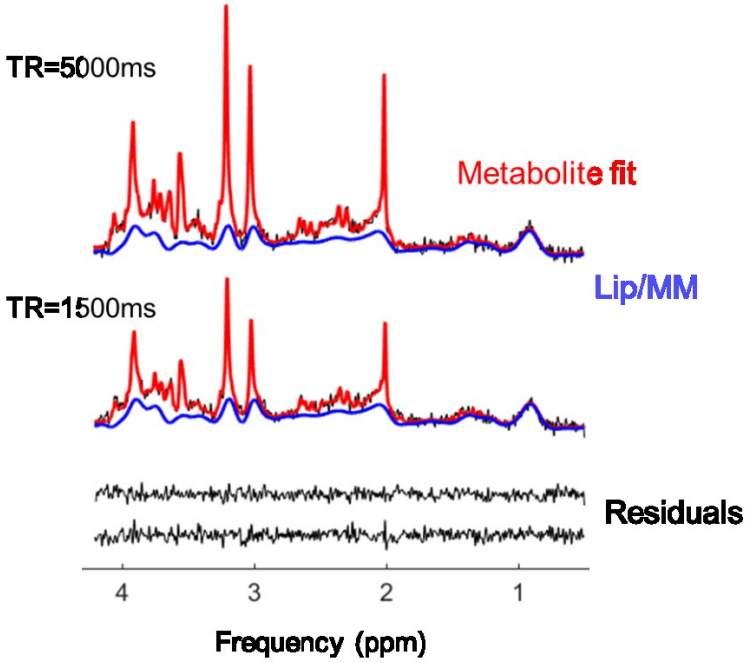

Figure 7.13: Representative spectral fit for the STEAM dataset in one of the subjects, with Lip/MM components (blue) extracted without prior knowledge. Note the resemblance of the Lip/MM component with metabolite-nulled spectra presented elsewhere (Schaller et al., 2014, Snoussi et al., 2015, Terpstra et al., 2016, Giapitzakis et al., 2018).

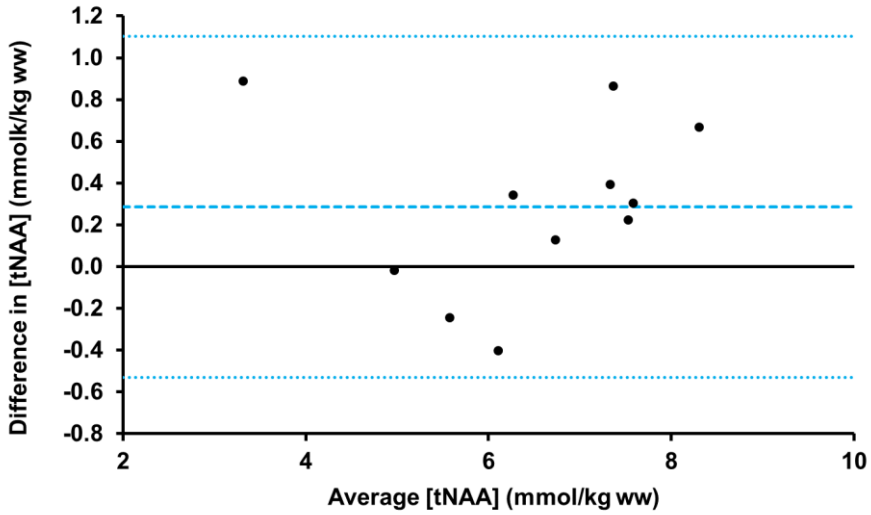

Figure 7.14: Bland-Altman plot illustrating the difference between estimated [tNAA] and the assumed ground truth values from the PRESS acquisition in Section 6. Here the Lip/MM components have been estimated without prior knowledge.

7.4.4 Discussion

The proposed fitting method results in [tNAA] estimates which are in agreement with the assumed ground truth values, even in the presence of substantial metabolic variation across neonatal subjects.

As expected, there is a strong dependence of the fitted metabolite amplitudes on the strength of the regularisation on the spline baseline, λ. However, in the heavily smoothed regime (i.e. $\lambda \sim 1000$) where results approach an asymptote, the estimated [tNAA] closely matches the assumed ground truth.

The Lip/MM component which can be automatically extracted in the absence of a defined model (Figure 7.13) is similar in appearance to many of the metabolite-nulled spectra presented by other investigators in other contexts (Schaller et al., 2014, Snoussi et al., 2015, Terpstra et al., 2016, Giapitzakis et al., 2018). The close approximation to ground truth [tNAA] is particularly promising, as it suggests that fitting could be made robust to variable Lip/MM characteristics. Future work should compare this directly to metabolite-nulled spectra in the same subjects. If this proves to be sufficiently accurate, this would negate the requirement for lengthy acquisitions of metabolite-nulled spectra, as the full dataset can be collected in one short protocol. Furthermore, the strategy is insensitive to the localisation method used and so can be incorporated into many existing MRS protocols.

Different Lip/MM parameterisations can also be easily accommodated into the fit, without the need for additional soft constraints on relative amplitudes. In fact, such soft constraints should be avoided, as these make *a priori* assumptions which may be inaccurate, particularly in the case of pathology. The improved characterisation of the Lip/MM peaks with the proposed strategy can be appreciated by comparing the correlation matrices in Figure 7.12 and Figure 4.13. In particular, the signal attributed to NAA in the proposed strategy is less influenced by underlying MM20 signals.

7.4.5 Limitations

There are some important limitations in the fitting model, which if overcome could improve the methodology. Firstly, the lineshape term does not account for deviations from an ideal Voigt profile. This can be readily incorporated via multiplication (time domain) or convolution (frequency domain) with a series of delta functions along with a regularisation constraint. In the work presented here, all spectra were corrected using the method proposed by Klose (1990) to minimise eddy current artefacts. However, this does not account for transient changes between the acquisition of the water suppressed and unsuppressed signals.

Secondly, all metabolite basis sets have been expanded so that the methyl singlet peaks are separated from multiplet resonances. This reduces the number of degrees of freedom in the model, but is necessary due to the known variations in relaxivities of different moieties (Deelchand et al., 2018), and

anomalous J modulation effects (see Section 2.4.4). One method of stabilising the fit would be to add soft constraints such that the relative amplitudes of different resonances on the same metabolite are all equal. While this would not account for anomalous J modulation losses, measured RF pulse waveforms can be incorporated into various basis set simulation packages (Soher et al., 2011, Simpson et al., 2017) which could address this issue.

Another potential limitation of the model is the assumption of perfect 90° excitation pulses, such that the signal is completely saturated. This could be addressed with the addition of another term, ξ, in the expressions for partial saturation effects:

$$\eta_{TR} = 1 - \xi \exp(-TD/T_1) \tag{7.21}$$

However, a two-point experiment is insufficient to characterise this behaviour. At least one additional TR should be included in the experiment in this case.

Finally, the process of numerically determining correction terms in the Jacobian (Equation 7.19) is slow. However, in the case where the baseline is heavily smoothed, the problem becomes considerably simpler and the additional perturbations tend to zero. Considering the dependence of the metabolite signal estimates on λ, it would be advisable to use heavily smoothed baselines.

7.4.6 Conclusion

The proposed fitting strategy gives accurate results for [tNAA] in the neonatal brain during HIE, in agreement with the assumed ground truth values, even in the presence of severe pathology.

8 DISCUSSION AND FUTURE WORK

In-vivo magnetic resonance spectroscopy has been behind many key discoveries over the last 35 years and has made an important clinical impact in many areas, not least in HIE with the development of therapeutic hypothermia. However, even though Cady et al. (1983) suggested MRS could provide prognostic value in HIE as early as 1983, this is far from routine practice in the NHS in 2018. This thesis examines the prognostic utility of MRS in a prospective multicentre study, and outlines a novel solution to obtaining accurate quantitative information within the constraints of a typical clinical setting.

8.1 THE PROMISE OF MRS IN HIE

This thesis first presents strong evidence in favour of performing thalamic MRS in infants with HIE. In the largest prospective study of its kind, [tNAA] explained 51% of the variance in motor scores at two-years of age, with the measurement obtained within two weeks of birth. In addition, it had an area under the receiver operating curve (AUC) of 0.99 in predicting death or moderate or severe disability at two years. However, one important limitation was the lengthy protocol (25 min), and the need for research privileges on some scanners to perform this successfully. As a result, only 41% of the infants in the study had metabolite concentrations for analysis. On the other hand, tLac/tNAA, which could be performed in all centres and had an acquisition time of approximately 6 minutes, was available for analysis in 85%, and had an AUC of 0.94.

Although changes in tLac/tNAA are less informative than [tNAA], its value provides a rapid and simple prognostic measure amenable to clinical settings. In clinical trials, the added predictive value of [tNAA] would increase the study power compared to tLac/tNAA when used as a surrogate outcome measure. The proposed STEAM protocol is currently being applied in a number of interventional trials, including the HELIX (Hypothermia for Encephalopathy in Low and Middle-Income Countries) Trial (n=408), a randomised controlled trial taking place in public sector hospitals in India, Sri Lanka and Bangladesh with access to 3T MR scanners (Thayyil et al., 2017). At the time of writing, this is on course to reach its recruitment target in December 2018. Although the trial is powered based on a primary outcome of death or moderate or severe disability at 18 months of age, all infants undergo thalamic [tNAA] measurements. As a result, there will be substantial power to investigate the importance of comorbidities such as infection on the efficacy of therapeutic hypothermia in HIE, using [tNAA] as a surrogate outcome.

8.2 INSIGHT INTO NEURONAL INJURY

The MRS results from the MARBLE study also provide some key insights into the evolution of injury in the first weeks after birth, as previous studies were conducted only in the initial days (Cady, 1996, Cheong et al., 2006). Thalamic [tNAA] demonstrated a linear dependence on gestational age at birth, in agreement with Kreis et al. (2002), though no dependence on postnatal age was observed. This suggests that the normal developmental changes are interrupted by the acute hypoxic-ischaemic insult. Meanwhile, the elevated [tCho] suggests ongoing inflammation in a long lasting tertiary phase of injury (Fleiss and Gressens, 2012). In addition, the absence of tNAA T_2 changes in more severe injury suggests acute neuronal swelling has subsided, although the increased parenchyma water T_2 indicates the onset of vasogenic oedema (Bednarek et al., 2012). The increased tCr T_2 in more severe injury suggests depletion of PCr in order to maintain ATP in the presence of persistent mitochondrial dysfunction, lasting for weeks.

8.3 TWO-POINT SATURATION RECOVERY STEAM

The final sections of this thesis outline a novel and simple technique for the accurate measurement of [tNAA] on any clinical 3T scanner with STEAM, lasting approximately 12 minutes (with the second TR extended to 5000ms). This overcomes the difficulties with separating key metabolite resonances from the broader underlying Lip/MM resonances, providing accurate and precise results even in the presence of severe pathological changes, and imposing fewer constraints based on *a priori* assumptions.

