

**Predicting recovery of ability to walk ten meters or more
independently at eight weeks and six months after stroke**

John Pearn

Submitted in accordance with the requirements for the degree of
Doctor of Medicine

The University of Leeds
Leeds Institute of Rheumatic and Musculoskeletal Medicine
Section of Rehabilitation Medicine

June, 2018



I dedicate this thesis to the memory of my supervisor,

Professor Bipinchandra Bhakta

BSc(Hons), MBChB, MD, FRCP

1960 – 2014

Emeritus Professor of Rehabilitation Medicine, University of Leeds

“How hard can it be?”

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others. This research has been carried out by a team which has included: Professor Bipin Bhakta and Professor Gary Ford, Chief Investigators; Professor Amanda Farrin, Professor of trial methodology and lead trial statistician; Ms Suzanne Hartley, head of trial management; Dr Tufail Patankar, consultant in interventional neuroradiology; Dr Jeremy Macmullen-Price, consultant in neuroradiology; and Ms Bonnie Cundill, trial statistician.

My own contributions, fully and explicitly indicated in the thesis, have been: to script, film, edit, and produce a patient information DVD; to prepare and deliver training material for investigators at recruiting centres; to support (as a co-investigator) the recruitment and randomisation of participants in Leeds Teaching Hospitals NHS Trust; to maintain an annual systematic literature review to ensure that no new trials had been published that might call into question the safety or clinical equipoise of the trial; to provide support to the Chief Investigator in categorising and responding to reports of adverse events received from recruiting sites; to develop processes for the despatch of scan images to the Clinical Trials Research Unit; to report (in conjunction with neuroradiology colleagues) CT and MRI scans from trial participants; and to perform the statistical analysis presented below. The candidate is also the author of the first chapter (“Introduction and scientific background”) of the trial monograph, which is currently in preparation for submission to the funding body. Elements of this Chapter were drawn (in edited form) from Chapter 1 of this Thesis.

The other members of the group and their contributions have been as follows. The late Professor Bipin Bhakta was the intellectual originator of the DARS trial, and together with Professor Amanda Farrin drafted the trial protocol. Professor Gary Ford was appointed as Chief Investigator following the retirement of Professor Bhakta. Ms Suzanne Hartley was the trial manager at the University of Leeds Clinical Trials Research Unit. Dr Tufail Patankar and Dr Jeremy Macmullen-Price provided expert input in to the layout of the radiology reporting case report form, training for the candidate in interpreting CT imaging, and co-reviewed all scans jointly with the candidate. Ms Bonnie

Cundill and Professor Amanda Farrin provided training and advice to the candidate on the principles of regression modelling. In addition, the standardised proforma used to analyse the scans was developed by Professor Joanna Wardlaw of the Edinburgh Brain Research Imaging Centre. It was used in here with her kind permission.

The trial was sponsored by the University of Leeds Clinical Trials Research Unit (CTRU), which provided support in the following areas: authorship of the trial protocol; submission of the initial ethics application and subsequent substantial amendments; recruitment and setup of local recruiting centres; design of trial paperwork, including case report forms; liaison with medication suppliers; maintaining the 24-hour randomisation telephone hotline; design of databases for data entry; receipt of patient case report forms, including outcome measures; categorising and notification to statutory bodies of reports of adverse events; and analysis of primary outcome data (influence of levodopa on motor outcome after stroke). Independent oversight was provided by an independent Data Monitoring and Ethics Committee, who periodically reviewed unblinded trial data and reports of adverse events.

The trial was additionally supported by a network of local collaborators at individual recruiting centre. They are too numerous to list here by name, but at each centre would typically include: a local Principal Investigator (a clinician with overall responsibility for the conduct of the trial at that centre, including patient recruitment and the day-to-day clinical management of trial participants); research nurses (who were responsible for randomising patients in to the trial and for the collection of outcome measures at baseline and at each follow-up visit); and local nursing and therapy staff (who were responsible for ensuring that the trial intervention was delivered in a timely manner and for the accurate recording of therapy sessions completed whilst taking the trial drug).

The DARS project was awarded by the Efficacy and Mechanism Evaluation (EME) Programme (Grant reference number 08/43/61) and is funded by the Medical Research Council (MRC) and managed by the National Institute for Health Research (NIHR) on behalf of the MRC-NIHR partnership. The views

expressed in this publication are those of the author(s) and not necessarily those of the MRC, NHS, NIHR or the Department of Health.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement. The right of Dr John Pearn to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

© 2018 The University of Leeds and Dr John Pearn

Acknowledgements

I would like to thank the following individuals and organisations for their assistance in completing this research and preparing this thesis:

My supervisors, Professor Amanda Farrin, and Professor Alan Tennant for their guidance and advice throughout this study.

Dr Tufail Patankar and Dr Jeremy Macmullen-Price: for providing their expert opinion on the configuration and layout of the imaging review case report forms, for teaching me how to interpret CT and MRI imaging, and for providing an expert co-review of scans.

Professor Joanna Wardlaw, of the Brain Research Imaging Centre, University of Edinburgh: for her kind permission to use here the imaging analysis template she had developed for the Third International Stroke Trial (IST-3).

The trials team at the Clinical Trials Research Unit, University of Leeds: for their support in obtaining and archiving CDs containing scan images, entering scan findings on to the trials database, and providing access to other trials data needed for these analyses. My thanks in particular to Bonnie Cundill, trial statistician, for her one-to-one tutorials on the principles of logistic regression modelling. Also to Suzanne Hartley for making available CTRU time and resources to support this work. To Andrew Carter, for his assistance in obtaining imaging from recruiting centres. And to Alison Ferguson and the data management team, for entering data from the imaging analysis case report forms on to the trial database.

The National Institute for Health Research: not only for their funding of the DARS trial (as part of the Efficacy and Mechanism Evaluation Programme), but also for subsequently appointing me to an Academic Clinical Fellow post in 2014.

Dr Collete Isabel Stadler, Academic Clinical Fellow, University of Cambridge: for her kind assistance in translating the abstract in to French.

In the course of preparing this work, I have benefited greatly from the wisdom and support of three generations of Charterhouse Professors of Rehabilitation Medicine. Each have, in their own way, made a unique and significant contribution both to this thesis and to my development as a researcher and as a physician.

The late Professor Bipin Bhakta was the Chief Investigator and intellectual originator of the DARS trial, and was instrumental in its design and opening to recruitment. It was he who invited me to join the trial in 2010, and until his retirement in 2013 he was my lead supervisor. He was an inspiration to me and to many early-career researchers in the Academic Department of Rehabilitation Medicine. Bipin had an enquiring mind, and a meticulous eye for detail. These traits served him well not only as one of the pre-eminent rehabilitation researchers of his day, but also as a first-rate physician. Above all he was a man of compassion, who took a keen interest in the careers and family lives of juniors fortunate enough to work alongside him. He was much loved by his colleagues and patients alike, and his untimely death in 2014 was an enormous loss to this department. I hope that this thesis, and those of others he mentored and supervised, stand as a fitting tribute to him.

Professor Bhakta was succeeded to the Charterhouse chair by Professor Rory O'Connor. Rory took over the stewardship of both the Academic Department of Rehabilitation Medicine and also of NHS rehabilitation services in Leeds at a time of enormous turmoil and change. With good humour and far-sighted vision he has over the past few years built upon the strong foundation that Bipin laid, and has presided over not only an expansion of NHS rehabilitation services in Leeds but also a thriving and happy Academic Department. I am indebted to him for his mentorship and guiding hand in both my research and clinical careers. I also gratefully acknowledge the time he has spent in supervising me, and his kindness in taking on the role of lead supervisor following Bipin's retirement. I trust that in the years ahead the department will go from strength to strength under his leadership.

In noting the achievements of the current and most recent incumbents of the Charterhouse Chair, it must be acknowledged that the Academic Department of Rehabilitation Medicine (or indeed, any of the current clinical rehabilitation

services in Leeds) would never have come in to being were it not for Professor M. Anne Chamberlain. Anne was the first Charterhouse Professor of Rehabilitation Medicine. Although now nominally “retired”, she remains very much at the centre of life in our department and she continues to inspire young rehabilitation physicians both here and in Madagascar by her example. Anne is incredibly generous with her time and wisdom. For the past eight years I have benefited enormously from her experience and sage advice, and her wealth of anecdotes have often been a source of inspiration, hope, and humour. I am particularly grateful to her for taking the time to read and critique this thesis, despite her busy timetable and her having no formal role in my supervision.

Finally I wish to acknowledge the love and support of my parents, Ann and Terry; but most of all that of my wife, Melanie, and my children, Oliver, Thomas, and Emily. For without them, I am nothing.

Abstract

Background: Models predicting return to walking after stroke might allow accurate prognostication, and planning of rehabilitation. This study aimed to construct prognostic models to predict ability to walk 10m or more independently at 8weeks and 6months after stroke.

Method: All participants (N=593) were enrolled in the “Dopamine Augmented Rehabilitation in Stroke” trial, and were unable to walk 10m or more independently at baseline.

Imaging predictor variables (from the first available CT scan) included where relevant: infarct size and location; vascular territory affected; haematoma location and volume; presence of atrophy, white matter hypodensities, old stroke lesions, mass effect, or hydrocephalus. Demographic variables included age, gender, and Oxford Community Stroke Project syndrome. Clinical outcomes recorded at baseline, 8weeks, and 6months included: the Rivermead Mobility Index (RMI); the Montreal Cognitive Assessment (MoCA), General Health Questionnaire-12; presence of musculoskeletal pain.

Using forward stepwise binary logistic regression, six models were constructed: models 1 and 2 (walking ability after ischaemic stroke); models 3 and 4 (walking ability after intracerebral haemorrhage); and models 5 and 6 (walking ability in the whole DARS sample). Models 1-4 utilised imaging, demographic, and clinical predictors; models 5 and 6 included demographic and clinical predictors only.

Results: No imaging variables predicted outcome. Baseline RMI was most consistently predictive across all models. Baseline MoCA was also predictive, but with a smaller effect size than RMI. Only 68%-73% of patients were correctly classified by the models. The percentage of variance they explained was modest (20-30%).

Discussion: This exploratory analysis utilised existing data, excluding predictors that might have explained additional variance. Within these limitations, this study suggests that initial level of mobility offers a more useful

prediction of mobility at up to 6months than assessment of structural brain impairment on CT.

Résumé Français

Contexte: les modèles prédisent que le retour à la marche après un accident vasculaire cérébral pourraient permettre un pronostic précis et la planification de la réadaptation. Cette étude visait à construire des modèles pronostiques pour prédire la capacité des patients à marcher 10 mètres ou plus indépendamment à 8 semaines et 6 mois après un accident vasculaire cérébral.