The disagreement between results from different expert investigators discussed by Bhogal et al. (2017) is largely attributed to the wide range of software used, and the large number of tuneable parameters. One way of addressing this problem is to routinely acquire richer datasets, as advocated here. The proposed method offers the potential to minimise these biases, as fewer constraints are made on the fitting model, and more information is available to constrain the fit based on simple relationships between spectra. Since a short TR spectrum can be acquired rapidly as an additional step in the protocol, it adds only a small additional time requirement, and the added benefit outweighs the time which would otherwise be spent on a single TR (see Section 5.7.4).

8.4 FUTURE WORK

The following sections are suggestions for further work, based on the work presented in this thesis.

8.4.1 Incorporation into semi-LASER acquisition schemes

As the semi-LASER pulse sequence becomes more widely available on clinical scanners, this will substantially reduce the chemical shift displacement artefacts and increase the signal-to-noise ratio available in routine clinical spectroscopy. The proposed method for improved quantification in this thesis is independent of the localisation scheme, and so this presents an excellent opportunity to improve the accuracy of clinical MRS even further, all within shorter acquisition times.

8.4.2 Implementation of MRS in other contexts

Thalamic MRS evidently has important prognostic value, however there must now be a concerted effort to realise this more widely, both to provide early prognostic information and to accelerate early phase clinical trials of novel putative neuroprotectants.

8.4.3 Acquisition of longitudinal data and further follow-up

It would be useful to investigate how an individual subject's metabolism changes in the first weeks after birth, both in HIE and in control subjects. This would give further information about the optimal timing of MRS studies, and the possibility of monitoring response to therapeutic interventions. In addition, the infants from the MARBLE study should be followed-up throughout childhood to examine the lasting impact of HIE on neurodevelopment.

8.4.4 Flexible lineshapes and imperfect RF excitation

Non-Voigt lineshapes can be readily incorporated into the fitting model as a series of frequency-separated delta functions, which would improve the robustness of the technique to lineshape distortions. In addition, a three-point fitting approach could be investigated to explore the impact of imperfect excitation pulses on the saturation recovery experiment.

8.4.5 Linking the behaviour of metabolite resonances

Soft constraints can be readily incorporated into the fitting model to encourage all groups on a given metabolite to have an equal abundance.

8.5 SUMMARY

The work presented in this thesis outlines the high prognostic accuracy of thalamic MR spectroscopy, performed within two weeks of birth, in predicting adverse neurodevelopmental outcomes years after hypoxic ischaemic encephalopathy. A simple method is then proposed to enable the accurate and precise measurement of [tNAA] within the constraints of a typical clinical environment, overcoming the biases associated with alternative techniques.

BIBLIOGRAPHY

ANCORA, G., SOFFRITTI, S., LODI, R., TONON, C., GRANDI, S., LOCATELLI, C., NARDI, L., BISACCHI, N., TESTA, C., TANI, G., AMBROSETTO, P. & FALDELLA, G. 2010. A combined a-EEG and MR spectroscopy study in term newborns with hypoxic–ischemic encephalopathy. *Brain and Development,* 32**,** 835-842.

ANCORA, G., TESTA, C., GRANDI, S., TONON, C., SBRAVATI, F., SAVINI, S., MANNERS, D. N., GRAMEGNA, L. L., TANI, G., MALUCELLI, E., CORVAGLIA, L. T., FALDELLA, G. & LODI, R. 2013. Prognostic value of brain proton MR spectroscopy and diffusion tensor imaging in newborns with hypoxic-ischemic encephalopathy treated by brain cooling. *Neuroradiology,* 55**,** 1017-1025.

ANDERSON, A. 2013. Ten years of maternity claims: an analysis of the NHS Litigation Authority data–key findings. *Clinical Risk,* 19**,** 24-31.

AZZOPARDI, D., ROBERTSON, N. J., BAINBRIDGE, A., CADY, E., CHARLES-EDWARDS, G., DEIERL, A., FAGIOLO, G., FRANKS, N. P., GRIFFITHS, J., HAJNAL, J., JUSZCZAK, E., KAPETANAKIS, B., LINSELL, L., MAZE, M., OMAR, O., STROHM, B., TUSOR, N. & EDWARDS, A. D. 2016. Moderate hypothermia within 6 h of birth plus inhaled xenon versus moderate hypothermia alone after birth asphyxia (TOBY-Xe): a proof-of-concept, open-label, randomised controlled trial. *The Lancet. Neurology,* 15**,** 145-153.

AZZOPARDI, D., STROHM, B., LINSELL, L., HOBSON, A., JUSZCZAK, E., KURINCZUK, J. J., BROCKLEHURST, P., EDWARDS, A. D. & REGISTER, U. T. C. 2012. Implementation and conduct of therapeutic hypothermia for perinatal asphyxial encephalopathy in the UK--analysis of national data. *PLoS One,* 7**,** e38504.

AZZOPARDI, D., WYATT, J. S., CADY, E. B., DELPY, D. T., BAUDIN, J., STEWART, A. L., HOPE, P. L., HAMILTON, P. A. & REYNOLDS, E. O. 1989. Prognosis of newborn infants with hypoxic-ischemic brain injury assessed by phosphorus magnetic resonance spectroscopy. *Pediatr Res,* 25**,** 445-51.

BADII, S., URIA-AVELLANAL, C., BAINBRIDGE, A., PRICE, D., DINAN, M., KENDALL, J. & ROBERTSON, N. 2014. PC.30 3T proton magnetic resonance spectroscopy in neonatal encephalopathy: lactate/n-acetylaspartate prognosis. *Archives of Disease in Childhood - Fetal and Neonatal Edition,* 99**,** A46-A46.

BAGORY, M., DURAND-DUBIEF, F., IBARROLA, D., CONFAVREUX, C. & SAPPEY-MARINIER, D. 2007. "Absolute" quantification in magnetic resonance spectroscopy: validation of a clinical protocol in multiple sclerosis. *Conf Proc IEEE Eng Med Biol Soc,* 2007**,** 3458-61.

BAINBRIDGE, A., TACHTSIDIS, I., FAULKNER, S. D., PRICE, D., ZHU, T., BAER, E., BROAD, K. D., THOMAS, D. L., CADY, E. B., ROBERTSON, N. J. & GOLAY, X. 2014. Brain mitochondrial oxidative metabolism during and after cerebral hypoxia–ischemia studied by simultaneous phosphorus magnetic-resonance and broadband near-infrared spectroscopy(). *Neuroimage,* 102**,** 173-183.

BARTHA, R. 2007. Effect of signal-to-noise ratio and spectral linewidth on metabolite quantification at 4 T. *NMR in Biomedicine,* 20**,** 512-521.

BARTHA, R., DROST, D. J., MENON, R. S. & WILLIAMSON, P. C. 2000. Spectroscopic lineshape correction by QUECC: Combined QUALITY deconvolution and eddy current correction. *Magnetic Resonance in Medicine,* 44**,** 641-645.

BEDNAREK, N., MATHUR, A., INDER, T., WILKINSON, J., NEIL, J. & SHIMONY, J. 2012. Impact of therapeutic hypothermia on MRI diffusion changes in neonatal encephalopathy. *Neurology,* 78**,** 1420-7.

BEDNAŘÍK, P., TKÁČ, I., GIOVE, F., DINUZZO, M., DEELCHAND, D. K., EMIR, U. E., EBERLY, L. E. & MANGIA, S. 2015. Neurochemical and BOLD responses during neuronal activation measured in the human visual cortex at 7 Tesla. *Journal of Cerebral Blood Flow & Metabolism,* 35**,** 601-610.

BEDNARIK, P., TKAC, I., GIOVE, F., EBERLY, L. E., DEELCHAND, D. K., BARRETO, F. R. & MANGIA, S. 2018. Neurochemical responses to chromatic and achromatic stimuli in the human visual cortex. *J Cereb Blood Flow Metab,* 38**,** 347-359.

BEHAR, K. L., DEN HOLLANDER, J. A., STROMSKI, M. E., OGINO, T., SHULMAN, R. G., PETROFF, O. A. & PRICHARD, J. W. 1983. High-resolution 1H nuclear magnetic resonance study of cerebral hypoxia in vivo. *Proc Natl Acad Sci U S A,* 80**,** 4945-8.

BHOGAL, A. A., SCHÜR, R. R., HOUTEPEN, L. C., VAN DE BANK, B., BOER, V. O., MARSMAN, A., BARKER, P. B., SCHEENEN, T. W. J., WIJNEN, J. P., VINKERS, C. H. & KLOMP, D. W. J. 2017. 1H–MRS processing parameters affect metabolite quantification: The urgent need for uniform and transparent standardization. *NMR in Biomedicine,* 30**,** e3804-n/a.

BIVARD, A., KRISHNAMURTHY, V., STANWELL, P., YASSI, N., SPRATT, N. J., NILSSON, M., LEVI, C. R., DAVIS, S. & PARSONS, M. W. 2014. Spectroscopy of reperfused tissue after stroke reveals heightened metabolism in patients with good clinical outcomes. *Journal of Cerebral Blood Flow & Metabolism,* 34**,** 1944-1950.

BLOCH, F., HANSEN, W. W. & PACKARD, M. 1946. The Nuclear Induction Experiment. *Physical Review,* 70**,** 474-485.

BLÜML, S., WISNOWSKI, J. L., NELSON, M. D., PAQUETTE, L., GILLES, F. H., KINNEY, H. C. & PANIGRAHY, A. 2013. Metabolic Maturation of the Human Brain From Birth Through Adolescence: Insights From In Vivo Magnetic Resonance Spectroscopy. *Cerebral Cortex (New York, NY),* 23**,** 2944-2955.

BOTTOMLEY, P. A. 1984. Selective volume method for performing localized NMR spectroscopy. Google Patents.

BOTTOMLEY, P. A. 2016. Quantifying Metabolite Ratios and Concentrations by Non- 1H MRS. *eMagRes.* John Wiley & Sons, Ltd.

BOTTOMLEY, P. A. & GRIFFITHS, J. R. 2016. *Handbook of Magnetic Resonance Spectroscopy In Vivo: MRS Theory, Practice and Applications*, Wiley.

BREKKE, E., MORKEN, T. S. & SONNEWALD, U. 2015. Glucose metabolism and astrocyte–neuron interactions in the neonatal brain. *Neurochemistry International,* 82**,** 33-41.

BROWN, R. W., HAACKE, E. M., CHENG, Y. C. N., THOMPSON, M. R. & VENKATESAN, R. 2014. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*, Wiley.

BRUHN, H., FRAHM, J., GYNGELL, M. L., MERBOLDT, K. D., HANICKE, W. & SAUTER, R. 1989. Cerebral metabolism in man after acute stroke: new observations using localized proton NMR spectroscopy. *Magn Reson Med,* 9**,** 126-31.

BRULATOUT, S., MÉRIC, P., LOUBINOUX, I., BORREDON, J., CORRÈZE, J. L., ROUCHER, P., GILLET, B., BÉRENGER, G., BELOEIL, J. C., TIFFON, B., MISPELTER, J. & SEYLAZ, J. 1996. A One-Dimensional (Proton and Phosphorus) and Two-Dimensional (Proton) In Vivo NMR Spectroscopic Study of Reversible Global Cerebral Ischemia. *Journal of Neurochemistry,* 66**,** 2491-2499.