Méthode: Tous les participants (N = 593) ont été inscrits dans le procès "Dopamine Augmented Rehabilitation in Stroke", et ont été incapables de marcher 10 mètres indépendamment au départ. Les variables prédictives d'imagerie (à partir de la première tomodensitométrie disponible) incluses étaient significatives, et, le cas échéant, la taille et l'emplacement de l'infarctus; Territoire vasculaire affecté; Emplacement et volume de l'hématome; Présence d'atrophie, hyperintensités de la matière blanche, vieilles lésions d'accident vasculaire cérébral (AVC), effet de masse ou hydrocéphalie. Les variables démographiques comprenaient l'âge, le sexe et le syndrome de Oxford Community Stroke Project. Les résultats cliniques enregistrés à la ligne de base, 8 semaines et 6 mois comprises: Rivermead Mobility Index (RMI); Évaluation cognitive de Montréal (MoCA); Santé générale. Questionnaire-12; Présence de douleurs musculo-squelettiques. En utilisant la régression logistique binaire progressive par étapes, six modèles ont été construits: modèles 1 et 2 (capacité de marche après accident vasculaire ischémique); Modèles 3 et 4 (capacité de marche après une hémorragie); Et les modèles 5 et 6 (capacité de marche dans l'ensemble de l'échantillon DARS). Les modèles 1-4 utilisaient des prédicteurs d'imagerie, démographiques et cliniques; Les modèles 5 et 6 incluaient les prédicteurs démographiques et cliniques seulement.

Résultats: aucune variable d'imagerie n'a prédit de résultat. Le RMI de référence était la variable prédictive la plus constante dans tous les modèles. Le MoCA de base était également prédictif, mais avec une taille d'effet plus petite que le RMI. Seulement 68% -73% des patients ont été correctement

classés par les modèles. Le pourcentage de variance qu'ils ont expliqué était modeste (20 à 30%).

Discussion: Cette analyse exploratoire a utilisé les données existantes, à l'exclusion des prédicteurs qui auraient expliqué une variance supplémentaire. Dans le cadre de ces limites, cette étude suggère que le niveau initial de mobilité offre une prédiction plus utile de la mobilité jusqu'à 6 mois que l'évaluation de l'insuffisance cérébrale structurelle sur la tomodensitométrie.

Traduit par Dr Collette Isabel Stadler, Academic Clinical Fellow, University of Cambridge.

Table of Contents

Acknowledgements	iv
Abstract	vii
Résumé Français	ix
Table of Contents	xi
List of Tables	xv
List of Figures	xvii
Chapter 1. Introduction	1
Part 1.1 Stroke in context.....	1
1.1.1. The epidemiology of stroke.....	1
1.1.2. Defining “stroke”.....	3
1.1.3. What constitutes optimal care for people with stroke?.....	5
1.1.4. What is “rehabilitation”?.....	14
1.1.5. The trajectory of recovery after stroke.....	17
1.1.6. What rehabilitation interventions are effective after stroke?.....	18
1.1.7. The biological basis of rehabilitation.....	21
Part 1.2 Motor learning and recovery from stroke.....	23
1.2.1. Motor learning after stroke.....	23
1.2.2. Cognitive dysfunction after stroke.....	29
1.2.3. Dopamine augmentation of rehabilitation in stroke: a theoretical background.....	40
1.2.4. Other impairments that might influence recovery from stroke.....	44
1.2.5. Might an assessment of brain structure contribute to predicting prognosis in rehabilitation?.....	46
Part 1.3 Standardised image analysis instruments.....	48
1.3.1. Mechanisms of brain injury in stroke.....	48
1.3.2. Mechanisms of brain injury in ischaemic stroke.....	48
1.3.3. Mechanisms of brain injury in ICH.....	50
1.3.4. The need for a standardised system to code scan findings.....	54
1.3.5. Standardised systems for coding ischaemic stroke.....	55
1.3.6. Standardised systems for coding ICH.....	61
Part 1.4 Developing a prognostic model to predict recovery of mobility after stroke.....	65

1.4.1. The need to predict specific rehabilitation outcomes.....	65
1.4.2. The challenges of developing prognostic models.....	66
1.4.3. Aims of this Thesis	69
Chapter 2. Methods	71
Part 2.1 Methods for the DARS trial	71
2.1.1. What was the DARS trial?.....	71
2.1.2. Trial setup.....	71
2.1.3. Participants.....	73
Part 2.2 Treatment and follow-up	76
2.2.1. Timing of assessments.....	76
2.2.2. Summary of outcome measures used in the DARS trial ...	77
2.2.3. Assessment at T ₀ and treatment regimen	79
2.2.4. Follow-up of participants	81
Part 2.3 Outcome measurement in the DARS trial	84
2.3.1. The principles of measurement	84
2.3.2. Outcome measures used in the DARS trial	85
2.3.3. Interpreting the output of rehabilitation outcome measures	93
Part 2.4 Analysis of neuroimaging.....	96
2.4.1. Analysis of imaging in the DARS trial	96
2.4.2. Coding scan findings	99
Part 2.5 Statistical analysis procedures.....	102
2.5.1. An overview of regression modelling.....	102
2.5.2. Modelling walking ability at T1 and T2 in the DARS data- set	107
2.5.3. Treatment of predictor variables.....	109
2.5.4. Constructing the models: general considerations.....	113
2.5.5. Procedure for constructing the models	115
2.5.6. Evaluating the models: assumption testing	117
2.5.7. Summarising the models.....	120
2.5.8. Testing assumptions made for missing outcome data.....	120
Chapter 3. Results	123
Part 3.1 Summary of recruitment.....	123
3.1.1. Summary of recruitment, and sample characteristics	123
3.1.2. The modelling process	129
Part 3.2 Modelling walking ability in IWS group.....	136