CADY, E. B. 1996. Metabolite concentrations and relaxation in perinatal cerebral hypoxic-ischemic injury. *Neurochem Res,* 21**,** 1043-52.

CADY, E. B., COSTELLO, A. M., DAWSON, M. J., DELPY, D. T., HOPE, P. L., REYNOLDS, E. O., TOFTS, P. S. & WILKIE, D. R. 1983. Non-invasive investigation of cerebral metabolism in newborn infants by phosphorus nuclear magnetic resonance spectroscopy. *Lancet,* 1**,** 1059-62.

CADY, E. B., LOREK, A., PENRICE, J., ALDRIDGE, R., WYLEZINSKA, M., WYATT, J. S., OSMUND, E. & REYNOLDS, R. 1994a. 1H magnetic resonance spectroscopy of the brains of normal infants and after perinatal hypoxia-ischemia. *Magnetic Resonance Materials in Physics, Biology and Medicine,* 2**,** 353-355.

CADY, E. B., LOREK, A., PENRICE, J., WYLEZINSKA, M., COOPER, C. E., BROWN, G. C., OWEN-REECE, H., KIRKBRIDE, V., WYATT, J. S., OSMUND, E. & REYNOLDS, R. 1994b. Brain-metabolite transverse relaxation times in magnetic resonance spectroscopy increase as adenosine triphosphate depletes during secondary energy failure following acute hypoxia-ischaemia in the newborn piglet. *Neuroscience Letters,* 182**,** 201-204.

CADY, E. B., PENRICE, J., AMESS, P. N., LOREK, A., WYLEZINSKA, M., ALDRIDGE, R. F., FRANCONI, F., WYATT, J. S. & REYNOLDS, E. O. R. 1996. Lactate, N-acetylaspartate, choline and creatine concentrations, and spin-spin relaxation in thalamic and occipito-parietal regions of developing human brain. *Magnetic Resonance in Medicine,* 36**,** 878-886.

CHANDRASEKARAN, M., CHABAN, B., MONTALDO, P. & THAYYIL, S. 2017. Predictive value of amplitude-integrated EEG (aEEG) after rescue hypothermic neuroprotection for hypoxic ischemic encephalopathy: a meta-analysis. *J Perinatol,* 37**,** 684-689.

CHANDRASEKARAN, M., THAYYIL, S., BAINBRIDGE, A., KENDALL, G., GUNNY, R., CADY, E., MARLOW, N. & ROBERTSON, N. J. 2011. Prognostic Accuracy of Thalamic Proton Magnetic Resonance Spectroscopy Following Therapeutic Hypothermia in Neonatal Encephalopathy. *Pediatric Research,* 70**,** 161.

CHEONG, J. L. Y., CADY, E. B., PENRICE, J., WYATT, J. S., COX, I. J. & ROBERTSON, N. J. 2006. Proton MR Spectroscopy in Neonates with Perinatal Cerebral Hypoxic-Ischemic Injury: Metabolite Peak-Area Ratios, Relaxation Times, and Absolute Concentrations. *American Journal of Neuroradiology,* 27**,** 1546-1554.

CHEONG, J. Y., COLEMAN, L., HUNT, R. W. & ET AL. 2012. Prognostic utility of magnetic resonance imaging in neonatal hypoxic-ischemic encephalopathy: Substudy of a randomized trial. *Archives of Pediatrics & Adolescent Medicine,* 166**,** 634-640.

CHONG, D. G., KREIS, R., BOLLIGER, C. S., BOESCH, C. & SLOTBOOM, J. 2011. Two-dimensional linear-combination model fitting of magnetic resonance spectra to define the macromolecule baseline using FiTAID, a Fitting Tool for Arrays of Interrelated Datasets. *Magma,* 24**,** 147-64.

CHUGANI, H. & PHELPS, M. 1986. Maturational changes in cerebral function in infants determined by 18FDG positron emission tomography. *Science,* 231**,** 840-843.

CRAMÉR, H. 1946. *Mathematical Methods of Statistics (PMS-9)*, Princeton University Press.

DE GRAAF, A. A., VAN DIJK, J. E. & BOÉE, W. M. M. J. 1990. Quality: quantification improvement by converting lineshapes to the lorentzian type. *Magnetic Resonance in Medicine,* 13**,** 343-357.

DE GRAAF, R. A. 2013. *In Vivo NMR Spectroscopy: Principles and Techniques*, Wiley.

DE STEFANO, N., MATTHEWS, P. M., ANTEL, J. P., PREUL, M., FRANCIS, G. & ARNOLD, D. L. 1995. Chemical pathology of acute demyelinating lesions and its correlation with disability. *Annals of Neurology,* 38**,** 901-909.

DE VIS, J. B., HENDRIKSE, J., PETERSEN, E. T., DE VRIES, L. S., VAN BEL, F., ALDERLIESTEN, T., NEGRO, S., GROENENDAAL, F. & BENDERS, M. J. 2015. Arterial spin-labelling perfusion MRI and outcome in neonates with hypoxic-ischemic encephalopathy. *Eur Radiol,* 25**,** 113-21.

DE VRIES, L. S. & JONGMANS, M. J. 2010. Long-term outcome after neonatal hypoxic-ischaemic encephalopathy. *Arch Dis Child Fetal Neonatal Ed,* 95**,** F220-4.

DEELCHAND, D. K., AUERBACH, E. J., KOBAYASHI, N. & MARJAŃSKA, M. 2018. Transverse relaxation time constants of the five major metabolites in human brain measured in vivo using LASER and PRESS at 3 T. *Magnetic Resonance in Medicine,* 79**,** 1260-1265.

DELPY, D. T., GORDON, R. E., HOPE, P. L., PARKER, D., REYNOLDS, E. O., SHAW, D. & WHITEHEAD, M. D. 1982. Noninvasive investigation of cerebral ischemia by phosphorus nuclear magnetic resonance. *Pediatrics,* 70**,** 310-3.

DICKINSON, R. J., HALL, A. S., HIND, A. J. & YOUNG, I. R. 1986. Measurement of changes in tissue temperature using MR imaging. *Journal of computer assisted tomography,* 10**,** 468-472.

DICKINSON, W. C. 1950. Dependence of the ${\mathrm{F}}^{19}$ Nuclear Resonance Position on Chemical Compound. *Physical Review,* 77**,** 736-737.

EDDEN, R. A. E. & HARRIS, A. D. 2016. N-Acetyl Aspartate. *eMagRes.* John Wiley & Sons, Ltd.

EDWARDS, A. D., YUE, X., SQUIER, M. V., THORESEN, M., CADY, E. B., PENRICE, J., COOPER, C. E., WYATT, J. S., REYNOLDS, E. O. & MEHMET, H. 1995. Specific inhibition of apoptosis after cerebral hypoxia-ischaemia by moderate post-insult hypothermia. *Biochem Biophys Res Commun,* 217**,** 1193-9.

ELLEGOOD, J., HANSTOCK, C. C. & BEAULIEU, C. 2005. Trace apparent diffusion coefficients of metabolites in human brain using diffusion weighted magnetic resonance spectroscopy. *Magnetic Resonance in Medicine,* 53**,** 1025-1032.

EMIR, U. E., DEELCHAND, D., HENRY, P. G. & TERPSTRA, M. 2011. Noninvasive quantification of T2 and concentrations of ascorbate and glutathione in the human brain from the same double-edited spectra. *NMR Biomed,* 24**,** 263-9.

ERNST, T., KREIS, R. & ROSS, B. D. 1993. Absolute Quantitation of Water and Metabolites in the Human Brain. I. Compartments and Water. *Journal of Magnetic Resonance, Series B,* 102**,** 1-8.

FLEISS, B. & GRESSENS, P. 2012. Tertiary mechanisms of brain damage: a new hope for treatment of cerebral palsy? *The Lancet Neurology,* 11**,** 556-566.

FORTIN, J. P., PARKER, D., TUNC, B., WATANABE, T., ELLIOTT, M. A., RUPAREL, K., ROALF, D. R., SATTERTHWAITE, T. D., GUR, R. C., GUR, R. E., SCHULTZ, R. T., VERMA, R. & SHINOHARA, R. T. 2017. Harmonization of multi-site diffusion tensor imaging data. *Neuroimage,* 161**,** 149-170.

FRIEDMAN, J. H., HASTIE, T. & TIBSHIRANI, R. 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. *2010,* 33**,** 22.

GIAPITZAKIS, I. A., AVDIEVICH, N. & HENNING, A. 2018. Characterization of macromolecular baseline of human brain using metabolite cycled semi-LASER at 9.4 T. *Magnetic resonance in medicine*.

GOLUB, G. H. & PEREYRA, V. 1973. The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate. *SIAM Journal on Numerical Analysis,* 10**,** 413-432.

GOVINDARAJU, V., YOUNG, K. & MAUDSLEY, A. A. 2000. Proton NMR chemical shifts and coupling constants for brain metabolites. *NMR in Biomedicine,* 13**,** 129-153.

GRAHAM, G. D., HWANG, J. H., ROTHMAN, D. L. & PRICHARD, J. W. 2001. Spectroscopic assessment of alterations in macromolecule and small-molecule metabolites in human brain after stroke. *Stroke,* 32**,** 2797-802.

GRANOT, J. 1986. Selected Volume Excitation Using Stimulated Echoes (VEST). Applications to spatially localized spectroscopy and imaging. *Journal of Magnetic Resonance (1969),* 70**,** 488-492.

GROENENDAAL, F., VEENHOVEN, R. H., VAN DER GROND, J., JANSEN, G. H., WITKAMP, T. D. & DE VRIES, L. S. 1994. Cerebral lactate and N-acetyl-aspartate/choline ratios in asphyxiated full-term neonates demonstrated in vivo using proton magnetic resonance spectroscopy. *Pediatr Res,* 35**,** 148-51.

HERMINGHAUS, S., PILATUS, U., MOLLER-HARTMANN, W., RAAB, P., LANFERMANN, H., SCHLOTE, W. & ZANELLA, F. E. 2002. Increased choline levels coincide with enhanced proliferative activity of human neuroepithelial brain tumors. *NMR Biomed,* 15**,** 385-92.

HOFMANN, L., SLOTBOOM, J., BOESCH, C. & KREIS, R. 2001. Characterization of the macromolecule baseline in localized 1H-MR spectra of human brain. *Magnetic Resonance in Medicine,* 46**,** 855-863.

HOFMANN, L., SLOTBOOM, J., JUNG, B., MALOCA, P., BOESCH, C. & KREIS, R. 2002. Quantitative 1H-magnetic resonance spectroscopy of human brain: Influence of composition and parameterization of the basis set in linear combination model-fitting. *Magnetic Resonance in Medicine,* 48**,** 440-453.

HOPE, P. L., COSTELLO, A. M., CADY, E. B., DELPY, D. T., TOFTS, P. S., CHU, A., HAMILTON, P. A., REYNOLDS, E. O. & WILKIE, D. R. 1984. Cerebral energy metabolism studied with phosphorus NMR spectroscopy in normal and birth-asphyxiated infants. *Lancet,* 2**,** 366-70.