3.2.1. Characteristics of the IWS group.....	136
3.2.2. Model 1: return to walking at T ₁ in IWS group	136
3.2.3. Model 2: return to walking at T ₂ in IWS group	144
Part 3.3 Modelling walking ability in intracerebral haemorrhage	153
3.3.1. Characteristics of the HWS group	153
3.3.2. Model 3: return to walking at T ₁ in HWS group	153
3.3.3. Model 4: return to walking at T ₂ in HWS group.....	160
Part 3.4 Modelling walking ability in the whole DARS sample.....	168
3.4.1. Characteristics of the DARS sample	168
3.4.2. Model 5: return to walking at T ₁ in the whole DARS sample	169
3.4.3. Model 6: return to walking at T ₂ in the whole DARS sample	175
Chapter 4. Discussion.....	183
Part 4.1 Summary of results.....	183
4.1.1. The aims of this study	183
4.1.2. Summary of models	183
4.1.3. The need for formal validation of the models	187
Part 4.2 Comparing the findings of this study with previous literature	190
4.2.1. Variables that were predictors of walking ability in the models presented here	190
4.2.2. Clinical predictor variables that did not make a contribution to models 1-6.....	194
4.2.3. Imaging predictor variables	197
Part 4.3 Limitations of this Thesis	202
4.3.1. Analytical considerations.....	202
4.3.2. Imaging considerations	212
Part 4.4 Directions for future work.....	217
4.4.1. The need for prognostic modelling in stroke.....	217
4.4.2. The need for robust outcome measurement in stroke research	223
Part 4.5 Concluding remarks.....	230
4.5.1. Potential future uses of outputs from this Thesis.....	230
4.5.2. The implications of this Thesis for clinical practice and research	233

References	237
List of abbreviations	269
Appendix A. Table summarising prognostic models for predicting outcome following ICH	273
Appendix B. The AISCT template for coding ischaemic stroke lesions	283
B.1. AISCT template for coding acute ischaemic change in the MCA territory	284
B.2. AISCT template for lacunar lesions and border zone ischaemia	285
B.3. AISCT template for tissue swelling	286
Appendix C. Summary of standardised outcome measures used in the DARS trial	287
Appendix D. The Rivermead Mobility Index	289
Appendix E. Flow diagram summarising the process of image analysis	291
Appendix F. CT Image Interpretation Case Report Form	293

List of Tables

Table 2.1. Summary of models presented for the “primary infarction, with scan available” (IWS), “primary intracerebral haemorrhage, with scan available” (HWS) and “whole DARS sample” groups.	107
Table 2.2. Classification of infarct size.....	111
Table 2.3 Example classification table, illustrating calculation of sensitivity, specificity, and positive and negative predictive values. ..	120
Table 2.4. Summary of alternative assumptions for missing data that were tested.	121
Table 3.1. Demographic characteristics and clinical impairment in the DARS sample at T ₀	128
Table 3.2. Correlation matrix for predictor variables	134
Table 3.3. Univariate predictors of independent walking ability at T ₁	138
Table 3.4. Univariate models of ability to walk 10m or more independently at T ₁	139
Table 3.5. Final version of Model 1.	142
Table 3.6. Classification table for Model 1.	142
Table 3.7. Properties of Model 1 when fitted under alternative assumptions for missing SR-RMI scores at T ₁	143
Table 3.8. Univariate predictors of independent walking ability at T ₂	145
Table 3.9. Univariate models of independent walking ability at T ₂	146
Table 3.10. Final version of Model 2.....	149
Table 3.11. Classification table for Model 2	150
Table 3.12. Properties of Model 2 when fitted under alternative assumptions for missing SR-RMI scores at T ₂	151
Table 3.13. Univariate predictors of independent walking ability at T ₁	155
Table 3.14. Final iteration of Model 3.....	157
Table 3.15. Classification table for Model 3.	158
Table 3.16. Properties of Model 3 when fitted under alternative assumptions for missing SR-RMI scores at T ₁	159
Table 3.17. Univariate models of independent walking ability at T ₂	161
Table 3.18. Univariate predictors of independent walking ability at T ₂	161
Table 3.19. Final version of Model 4.	164
Table 3.20. Classification table for Model 4.	165
Table 3.21. Properties of Model 4 when fitted under alternative assumptions for missing SR-RMI scores at T ₂	166
Table 3.22. Univariate predictors independent walking ability at T ₁	169
Table 3.23. Final version of Model 5.	172

Table 3.24. Classification table for Model 5.	173
Table 3.25. Properties of Model 5 under alternative assumptions for missing SR-RMI scores at T ₁	174
Table 3.26. Univariate models of independent walking ability at T ₁	176
Table 3.27. Final version of Model 6	178
Table 3.28. Classification table for Model 6.	179
Table 3.29. Properties of Model 6 when fitted under alternative assumptions for missing SR-RMI scores at T ₂	180
Table A.1. Summary of prognostic models for predicting outcome following ICH.	273
Table C.1. Standardised outcome measures used in the DARS trial.....	288

List of Figures

Figure 1.1. The WHO International Classification of Functioning, Disability, and Health (ICF)	14
Figure 1.2. Schematic illustration of the architecture of the cortico-basal loop circuits	26
Figure 1.3. Schematic illustration of the architecture of the subcortical loop circuits.	27
Figure 1.4. schematic representation of the overlap between syndromes of cognitive impairment.	31
Figure 1.5. Mechanisms of injury to cortico-basal and subcortical loop circuits.	32
Figure 2.1. Summary of timeline for screening, recruitment, treatment, and follow-up of DARS participants.....	77
Figure 2.2. Simplified algorithm for assigning mRS score.	87
Figure 2.3. The musculoskeletal signs, symptoms, and pain manikin (MSK-SSP)	89
Figure 2.4. Linear relationship between a continuous predictor variable, x , and a continuous outcome variable, Y	103
Figure 2.5. Algorithm for dichotomising RMI scores.	108
Figure 3.1. Summary of patients screened, eligible, recruited, randomised, and followed up in the DARS trial.....	124
Figure 3.2. Summary of the availability of brain imaging for DARS participants.....	126
Figure 3.3. Example of a weak correlation between two predictor variables.....	130
Figure 3.4. Example of a moderate correlation between two predictor variables.....	131
Figure 3.5. Example of a strong correlation between two predictor variables.....	132
Figure 3.6. Example of a negative correlation between two predictor variables.....	133
Figure E.1. Summary of the process of imaging analysis in the DARS trial.	291

Chapter 1. Introduction

Part 1.1 Stroke in context

1.1.1. The epidemiology of stroke

Every two seconds, someone in the world sustains a stroke for the first time (The Stroke Association, 2017c). Around 85% of strokes result from an occlusion of an artery, with the remaining 15% resulting from a bleed in to the brain parenchyma (Feigin et al., 2014).

Although the overall incidence of stroke in the UK has reduced by 19% between 1990 and 2010, this still equates to around 100,000 new cases a year or roughly one person every 5 minutes (The Stroke Association, 2017c). Women tend to be slightly older than men at the time of first stroke, with a mean age of 80 (versus 74 for men) in England, Wales, and Northern Ireland, and 76 (versus 71 for men) in Scotland (The Stroke Association, 2017c). However, around 25% of strokes happen in adults of working age. Although the peak incidence of stroke remains in the over-70s age group, in England the proportion of strokes sustained by those aged 40-69 has risen from 33.7% in 2007 to 38.2% in 2016 (Public Health England, 2018). This is significant, as those who survive a stroke at a younger age might be expected to spend a greater number of years of their lives living with disability and in many cases requiring carer support.

The number of deaths attributable to stroke in the UK have almost halved between 1990 and 2010; however, it remains the UK's fourth largest cause of death accounting for 7% of all deaths overall (The Stroke Association, 2017c). There are a greater number of stroke-related deaths in women (8% of female deaths) than in men (6% of all male deaths); presumably due in part to the longer life expectancy of women, and the fact that they tend to be older (and thus more frail) at the time of stroke (The Stroke Association, 2017c). The greatest mortality from stroke is within the first 30 days, with around one

person in eight who has a stroke dying within this time (The Stroke Association, 2017c).

Those who survive a stroke are often left with profound impairments. There are 1.2million stroke survivors in the UK, of whom: 75% have arm and/ or leg weakness; 50% have problems with bladder control; 45% have swallowing problems; 30% have aphasia; and 20% have long-term visual problems (The Stroke Association, 2017c). The Auckland Stroke Outcomes Study found that after five years, 15% of stroke survivors were living in institutional care (Feigin et al., 2010).

1.1.1.1. The wider impact of stroke

In 2010, the direct cost of stroke care to the UK's National Health Service (NHS) was estimated at around £3billion annually (Department of Health, 2010). A more recent study by The Stroke Association (2017a) estimated that stroke care cost the NHS £3.5billion in 2015, and forecast that this figure could rise to £10.2billion by 2035. The annual cost of stroke care to UK society as a whole is £25.6billion (The Stroke Association, 2017b). This equates to an average societal cost of £45,409 per patient in the first year, and £24,778 per patient per year thereafter (The Stroke Association, 2017b). Around £15.8billion of this £25.5billion annual cost is the value of "informal" or unpaid care provided to stroke survivors by family members and friends (The Stroke Association, 2017b). The cost of lost economic productivity is more modest by comparison; around £1.6billion per year (The Stroke Association, 2017b).

The emotional cost to stroke survivors and their families is, of course, substantial. A recent survey of stroke survivors by the Stroke Association found that 67% had experienced anxiety and 59% had felt depressed: but two-thirds did not feel that their emotional needs were adequately addressed (The Stroke Association, 2013). A similar proportion of partners reported relationship strain, and one-in-ten had either ended their relationship or had considered doing so (The Stroke Association, 2013). Rates of anxiety and depression amongst carers were comparable to those seen in stroke survivors themselves, at 79% and 56% respectively.

1.1.1.2. The global impact of stroke

On a global scale, stroke accounts for 6.7million deaths per year: almost one-third of the total number of deaths worldwide that are attributable to cardiovascular disease (The World Health Organisation, 2014c). Another way of conceptualising the impact of stroke is to measure the number of years of healthy life lost to this condition each year (either by death, or by survival with disability): a concept termed “Disability-Adjusted Life Years” (DALYs; The World Health Organisation (2014b)). Viewed in these terms, in 2012 stroke accounted for the loss of over 141million years of healthy life worldwide (The World Health Organisation, 2014b). Worryingly, the incidence of stroke is projected to increase, due to a general aging of the world population (The World Health Organisation, 2014a) and an increased prevalence of modifiable risk factors such as hypertension, diabetes, tobacco use, and obesity (Mendis, 2013).

Internationally, the burden of stroke is not evenly borne. Mortality and disability rates vary ten-fold between the most- and least-affected countries (Johnston et al., 2009b). Regions with the highest mortality are Eastern Europe, North Asia, central Africa, and the South Pacific (Johnston et al., 2009b). The countries most profoundly affected are those with low- and middle-incomes, which account for 85% of all strokes each year, and which bear 87% of all the DALYs lost to stroke (Johnston et al., 2009b). The period between 1970 and 2008 has seen a 42% *fall* in stroke incidence in high income countries, but a 100% *increase* in incidence in low- and middle-income nations (Johnston et al., 2009b). In the absence of effective acute stroke services (as is the case in many developing countries), 62% of those who sustain a stroke will be dead or dependent at six months (Johnston et al., 2009b).