HOWE, F. A., BARTON, S. J., CUDLIP, S. A., STUBBS, M., SAUNDERS, D. E., MURPHY, M., WILKINS, P., OPSTAD, K. S., DOYLE, V. L., MCLEAN, M. A., BELL, B. A. & GRIFFITHS, J. R. 2003. Metabolic profiles of human brain tumors using quantitative in vivo 1H magnetic resonance spectroscopy. *Magn Reson Med,* 49**,** 223-32.

JACOBS, S. E., BERG, M., HUNT, R., TARNOW-MORDI, W. O., INDER, T. E. & DAVIS, P. G. 2013. Cooling for newborns with hypoxic ischaemic encephalopathy. *Cochrane Database of Systematic Reviews*.

KASSEM, M. N. & BARTHA, R. 2003. Quantitative proton short-echo-time LASER spectroscopy of normal human white matter and hippocampus at 4 Tesla incorporating macromolecule subtraction. *Magn Reson Med,* 49**,** 918-27.

KE, Y., COHEN, B. M., LOWEN, S., HIRASHIMA, F., NASSAR, L. & RENSHAW, P. F. 2002. Biexponential transverse relaxation (T2) of the proton MRS creatine resonance in human brain. *Magnetic Resonance in Medicine,* 47**,** 232-238.

KHODANOVICH, M. Y., KISEL, A. A., AKULOV, A. E., ATOCHIN, D. N., KUDABAEVA, M. S., GLAZACHEVA, V. Y., SVETLIK, M. V., MEDVEDNIKOVA, Y. A., MUSTAFINA, L. R. & YARNYKH, V. L. 2018. Quantitative assessment of demyelination in ischemic stroke in vivo using macromolecular proton fraction mapping. *J Cereb Blood Flow Metab***,** 271678x18755203.

KIROV, I. I., FLEYSHER, L., FLEYSHER, R., PATIL, V., LIU, S. & GONEN, O. 2008. The Age Dependence of Regional Proton Metabolites T(2) Relaxation Times in the Human Brain at 3 Tesla. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine,* 60**,** 790-795.

KLOSE, U. 1990. In vivo proton spectroscopy in presence of eddy currents. *Magnetic Resonance in Medicine,* 14**,** 26-30.

KNIGHT-SCOTT, J. Analysis of Saturated T2 Curves for Rapid Relaxometry Measurements in PRESS Localization. ISMRM 19th scientific meeting, Montréal, Québec, Canada, 2011.

KREIS, R. 2004. Issues of spectral quality in clinical 1H-magnetic resonance spectroscopy and a gallery of artifacts. *NMR in Biomedicine,* 17**,** 361-381.

KREIS, R. 2016. The trouble with quality filtering based on relative Cramer-Rao lower bounds. *Magn Reson Med,* 75**,** 15-8.

KREIS, R. & BOESCH, C. Bad spectra can be better than good spectra. ISMRM 11th scientific meeting, Toronto, Ontario, Canada, 2003.

KREIS, R., ERNST, T. & ROSS, B. D. 1993. Development of the human brain: In vivo quantification of metabolite and water content with proton magnetic resonance spectroscopy. *Magnetic Resonance in Medicine,* 30**,** 424-437.

KREIS, R., HOFMANN, L., KUHLMANN, B., BOESCH, C., BOSSI, E. & HUPPI, P. S. 2002. Brain metabolite composition during early human brain development as measured by quantitative in vivo 1H magnetic resonance spectroscopy. *Magn Reson Med,* 48**,** 949-58.

KREIS, R., SLOTBOOM, J., HOFMANN, L. & BOESCH, C. 2005. Integrated data acquisition and processing to determine metabolite contents, relaxation times, and macromolecule baseline in single examinations of individual subjects. *Magnetic Resonance in Medicine,* 54**,** 761-768.

KRUSE, M., MICHELSEN, S. I., FLACHS, E. M., BRONNUM-HANSEN, H., MADSEN, M. & ULDALL, P. 2009. Lifetime costs of cerebral palsy. *Dev Med Child Neurol,* 51**,** 622-8.

KUNZ, N., CUDALBU, C., MLYNARIK, V., HUPPI, P. S., SIZONENKO, S. V. & GRUETTER, R. 2010. Diffusion-weighted spectroscopy: a novel approach to determine macromolecule resonances in short-echo time 1H-MRS. *Magn Reson Med,* 64**,** 939-46.

LALLY, P. J., PAULIAH, S., MONTALDO, P., CHABAN, B., OLIVEIRA, V., BAINBRIDGE, A., SOE, A., PATTNAYAK, S., CLARKE, P., SATODIA, P., HARIGOPAL, S., ABERNETHY, L. J., TURNER, M. A., HUERTAS-CEBALLOS, A., SHANKARAN, S. & THAYYIL, S. 2015. Magnetic Resonance Biomarkers in Neonatal Encephalopathy (MARBLE): a prospective multicountry study. *BMJ Open,* 5.

LENTNER, C. 1981. *Geigy scientific tables*, Ciba-Geigy.

LOLLEY, R. N., BALFOUR, W. M. & SAMSON, F. E. 1961. The high-energy phosphates in developing brain. *Journal of Neurochemistry,* 7**,** 289-297.

LOREK, A., TAKEI, Y., CADY, E. B., WYATT, J. S., PENRICE, J., EDWARDS, A. D., PEEBLES, D., WYLEZINSKA, M., OWEN-REECE, H., KIRKBRIDE, V. & ET AL. 1994. Delayed ("secondary") cerebral energy

failure after acute hypoxia-ischemia in the newborn piglet: continuous 48-hour studies by phosphorus magnetic resonance spectroscopy. *Pediatr Res,* 36**,** 699-706.
MADER, I., SEEGER, U., WEISSERT, R., KLOSE, U., NAEGELE, T., MELMS, A. & GRODD, W. 2001. Proton MR spectroscopy with metabolite-nulling reveals elevated macromolecules in acute multiple sclerosis. *Brain,* 124**,** 953-61.
MANGIA, S., TKAC, I., GRUETTER, R., VAN DE MOORTELE, P. F., GIOVE, F., MARAVIGLIA, B. & UGURBIL, K. 2006. Sensitivity of single-voxel 1H-MRS in investigating the metabolism of the activated human visual cortex at 7 T. *Magn Reson Imaging,* 24**,** 343-8.
MARJAŃSKA, M., DEELCHAND, D. K., HODGES, J. S., MCCARTEN, J. R., HEMMY, L. S., GRANT, A. & TERPSTRA, M. 2018. Altered macromolecular pattern and content in the aging human brain. *NMR in Biomedicine,* 31**,** e3865-n/a.
MARSHALL, I., HIGINBOTHAM, J., BRUCE, S. & FREISE, A. 1997. Use of Voigt lineshape for quantification of in vivo 1H spectra. *Magn Reson Med,* 37**,** 651-7.
MASON, S. 2017. Lactate Shuttles in Neuroenergetics—Homeostasis, Allostasis and Beyond. *Frontiers in Neuroscience,* 11.
MCKINSTRY, R. C., MILLER, J. H., SNYDER, A. Z., MATHUR, A., SCHEFFT, G. L., ALMLI, C. R., SHIMONY, J. S., SHIRAN, S. I. & NEIL, J. J. 2002. A prospective, longitudinal diffusion tensor imaging study of brain injury in newborns. *Neurology,* 59**,** 824-33.
MELBOURNE, A., EATON-ROSEN, Z., ORASANU, E., PRICE, D., BAINBRIDGE, A., CARDOSO, M. J., KENDALL, G. S., ROBERTSON, N. J., MARLOW, N. & OURSELIN, S. 2016. Longitudinal development in the preterm thalamus and posterior white matter: MRI correlations between diffusion weighted imaging and T2 relaxometry. *Human Brain Mapping,* 37**,** 2479-2492.
METROPOLIS, N. & ULAM, S. 1949. The Monte Carlo Method. *Journal of the American Statistical Association,* 44**,** 335-341.
MOFFETT, J. R. & NAMBOODIRI, A. M. 2006. Expression of N-acetylaspartate and N-acetylaspartylglutamate in the nervous system. *N-Acetylaspartate.* Springer.
MOFFETT, J. R., ROSS, B., ARUN, P., MADHAVARAO, C. N. & NAMBOODIRI, M. A. A. 2007. N-Acetylaspartate in the CNS: From Neurodiagnostics to Neurobiology. *Progress in neurobiology,* 81**,** 89-131.
MOORCRAFT, J., BOLAS, N. M., IVES, N. K., OUWERKERK, R., SMYTH, J., RAJAGOPALAN, B., HOPE, P. L. & RADDA, G. K. 1991. Global and depth resolved phosphorus magnetic resonance spectroscopy to predict outcome after birth asphyxia. *Archives of Disease in Childhood,* 66**,** 1119-1123.
MUÑOZ MANIEGA, S., CVORO, V., CHAPPELL, F. M., ARMITAGE, P. A., MARSHALL, I., BASTIN, M. E. & WARDLAW, J. M. 2008. Changes in NAA and lactate following ischemic stroke. *A serial MR spectroscopic imaging study,* 71**,** 1993-1999.
NATIONAL INSTITUTE FOR CLINICAL EXCELLENCE 2010. Therapeutic hypothermia with intracorporeal temperature monitoring for hypoxic perinatal brain injury. *Guidance NICE Guidelines. Interventional Procedures*.
OLIVEIRA, V., SINGHVI, D. P., MONTALDO, P., LALLY, P. J., MENDOZA, J., MANERKAR, S., SHANKARAN, S. & THAYYIL, S. 2017. Therapeutic hypothermia in mild neonatal encephalopathy: a national survey of practice in the UK. *Arch Dis Child Fetal Neonatal Ed*.
OPSTAD, K. S., BELL, B. A., GRIFFITHS, J. R. & HOWE, F. A. 2008a. An investigation of human brain tumour lipids by high-resolution magic angle spinning 1H MRS and histological analysis. *NMR in Biomedicine,* 21**,** 677-685.
OPSTAD, K. S., GRIFFITHS, J. R., BELL, B. A. & HOWE, F. A. 2008b. Apparent T2 relaxation times of lipid and macromolecules: A study of high-grade tumor spectra. *Journal of Magnetic Resonance Imaging,* 27**,** 178-184.
ÖZ, G., ALGER, J. R., BARKER, P. B., BARTHA, R., BIZZI, A., BOESCH, C., BOLAN, P. J., BRINDLE, K. M., CUDALBU, C., DINÇER, A., DYDAK, U., EMIR, U. E., FRAHM, J., GONZÁLEZ, R. G., GRUBER, S., GRUETTER, R., GUPTA, R. K., HEERSCHAP, A., HENNING, A., HETHERINGTON, H. P., HOWE, F.

A., HÜPPI, P. S., HURD, R. E., KANTARCI, K., KLOMP, D. W. J., KREIS, R., KRUISKAMP, M. J., LEACH, M. O., LIN, A. P., LUIJTEN, P. R., MARJAŃSKA, M., MAUDSLEY, A. A., MEYERHOFF, D. J., MOUNTFORD, C. E., NELSON, S. J., PAMIR, M. N., PAN, J. W., PEET, A. C., POPTANI, H., POSSE, S., POUWELS, P. J. W., RATAI, E.-M., ROSS, B. D., SCHEENEN, T. W. J., SCHUSTER, C., SMITH, I. C. P., SOHER, B. J., TKÁČ, I., VIGNERON, D. B., KAUPPINEN, R. A. & GROUP, F. T. M. C. 2014. Clinical Proton MR Spectroscopy in Central Nervous System Disorders. *Radiology,* 270**,** 658-679.

PAPILE, L. A., BALEY, J. E., BENITZ, W., CUMMINGS, J., CARLO, W. A., EICHENWALD, E., KUMAR, P., POLIN, R. A., TAN, R. C. & WANG, K. S. 2014. Hypothermia and neonatal encephalopathy. *Pediatrics,* 133**,** 1146-50.

PEET, A. C., DAVIES, N. P., RIDLEY, L., BRUNDLER, M.-A., KOMBOGIORGAS, D., LATEEF, S., NATARAJAN, K., SGOUROS, S., MACPHERSON, L. & GRUNDY, R. G. 2007. Magnetic resonance spectroscopy suggests key differences in the metastatic behaviour of medulloblastoma. *European Journal of Cancer,* 43**,** 1037-1044.

PELLERIN, L. & MAGISTRETTI, P. J. 1994. Glutamate uptake into astrocytes stimulates aerobic glycolysis: a mechanism coupling neuronal activity to glucose utilization. *Proc Natl Acad Sci U S A,* 91**,** 10625-9.

PENNER, J., CURTIS, A., GILBERT, K., KLASSEN, M., GATI, J., SMITH, M., BORRIE, M. & BARTHA, R. 2010. *Metabolite Nulling to Measure the Macromolecule Baseline for Quantitative 1H Magnetic Resonance Spectroscopy at 7 Tesla*.

PENRICE, J., LOREK, A., CADY, E. B., AMESS, P. N., WYLEZINSKA, M., COOPER, C. E., D'SOUZA, P., BROWN, G. C., KIRKBRIDE, V., EDWARDS, A. D., WYATT, J. S. & REYNOLDS, E. O. R. 1997. Proton Magnetic Resonance Spectroscopy of the Brain during Acute Hypoxia-Ischemia and Delayed Cerebral Energy Failure in the Newborn Piglet. *Pediatric Research,* 41**,** 795.

PIJNAPPEL, W. W. F., VAN DEN BOOGAART, A., DE BEER, R. & VAN ORMONDT, D. 1992. SVD-based quantification of magnetic resonance signals. *Journal of Magnetic Resonance (1969),* 97**,** 122-134.

POULLET, J. B., SIMA, D. M., SIMONETTI, A. W., DE NEUTER, B., VANHAMME, L., LEMMERLING, P. & VAN HUFFEL, S. 2007. An automated quantitation of short echo time MRS spectra in an open source software environment: AQSES. *NMR Biomed,* 20**,** 493-504.

POVAŽAN, M., STRASSER, B., HANGEL, G., HECKOVA, E., GRUBER, S., TRATTNIG, S. & BOGNER, W. 2018. Simultaneous mapping of metabolites and individual macromolecular components via ultra-short acquisition delay 1H MRSI in the brain at 7T. *Magnetic Resonance in Medicine,* 79**,** 1231-1240.

PRICHARD, J., ROTHMAN, D., NOVOTNY, E., PETROFF, O., KUWABARA, T., AVISON, M., HOWSEMAN, A., HANSTOCK, C. & SHULMAN, R. 1991. Lactate rise detected by 1H NMR in human visual cortex during physiologic stimulation. *Proceedings of the National Academy of Sciences,* 88**,** 5829-5831.

PROCTOR, W. G. & YU, F. C. 1950. The Dependence of a Nuclear Magnetic Resonance Frequency upon Chemical Compound. *Physical Review,* 77**,** 717-717.

PROVENCHER, S. W. 1993. Estimation of metabolite concentrations from localized in vivo proton NMR spectra. *Magnetic Resonance in Medicine,* 30**,** 672-679.

PROVENCHER, S. W. 2001. Automatic quantitation of localized in vivo 1H spectra with LCModel. *NMR Biomed,* 14**,** 260-4.

PURCELL, E. M., TORREY, H. C. & POUND, R. V. 1946. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Physical Review,* 69**,** 37-38.

R CORE TEAM 2013. R: A language and environment for statistical computing.

RAO, C. R. 1992. Information and the Accuracy Attainable in the Estimation of Statistical Parameters. *In:* KOTZ, S. & JOHNSON, N. L. (eds.) *Breakthroughs in Statistics: Foundations and Basic Theory.* New York, NY: Springer New York.

RATINEY, H., COENRADIE, Y., CAVASSILA, S., VAN ORMONDT, D. & GRAVERON-DEMILLY, D. 2004. Time-domain quantitation of 1H short echo-time signals: background accommodation. *Magma,* 16**,** 284-96.

RUTHERFORD, M., RAMENGHI, L. A., EDWARDS, A. D., BROCKLEHURST, P., HALLIDAY, H., LEVENE, M., STROHM, B., THORESEN, M., WHITELAW, A. & AZZOPARDI, D. 2010. Assessment of brain tissue injury after moderate hypothermia in neonates with hypoxic-ischaemic encephalopathy: a nested substudy of a randomised controlled trial. *Lancet Neurol,* 9**,** 39-45.

SAGER, T. N., HANSEN, A. J. & LAURSEN, H. 2000. Correlation between N-acetylaspartate levels and histopathologic changes in cortical infarcts of mice after middle cerebral artery occlusion. *J Cereb Blood Flow Metab,* 20**,** 780-8.

SAGER, T. N., LAURSEN, H. & HANSEN, A. J. 1995. Changes in N-acetyl-aspartate content during focal and global brain ischemia of the rat. *J Cereb Blood Flow Metab,* 15**,** 639-46.

SAGER, T. N., TOPP, S., TORUP, L., HANSON, L. G., EGESTAD, B. & MOLLER, A. 2001. Evaluation of CA1 damage using single-voxel 1H-MRS and un-biased stereology: Can non-invasive measures of N-acetyl-asparate following global ischemia be used as a reliable measure of neuronal damage? *Brain Res,* 892**,** 166-75.

SAUNDERS, D. E., HOWE, F. A., BOOGAART, A. V. D., GRIFFITHS, J. R. & BROWN, M. M. 1997. Discrimination of Metabolite from Lipid and Macromolecule Resonances in Cerebral Infarction in Humans Using Short Echo Proton Spectroscopy. *Journal of Magnetic Resonance Imaging,* 7**,** 1116-1121.

SAUNDERS, D. E., HOWE, F. A., VAN DEN BOOGAART, A., MCLEAN, M. A., GRIFFITHS, J. R. & BROWN, M. M. 1995. Continuing ischemic damage after acute middle cerebral artery infarction in humans demonstrated by short-echo proton spectroscopy. *Stroke,* 26**,** 1007-13.

SCHALLER, B., XIN, L. & GRUETTER, R. 2014. Is the macromolecule signal tissue-specific in healthy human brain? A (1)H MRS study at 7 Tesla in the occipital lobe. *Magn Reson Med,* 72**,** 934-40.

SHANKARAN, S., MCDONALD, S. A., LAPTOOK, A. R., HINTZ, S. R., BARNES, P. D., DAS, A., PAPPAS, A., HIGGINS, R. D., ON BEHALF OF THE EUNICE KENNEDY SHRIVER NATIONAL INSTITUTE OF CHILD, H. & HUMAN DEVELOPMENT NEONATAL RESEARCH, N. 2015. Neonatal Magnetic Resonance Imaging Pattern of Brain Injury as a Biomarker of Childhood Outcomes following a Trial of Hypothermia for Neonatal Hypoxic-Ischemic Encephalopathy. *The Journal of pediatrics,* 167**,** 987-993.e3.

SHANNON, R. J., VAN DER HEIDE, S., CARTER, E. L., JALLOH, I., MENON, D. K., HUTCHINSON, P. J. & CARPENTER, K. L. 2016. Extracellular N-Acetylaspartate in Human Traumatic Brain Injury. *J Neurotrauma,* 33**,** 319-29.

SIGNORETTI, S., MARMAROU, A., TAVAZZI, B., LAZZARINO, G., BEAUMONT, A. & VAGNOZZI, R. 2001. N-Acetylaspartate reduction as a measure of injury severity and mitochondrial dysfunction following diffuse traumatic brain injury. *J Neurotrauma,* 18**,** 977-91.

SIMA, D. M., GARCIA, M. I. O., POULLET, J., SUVICHAKORN, A., ANTOINE, J. P., HUFFEL, S. V. & ORMONDT, D. V. 2009. Lineshape estimation for magnetic resonance spectroscopy (MRS) signals: self-deconvolution revisited. *Measurement Science and Technology,* 20**,** 104031.

SIMA, D. M. & VAN HUFFEL, S. 2007. Separable nonlinear least squares fitting with linear bound constraints and its application in magnetic resonance spectroscopy data quantification. *Journal of Computational and Applied Mathematics,* 203**,** 264-278.

SIMPSON, R., DEVENYI, G. A., JEZZARD, P., HENNESSY, T. J. & NEAR, J. 2017. Advanced processing and simulation of MRS data using the FID appliance (FID-A)—An open source, MATLAB-based toolkit. *Magnetic Resonance in Medicine,* 77**,** 23-33.

SLOTBOOM, J., BOESCH, C. & KREIS, R. 1998. Versatile frequency domain fitting using time domain models and prior knowledge. *Magnetic Resonance in Medicine,* 39**,** 899-911.

SNOUSSI, K., GILLEN, J. S., HORSKA, A., PUTS, N. A., PRADHAN, S., EDDEN, R. A. & BARKER, P. B. 2015. Comparison of brain gray and white matter macromolecule resonances at 3 and 7 Tesla. *Magn Reson Med,* 74**,** 607-13.

SOHER, B., SEMANCHUK, P., TODD, D., STEINBERG, J. & YOUNG, K. Vespa: Integrated applications for RF pulse design, spectral simulation and MRS data analysis. Proc Int Soc Magn Reson Med, 2011. 1410.

SOKOLSKA, M., URIA-AVELLANAL, C., JORGE CARDOSO, M., PROISY, M., BAINBRIDGE, A., OURSELIN, S., THOMAS, D., ROBERTSON, N. & GOLAY, X. Assessing brain damage after perinatal hypoxic-ischaemia using an automated protocol for combined regional analysis of the Cerebral Blood Flow and MR spectroscopy ISMRM 23rd scientific meeting, Toronto, Ontario, Canada, 2015.

TAYLOR, D. L., DAVIES, S. E. C., OBRENOVITCH, T. P., URENJAK, J., RICHARDS, D. A., CLARK, J. B. & SYMON, L. 1994. Extracellular N-Acetylaspartate in the Rat Brain: In Vivo Determination of Basal Levels and Changes Evoked by High K+. *Journal of Neurochemistry,* 62**,** 2349-2355.