1.1.2. Defining “stroke”

This Thesis is set within the context of a large randomised controlled trial of a novel stroke rehabilitation intervention (DARS; Dopamine Augmented Rehabilitation in Stroke, ISRCTN99643613) (Bhakta et al., 2014). The purpose and methods of this trial will be described later, but at present it must be noted that participants were enrolled on the basis of the World Health

Organisation's *clinical* criteria (Bhakta et al., 2014), which define "stroke" as "rapidly developed clinical signs of focal (or global) disturbance of cerebral function, lasting more than 24 hours or leading to death, with no apparent cause other than of vascular origin" (Aho et al., 1980). Although widely adopted, this definition has been criticised (Sacco et al., 2013). Firstly, stroke rarely results in "global" cerebral dysfunction (Sacco et al., 2013). Secondly, the 24-hour timeframe that distinguishes "stroke" from "transient ischaemic attack" was arbitrarily defined (Sacco et al., 2013). As imaging technology has matured, it has become apparent that patients with a "transient" clinical deficit may have signs of established infarction on scans; conversely, chronic ischaemia may cause persistent clinical manifestations in the absence of radiographic evidence of an infarct (Sacco et al., 2013). Finally, the WHO definition of stroke makes no reference to the underlying pathophysiology of this condition. It is therefore important to distinguish between a patient's clinical presentation ("stroke"), and terms that describe pathophysiological processes ("ischaemic stroke" or "haemorrhage") or scan findings ("infarct" or "haematoma").

The language used in the literature to describe a stroke syndrome attributable to haemorrhage is sometimes confusing. The term "intracerebral haemorrhage" is sometimes used to cover both spontaneous events and parenchymal bleeding due to trauma, and may also encompass subarachnoid haemorrhage (Sacco et al., 2013). The alternative expression "haemorrhagic stroke" is ambiguous since it may denote a primary haemorrhage or haemorrhagic transformation of an underlying infarct (Sacco et al., 2013). This Thesis will adopt the term "intracerebral haemorrhage" (ICH), defined as "a focal collection of blood within the brain parenchyma or ventricular system that is not caused by trauma" (Sacco et al., 2013). Patients with primary subarachnoid haemorrhage will not be considered here. It is important to note that the definition of ICH as stated above includes both primary ICH and confluent parenchymal haematomas which arise as a result of haemorrhagic transformation of an underlying infarct (Sacco et al., 2013). The case for including these lesions is that the management of a confluent secondary ICH is analogous to a spontaneous primary ICH (Sacco et al., 2013). Although the term "intracerebral" strictly refers to a process occurring within the cerebral

hemispheres, for simplicity the expression “ICH” will also be used here to encompass spontaneous haemorrhage in to the parenchyma of the cerebellum or brain stem.

1.1.3. What constitutes optimal care for people with stroke?

The care of people with stroke focuses initially upon minimising the extent of tissue injury, preventing secondary complications, and identifying and treating modifiable risk factors. Subsequently, multidisciplinary rehabilitation seeks to restore or compensate for lost functions, with the overall aim of maximising independence and autonomy. The principles of rehabilitation will be examined in detail below: but first, it is perhaps worth reviewing in brief what constitutes “optimal” acute care in stroke. In doing so, reference will be made to the most recent UK guidelines for stroke care (The Royal College of Physicians, 2016), whilst acknowledging that practice may vary between nations depending upon local resources and policies.

1.1.3.1. Raising public awareness of stroke symptoms

Optimum care perhaps begins with the patient themselves recognising the symptoms of an acute stroke, and presenting promptly to medical services. High-profile public health campaigns, such as the UK Department of Health’s recent “Act FAST” initiative (Public Health England Campaigns Resource Centre, 2009), use multi-media marketing techniques to raise public awareness of stroke symptoms and to emphasise the importance of seeking medical help promptly (Dombrowski et al., 2013). Nevertheless, a review of several such campaigns across a number of countries concluded that there is no evidence to suggest that they actually change public behaviour (Lecouturier et al., 2010). Although the Act FAST campaign resulted in a short-term fall in delays in seeking and receiving help in those with suspected stroke (Wolters et al., 2015), a survey of general practitioners reported scepticism of its longer-term impact on public behaviour (Dombrowski et al., 2013). Claims made by ministers at the time that the campaign saved 642 people from death or serious disability were based upon modelling conducted by the advertising agency involved in the design of the campaign (Dudley, 2011). These models formed part of a cost-effectiveness (“return on marketing investment”)

analysis, but have been criticised for assuming firstly that thrombolysis is a potentially life-saving intervention (the true aim of this treatment is to prevent disability), and secondly that an implausibly-high 56% of patients who presented to hospital within three hours of a stroke would be eligible to receive this intervention (Dudley, 2011).

1.1.3.2. Acute investigation and management of ischaemic stroke

An ischaemic stroke typically consists of a “core” of irreversibly-infarcted tissue, surrounded by a zone of less-severe ischaemia. This so-called “penumbra” contains tissue that is under oxidative stress, but is potentially salvageable with optimum treatment (Iadecola and Anrather, 2011). The basis of treatment of ischaemic stroke is therefore firstly to minimise the extent of permanent tissue injury, and secondly to maximise salvage of the “ischaemic penumbra”.

The primary purpose of acute imaging in stroke is to differentiate between ischaemic stroke and ICH, to exclude other conditions that might mimic stroke (The Royal College of Physicians, 2016), but also to identify substantial infarcts that carry a greater risk of haemorrhagic transformation (Kaste et al., 1995, Hacke et al., 1998). Non-contrast computerised tomography (CT) scanning performed early after ictus is highly sensitive for acute ICH, and is therefore the first-line investigation of choice (Macellari et al., 2014, Sacco et al., 2013). National clinical guidelines currently recommend that patients with suspected stroke are scanned urgently as soon as possible after presentation, within a maximum of one hour of presentation (The Royal College of Physicians, 2016).

For those with ischaemic stroke, intravenous thrombolysis using the recombinant tissue plasminogen activator alteplase has been shown to reduce the risk of death or dependency at 3-6 months (odds ratio (OR) 0.85, 95% confidence interval (CI) 0.78-0.93; Wardlaw et al. (2014b)). This treatment should be administered as quickly as possible; ideally within 3 hours of ictus if the time of onset of symptoms is known, although treatment may be considered at up to 4.5 hours on a case-by-case basis (The Royal College of Physicians, 2016). Although treatment with alteplase carries an increased risk

of both fatal and non-fatal ICH within the first 7 to 10 days (OR 3.75, CI 3.11-4.51; Wardlaw et al. (2014b)), mortality at six months does not appear to be increased (The Royal College of Physicians, 2016). Furthermore, older patients (aged 80 or over) benefitted equally from treatment when compared those under 80, particularly when the drug was administered within 3 hours of symptom onset (Wardlaw et al., 2014b). Alteplase must be given by staff experienced in stroke thrombolysis. Prior to initiation of treatment the patient's blood pressure must be controlled to less than 185/110mmHg. Following treatment, patients must be cared for on a dedicated hyper-acute stroke unit (The Royal College of Physicians, 2016).

Historically thrombolysis has been the only available means of restoring brain perfusion after ischaemic stroke. However, there is now growing evidence to suggest that endovascular clot retrieval ("thrombectomy") is effective in reducing disability after stroke (Berkhemer et al., 2015, Campbell et al., 2015, Goyal et al., 2015, Goyal et al., 2016, Jovin et al., 2015, Muir et al., 2017, Saver et al., 2015). A meta-analysis of trials enrolling a total of 1287 patients with anterior circulation ischaemic stroke found that for every 2.6 patients treated, one would achieve a reduction in Modified Rankin Scale (mRS) score of at least one point. Treatment effect sizes were similar in several sub-group analyses, including patients aged >80, those randomised more than 300 minutes after symptom onset, and those who were not eligible for intravenous alteplase (Goyal et al., 2016). The inclusion criteria for several of these trials specified a National Institute of Health Stroke Scale (NIHSS) score of 6 or more. The time between symptom onset and thrombectomy varied between trials; significantly, all trials that enrolled patients between 6 and 12 hours after stroke required evidence of viable penumbral tissue on CT perfusion (The Royal College of Physicians, 2016). For patients undergoing thrombectomy within 5 hours of symptom onset, demonstration of a large-vessel occlusion on CT angiography was regarded as sufficient (The Royal College of Physicians, 2016).

The most recent UK guidelines for stroke care recommend that thrombectomy be considered for patients with an NIHSS of 6 or more, who present within 5 hours of onset of anterior circulation symptoms attributable to a proximal

large-vessel thrombus. Patients with anterior circulation symptoms and evidence of viable brain tissue on perfusion imaging may be considered for thrombectomy up to 12 hours after symptom onset. Those with a large-vessel posterior circulation syndrome may be considered for thrombectomy within 24 hours of ictus. In patients presenting within the time window for thrombolysis and with no contraindications to this intervention, a combination of initial treatment with intravenous alteplase followed by thrombectomy may be offered. For those with known contra-indications to thrombolysis but not to clot retrieval, thrombectomy may be considered as a “stand alone” treatment within 5 hours of symptom onset. Although thrombectomy is a promising development in acute stroke care, delivery of this intervention is dependent upon the availability of resources such as operating theatres, interventional radiologists skilled in this procedure, and other support staff. At present, there remain significant barriers to the widespread introduction of this treatment (The Royal College of Physicians, 2016).

1.1.3.3. Secondary prevention of ischaemic stroke

In ischaemic stroke it is recommended that, unless otherwise contra-indicated, high-dose aspirin (300mg) be commenced as soon as possible within the first 24 hours and continued thereafter for two weeks (The Royal College of Physicians, 2016). In those without underlying atrial fibrillation (AF), long term antiplatelet therapy should then be initiated. Clopidogrel 75mg is recommended as first line; if this is not tolerated then combination treatment with aspirin 75mg and modified-release dipyridamole may be used (The Royal College of Physicians, 2016). There is no evidence to support the combined use of aspirin and clopidogrel (The Royal College of Physicians, 2016), nor for the use of other antiplatelet agents such as ticlopidine, cilostazol, satigrel, sarpolgrate, KBT 3022, or isbogrel (Sandercock et al., 2014).