TERPSTRA, M., CHEONG, I., LYU, T., DEELCHAND, D. K., EMIR, U. E., BEDNARIK, P., EBERLY, L. E. & OZ, G. 2016. Test-retest reproducibility of neurochemical profiles with short-echo, single-voxel MR spectroscopy at 3T and 7T. *Magn Reson Med,* 76**,** 1083-91.

THAYYIL, S., CHANDRASEKARAN, M., TAYLOR, A., BAINBRIDGE, A., CADY, E. B., CHONG, W. K., MURAD, S., OMAR, R. Z. & ROBERTSON, N. J. 2010. Cerebral magnetic resonance biomarkers in neonatal encephalopathy: a meta-analysis. *Pediatrics,* 125**,** e382-95.

THAYYIL, S., OLIVEIRA, V., LALLY, P. J., SWAMY, R., BASSETT, P., CHANDRASEKARAN, M., MONDKAR, J., MANGALABHARATHI, S., BENKAPPA, N., SEERALAR, A., SHAHIDULLAH, M., MONTALDO, P., HERBERG, J., MANERKAR, S., KUMARASWAMI, K., KAMALARATNAM, C., PRAKASH, V., CHANDRAMOHAN, R., BANDYA, P., MANNAN, M. A., RODRIGO, R., NAIR, M., RAMJI, S. & SHANKARAN, S. 2017. Hypothermia for encephalopathy in low and middle-income countries (HELIX): study protocol for a randomised controlled trial. *Trials,* 18**,** 432.

THORESEN, M., PENRICE, J., LOREK, A., CADY, E. B., WYLEZINSKA, M., KIRKBRIDE, V., COOPER, C. E., BROWN, G. C., EDWARDS, A. D., WYATT, J. S. & ET AL. 1995. Mild hypothermia after severe transient hypoxia-ischemia ameliorates delayed cerebral energy failure in the newborn piglet. *Pediatr Res,* 37**,** 667-70.

THORESEN, M., SATAS, S., PUKA-SUNDVALL, M., WHITELAW, A., HALLSTROM, A., LOBERG, E. M., UNGERSTEDT, U., STEEN, P. A. & HAGBERG, H. 1997. Post-hypoxic hypothermia reduces cerebrocortical release of NO and excitotoxins. *Neuroreport,* 8**,** 3359-62.

TIBSHIRANI, R. 1996. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological),* 58**,** 267-288.

TKAC, I., UGURBIL, K. & GRUETTER, R. On the quantification of low concentration metabolites by 1H NMR spectroscopy in the human brain at 7 Tesla. Proceedings of the 10th Meeting of the International Society of Magnetic Resonance in Medicine, Honolulu, USA, 2002. 528.

TOMIYASU, M., AIDA, N., ENDO, M., SHIBASAKI, J., NOZAWA, K., SHIMIZU, E., TSUJI, H. & OBATA, T. 2013. Neonatal Brain Metabolite Concentrations: An In Vivo Magnetic Resonance Spectroscopy Study with a Clinical MR System at 3 Tesla. *PLOS ONE,* 8**,** e82746.

TOORN, A. V. D., DIJKHUIZEN, R. M., TULLEKEN, C. A. F. & NICOLAY, K. 1995. T1 and T2 relaxation times of the major 1H-containing metabolites in rat brain after focal ischemia. *NMR in Biomedicine,* 8**,** 245-252.

VAN DER TOORN, A., VERHEUL, H. B., BERKELBACH VAN DER SPRENKEL, J. W., TULLEKEN, C. A. & NICOLAY, K. 1994. Changes in metabolites and tissue water status after focal ischemia in cat brain assessed with localized proton MR spectroscopy. *Magn Reson Med,* 32**,** 685-91.

VANHAMME, L., VAN DEN BOOGAART, A. & VAN HUFFEL, S. 1997. Improved method for accurate and efficient quantification of MRS data with use of prior knowledge. *J Magn Reson,* 129**,** 35-43.

VANNUCCI, S. J., SEAMAN, L. B., BRUCKLACHER, R. M. & VANNUCCI, R. C. 1994. Glucose transport in developing rat brain: Glucose transporter proteins, rate constants and cerebral glucose utilization. *Molecular and Cellular Biochemistry,* 140**,** 177-184.

VOLPE, J. J. 2008. *Neurology of the Newborn*, Saunders/Elsevier.

WALKER, P. M., BEN SALEM, D. E., LALANDE, A., GIROUD, M. & BRUNOTTE, F. 2004. Time course of NAA T2 and ADCw in ischaemic stroke patients: 1H MRS imaging and diffusion-weighted MRI. *Journal of the Neurological Sciences,* 220**,** 23-28.

WARDLAW, J. M., MARSHALL, I., WILD, J., DENNIS, M. S., CANNON, J. & LEWIS, S. C. 1998. Studies of acute ischemic stroke with proton magnetic resonance spectroscopy: relation between time from onset, neurological deficit, metabolite abnormalities in the infarct, blood flow, and clinical outcome. *Stroke,* 29**,** 1618-24.

WERMTER, F. C., MITSCHKE, N., BOCK, C. & DREHER, W. 2017. Temperature dependence of (1)H NMR chemical shifts and its influence on estimated metabolite concentrations. *Magma,* 30**,** 579-590.

WILSON, M., REYNOLDS, G., KAUPPINEN, R. A., ARVANITIS, T. N. & PEET, A. C. 2011. A constrained least-squares approach to the automated quantitation of in vivo (1)H magnetic resonance spectroscopy data. *Magn Reson Med,* 65**,** 1-12.

WISNOWSKI, J. L., WU, T. W., REITMAN, A. J., MCLEAN, C., FRIEDLICH, P., VANDERBILT, D., HO, E., NELSON, M. D., PANIGRAHY, A. & BLUML, S. 2016. The effects of therapeutic hypothermia on cerebral metabolism in neonates with hypoxic-ischemic encephalopathy: An in vivo 1H-MR spectroscopy study. *J Cereb Blood Flow Metab,* 36**,** 1075-86.

YABLONSKIY, D. A., NEIL, J. J., RAICHLE, M. E. & ACKERMAN, J. J. H. 1998. Homonuclear J coupling effects in volume localized NMR spectroscopy: Pitfalls and solutions. *Magnetic Resonance in Medicine,* 39**,** 169-178.

ZHANG, Y. & SHEN, J. 2014. Smoothness of In vivo Spectral Baseline Determined by Mean Squared Error. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine,* 72**,** 913-922.

APPENDIX I: RESEARCH OUTPUT

SELECTED CONFERENCE ABSTRACTS:

Lally PJ et al. (2018) Early MRS biomarkers accurately predict neurodevelopment after neonatal encephalopathy in a multicentre setting. ISMRM Annual Meeting, Paris, France [$525 Stipend: ISMRM]

Lally PJ (2017) Magnetic resonance biomarkers to speed up drug development in neonatal encephalopathy. NEOCON 2017, Gurgaon, India [Invited talk]

Lally PJ et al. (2017) Dual spectral analysis for metabolite quantification in the presence of lipids and macromolecules, ISMRM Annual Meeting, Honolulu, USA

Lally PJ et al. (2016) Diffusion-Based Segmentation of Long T2 Components for Partial Volume Correction in MRS, ISMRM Spectroscopy Workshop, Allensbach, Germany

Lally PJ et al. (2015) Basis Set Choice Biases Quantification in Long TE Proton Spectroscopy, ESMRMB Annual Scientific Meeting, Edinburgh, UK

SELECTED PUBLICATIONS:

Thayyil S, Oliveira V, Lally PJ et al. (2017) Hypothermia for Encephalopathy in Low and Middle-Income Countries (HELIX) Trial. Trials. Sep18;18(1):432. doi: 10.1186/s13063-017-2165-3.

Lally PJ, Montaldo P, Oliveira V et al. (2017) Residual brain injury after early discontinuation of cooling therapy in mild neonatal encephalopathy. ADC Fetal & Neonatal Edition. Online first doi: 10.1136/archdischild-2017-313321

Lally PJ et al. (2015) Magnetic Resonance Biomarkers in Neonatal Encephalopathy (MARBLE): a prospective multicountry study. BMJ Open 2015;5:e008912

GRANT INCOME:

NIHR Healthcare Science Doctoral Fellowship (Lead Applicant) March 2015 – February 2018: £281,590

ITMAT Imperial Biomedical Research Centre (Co-Applicant) April 2016 – March 2017: £37,000

NIHR Clinical Trials Fellowship (Lead Applicant) March 2018 – December 2018: £59,209

Cerebral Palsy Alliance Project Grant (Lead Applicant) March 2018 – February 2019: AUD 50,000

APPENDIX II: MARBLE STUDY FLOWCHART, MRI PROTOCOL AND TECHNICAL DETAILS

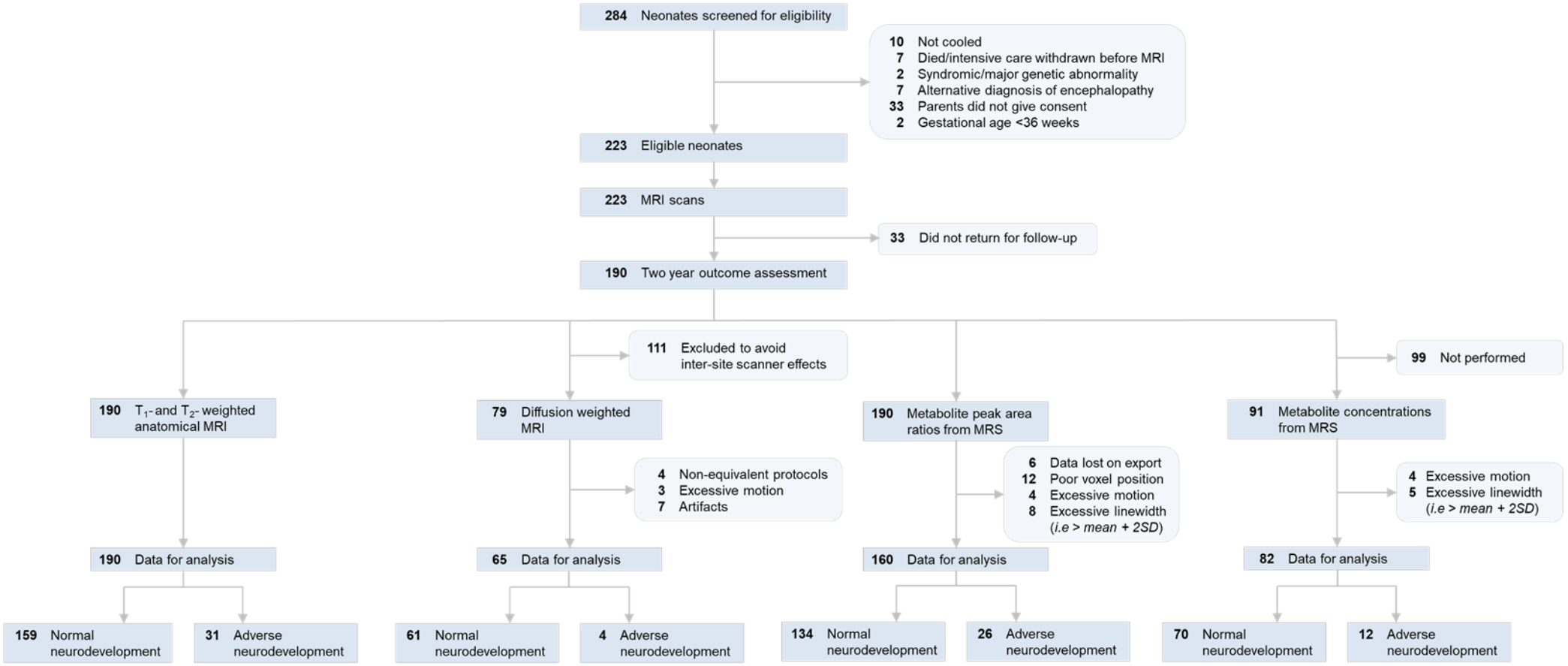

Figure A2.1: Flowchart depicting the screening and recruitment of neonates, and their inclusion or exclusion at various stages of MR data analysis

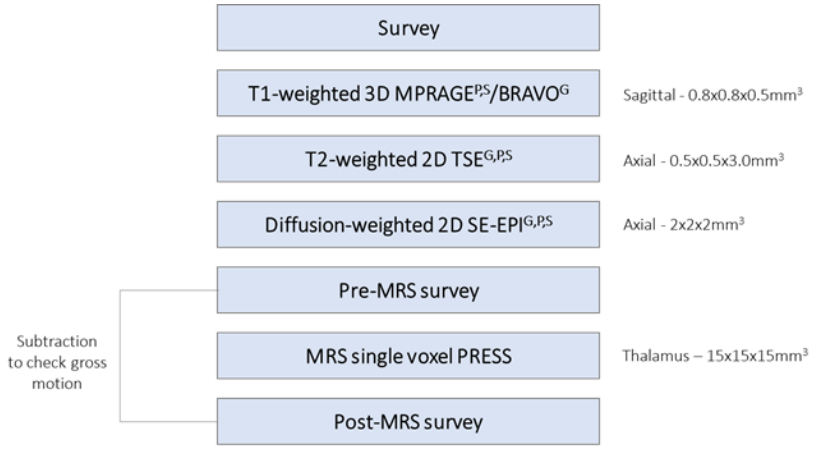

Figure A2.2: MR protocol for the MARBLE study. [G] GE scanners, [P] Philips scanners, [S] Siemens scanners

Site	Scanner	Software version	Receiver coil
Imperial	Philips Achieva 3T	3.2.1	Philips 8ch SENSE Head Array
UCL	Philips Achieva 3T	3.2.1	LMT 8ch Neonatal Head Array
Liverpool	Philips Achieva 3T	3.2.1	Philips 8ch SENSE Head Array
Medway	GE Signa HDXt 3T	23	GE 16ch Head/Neck/Spine Array
Coventry	GE Signa HDxt 3T	15, 24	GE 8ch HR Brain Array
Wayne State	GE Signa HDxt 3T	15	GE 8ch HR Brain Array
Norwich	GE Discovery 750w 3T	23, 24	GE 21ch Head and Neck Array
Newcastle	Siemens Verio 3T	VB17, VB19	Siemens 12ch Head Matrix Array
Bristol	Siemens Skyra 3T	VD13	Siemens 32ch Head Array

Table A2.1: MR scanners used in the MARBLE study (including Bristol, where only phantom data was collected).

Substance	Concentration (mM)	Mass in 1 litre (g)	Reason
N-acetylaspartate	10	1.79	Metabolite
Lactate	10	1.15	Metabolite
K_2HPO_4	72	12.7	pH buffer
KH_2PO_4	28	3.83	pH buffer
NaN_3	-	1.00	Bacteriostatic
NaOH	-	as needed	Titrate for pH
HCl	-	as needed	Titrate for pH

Table A2.2: Phantom ingredients. NaOH and concentrated HCl are used for careful titration to reach a pH of 7.2, and NaN_3 is added to stop bacterial growth, as outlined by Provencher in the LCModel manual. No gadolinium was included, so the relaxation rates were substantially longer than in-vivo.

Vendor	Sequence name	RF pulses	Flip angles	Outer volume suppression	Readout bandwidth (kHz)	Complex datapoints
Philips	SV-PRESS	Sharp	90°-180°-180° (Default)	Off	4.0	2048
Siemens	SVS-SE	Default	90°-180°-180° (Default)	Off	2.5	2048
GE	PROBE-P	Default	90°-137°-137° (Default)	On (Default)	5.0 / 4.0*	2048

*Table A2.3: Implementations of the PRESS sequence across different vendors. Only Philips allow specification of the shape of refocusing pulses in their product ('Sharp' chosen to minimise chemical shift displacement). GE refocusing pulses have a reduced flip angle to avoid stretching according to different patient weights. *At one site with research privileges, the readout bandwidth was reduced from the default (5.0 kHz) to 4.0 kHz in order to replicate the Philips data. Metabolite basis sets for fitting were created for each sequence, and in the case of GE, also for each bandwidth.*

APPENDIX III: EXAMPLE SPECTRA

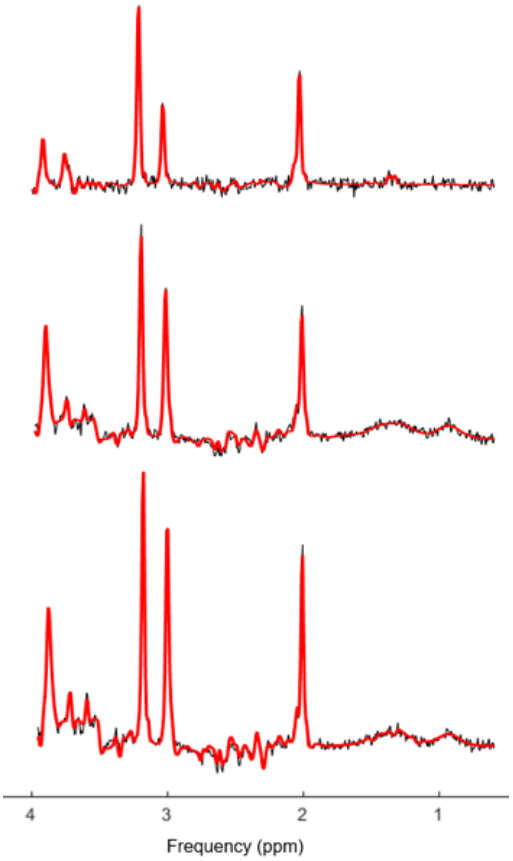

Figure A3.1: Example spectra using the PRESS protocol, in the same subject as shown in Figure A2.2. Top) TR/TE=2288/288ms, Middle) TR/TE=2060/60ms. Bottom) TR/TE=5000/60ms. LCModel fit in red.

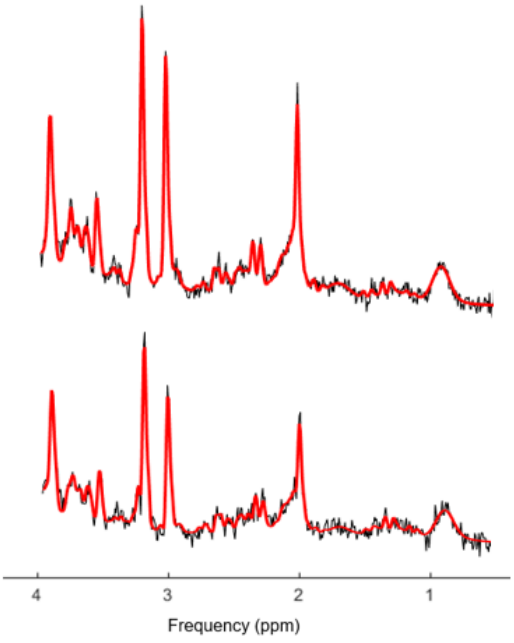

Figure A3.2: Example spectra using the STEAM protocol, in the same subject as shown in Figure A2.1. Top) TR/TM/TE=5000/20/20ms, Bottom) TR/TM/TE=1500/20/20ms. LCModel fit in red.

APPENDIX IV: SIMULATED BASIS SETS

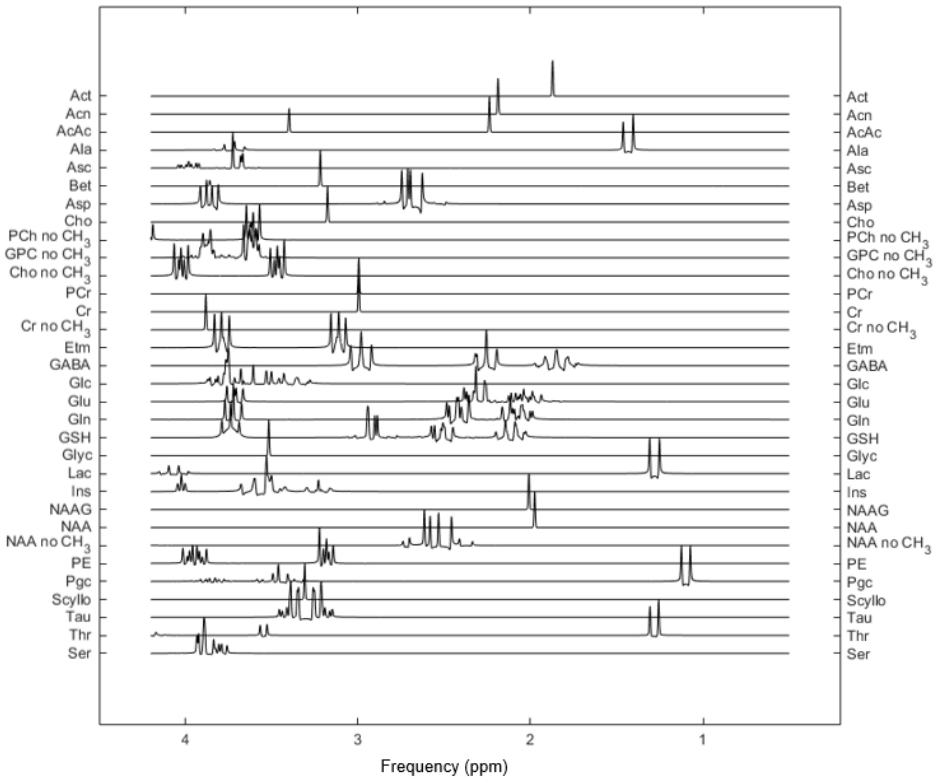

Figure A4.1: Basis set for STEAM TM/TE=20/20ms (Philips).

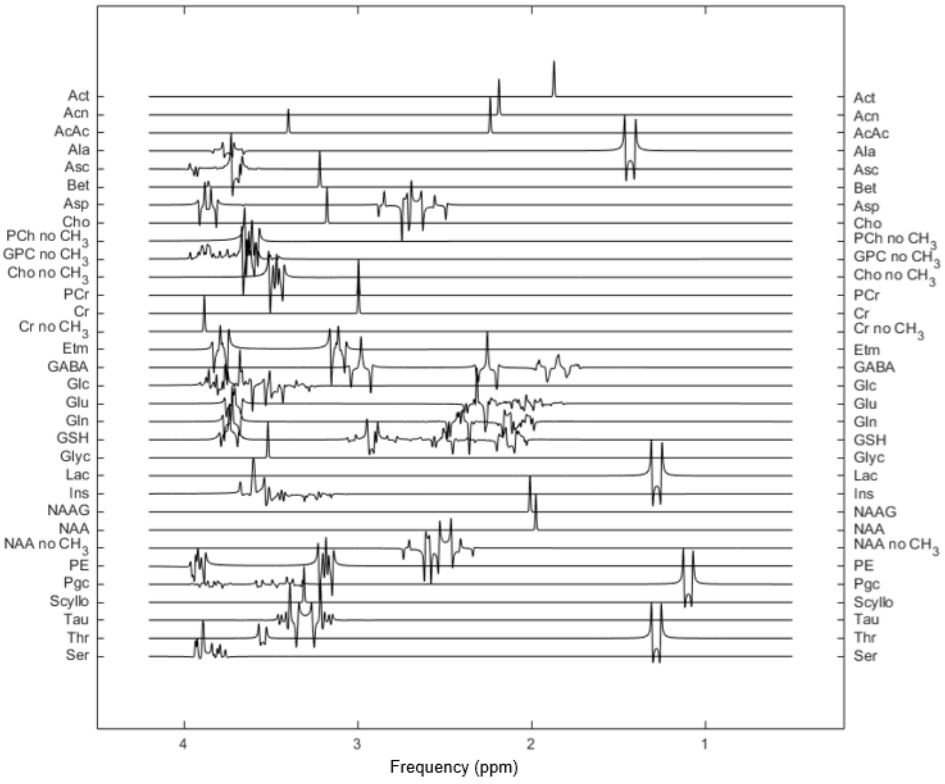

Figure A4.2: Basis set for PRESS TE=60ms (Philips).

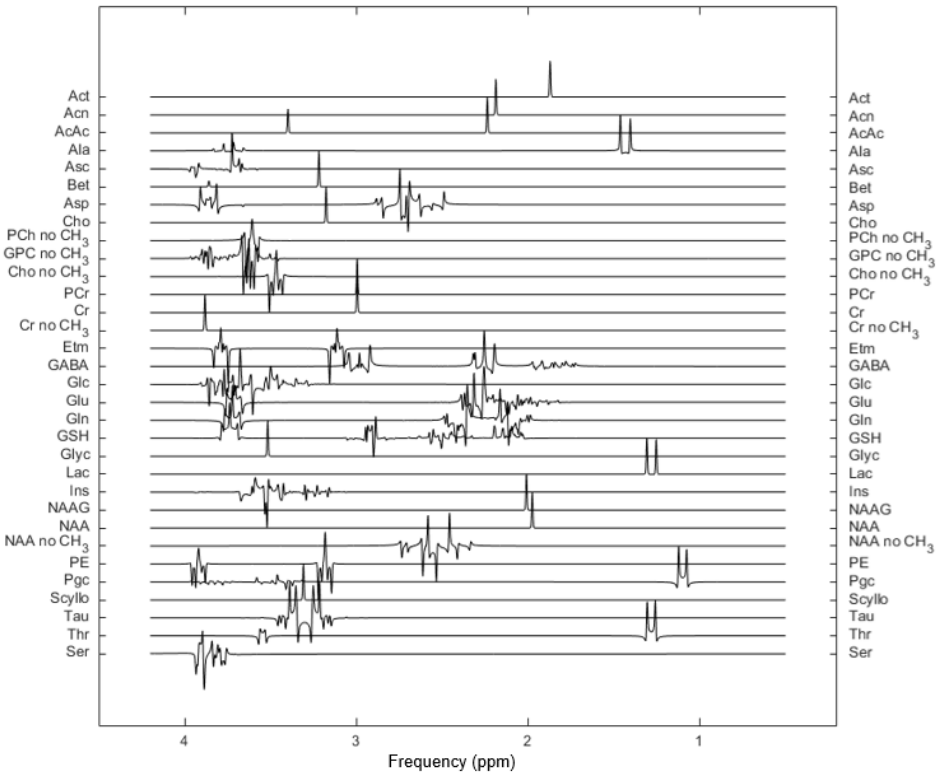

Figure A4.3: Basis set for PRESS TE=288ms (Philips).

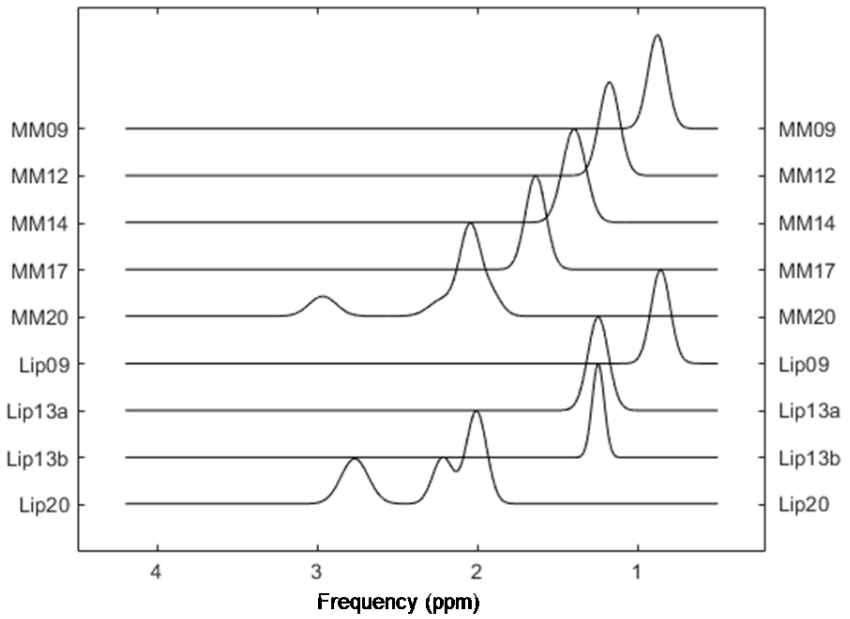

Figure A4.4: Basis set for Provencher's macromolecule and lipid parameterisation.

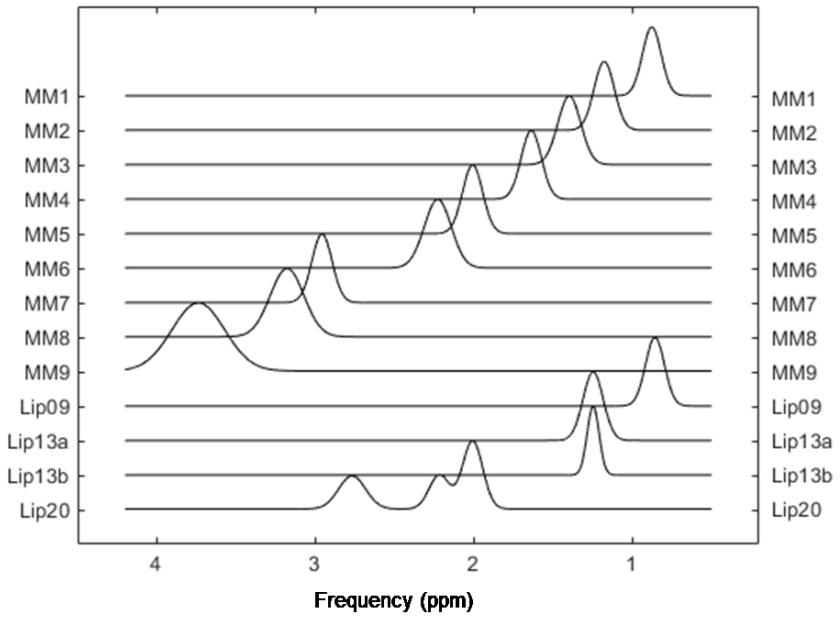

Figure A4.5: Basis set for the nine-component macromolecule parameterisation, including Provencher's parameterisation for the lipid components.

APPENDIX V: VOXEL POSITIONING CHECK FOR *IN-VIVO* DATA

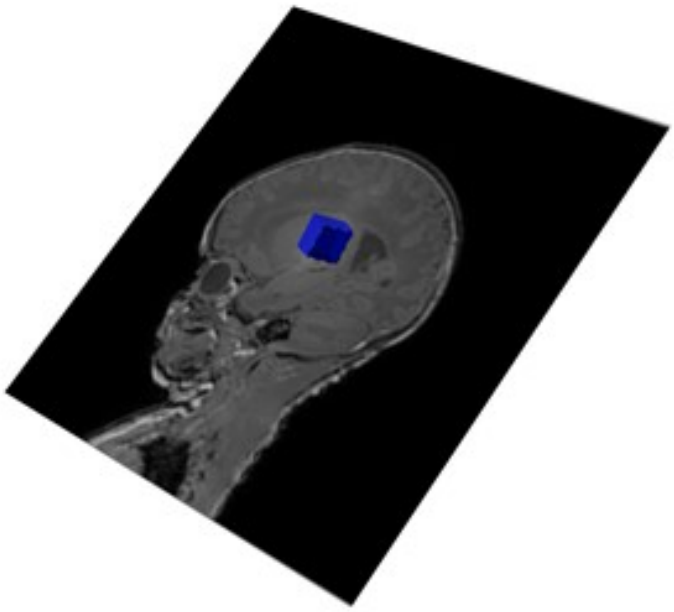

Figure A5: Example of the voxel visualisation software used to check positioning. I developed this tool in Python to be applicable to MRI and MRS data from Philips, Siemens and GE scanners.

APPENDIX VI: SENSITIVITY ANALYSIS FOR GE CORRECTION FACTOR

MRS measure	Normal	Abnormal	P-value	AUC (95% CI)
[tNAA]	7.1 (0.8)	4.0 (1.3)	**<0.001**	0.99 (0.94-1.00)
[tCr]	10.8 (1.1)	7.0 (2.4)	**<0.001**	0.98 (0.93-1.00)
[tCho]	2.8 (0.4)	2.0 (0.7)	**<0.001**	0.87 (0.77-0.97)

Table A6.1: Entries from Table 4.2 as given, including correction factor for GE concentrations

MRS measure	Normal	Abnormal	P-value	AUC (95% CI)
[tNAA]	7.0 (0.8)	4.0 (1.3)	**<0.001**	0.99 (0.94-1.00)
[tCr]	10.7 (1.1)	6.9 (2.5)	**<0.001**	0.96 (0.91-1.00)
[tCho]	2.8 (0.4)	2.0 (0.7)	**<0.001**	0.85 (0.74-0.96)

Table A6.2: Updated entries from Table 4.2, without applying the GE correction factor to concentrations

MRS measure	R^2 change	P-value	Coefficient
[tNAA]	0.55	**<0.001**	11.9
[tCr]	0.36	**<0.001**	7.1
[tCho]	0.32	**<0.001**	24.1

Table A6.3: Entries from Table 4.4 as given, including correction factor for GE concentrations

MRS measure	R^2 change	P-value	Coefficient
[tNAA]	0.55	**<0.001**	11.8
[tCr]	0.34	**<0.001**	6.5
[tCho]	0.30	**<0.001**	22.7

Table A6.4: Updated entries from Table 4.4, without applying the GE correction factor to concentrations