Patients who have sustained an embolic stroke as a result of permanent or paroxysmal AF have an annual risk of recurrent stroke of 12% (Anonymous, 1993). Oral anticoagulants are significantly more effective than antiplatelets in preventing stroke in this group, with around a one-third risk reduction when compared to aspirin alone (Aguilar et al., 2007). Vitamin K antagonists such as warfarin have previously been the drug of choice for stroke prevention in

AF. More recently newer drugs which act to inhibit thrombin or factor Xa (non-vitamin K oral anticoagulants; “NOAC”) have been found to be more effective than warfarin in stroke prevention (relative risk 0.87, 95% CI 0.77 to 0.99), with a lower risk of intracranial bleeding (relative risk 0.49, 95% CI 0.36 to 0.66; Miller et al. (2012)). In acute stroke, there is concern that early initiation of anticoagulation may increase the risk of haemorrhagic transformation of the infarct. Current guidelines therefore recommend a two-week initial course of aspirin 300mg, followed by anticoagulation with either a vitamin K antagonist or (in non-valvular AF) a NOAC (The Royal College of Physicians, 2016). Every effort should be made to identify and mitigate risk factors for bleeding before anticoagulation is initiated. In those judged to be at high risk of bleeding, aspirin alone cannot be considered a safer alternative to anticoagulation (The Royal College of Physicians, 2016).

Control of cardiovascular risk factors is crucial in the secondary prevention of ischaemic stroke. Aggressive control of hypertension is not recommended in the hyperacute stage, due to concerns about further compromising cerebral perfusion. Early treatment of hypertension is therefore only recommended if there is a definite indication to do so; for example, where there is evidence of end-organ injury (hypertensive encephalopathy, nephropathy, or heart failure), or in patients who might be eligible for thrombolysis if their blood pressure can be reduced to <185/110mmHg. For longer-term secondary prevention a target systolic blood pressure (SBP) of <130mmHg is recommended, except in those with severe bilateral carotid stenosis in whom a target of 140-150mmHg is appropriate (The Royal College of Physicians, 2016)

Lipid lowering therapy with atorvastatin 80mg following an ischaemic stroke or transient ischaemic attack has been shown to reduce the risk of a subsequent stroke by 15%, and that of major coronary events by 35% (Amarenco et al., 2006). High-intensity lipid lowering treatment with atorvastatin 20mg-80mg should therefore be initiated as soon as possible after ischaemic stroke, aiming for a 40% reduction in non-high density lipoprotein cholesterol (The Royal College of Physicians, 2016).

1.1.3.4. Acute management of ICH

Patients who sustain an ICH are at risk of deterioration, and must therefore be admitted to a hyper-acute stroke unit for monitoring of their conscious level (The Royal College of Physicians, 2016). For patients who sustain an ICH, treatment aimed at reducing SBP to a target of <140mmHg within the first six hours of stroke has been shown to reduce the risk of haematoma expansion within the first 72 hours (Anderson et al., 2010). For patients with deep ICH involving the basal ganglia, achieving a target SBP of <140mmHg does not appear to reduce the risk of haematoma expansion or death, but there is a possible benefit in reducing the final level of disability (defined by the mRS) and quality of life (Anderson et al., 2013). A more aggressive target of 110-139mmHg does not result in any significant benefit in terms of survival or final level of disability for those with small deep ICH (Qureshi et al., 2016). For those presenting with acute ICH and a SBP>150mmHg, initiation of antihypertensive treatment within 6 hours of stroke onset aiming for a target of <140mmHg is recommended (The Royal College of Physicians, 2016).

Those patients who were anticoagulated with warfarin or a NOAC prior to their stroke should have their anticoagulation reversed as soon as possible. A combination of intravenous vitamin K and prothrombin complex concentrate is recommended for those taking vitamin K antagonists. For those taking Factor Xa inhibitors, treatment with four-factor prothrombin complex is recommended (The Royal College of Physicians, 2016).

The majority of patients with ICH do not require surgery, and should be managed by control of hypertension and reversal of anticoagulation. The exceptions are patients with posterior fossa ICH and those who develop hydrocephalus, for whom surgical intervention may be considered (The Royal College of Physicians, 2016). There is no evidence to suggest that surgical evacuation of lobar ICH reduces the rate of death or disability at 6 months (Mendelow et al., 2013).

1.1.3.5. Evidence for management on an acute stroke unit

For both ischaemic stroke and ICH, the *environment* in which acute care is provided is of crucial importance. The use of dedicated stroke units, staffed

by a skilled multidisciplinary team with expertise in the care of stroke patients, has been shown to reduce significantly the odds of death at follow-up (median 1 year;) OR 0.81, CI 0.69-0.94; $p=0.005$), death or discharge to institutional care (OR 0.78, CI 0.68-0.89; $p=0.0003$), or death or dependency (OR 0.79, CI 0.68-0.90; $p=0.0007$) when compared with care provided on general medical wards or by a “roving” stroke team (Stroke Unit Trialists Collaboration, 2013). Admission to a stroke unit is therefore recommended for all patients with an acute stroke in England (National Institute for Health and Care Excellence, 2013, The Royal College of Physicians, 2016).

1.1.3.6. Recommendations for rehabilitation after stroke

Following initial stabilisation patients typically require a period of multidisciplinary rehabilitation. An initial rehabilitation assessment should be completed as soon as possible after admission to hospital. This should focus on impairments that might affect adversely the safety or comfort of the patient such as swallowing, nutritional status, skin pressure areas, continence, cognition, communication, and cognitive function (National Institute for Health and Care Excellence, 2016). A more detailed rehabilitation assessment should then follow, taking in to account a patient’s previous functional abilities, their impairment in bodily structures and functions, activity limitations and participation restrictions, and relevant environmental factors (National Institute for Health and Care Excellence, 2016).

A rehabilitation programme should be delivered on a specialist rehabilitation ward (National Institute for Health and Care Excellence, 2016). In England, it is currently recommended that those who are able to participate should have a minimum of 45 minutes per day of each relevant therapy on at least five days a week (National Institute for Health and Care Excellence, 2016, The Royal College of Physicians, 2016). For selected patients with mild to moderate disability, early discharge home from hospital with ongoing rehabilitation provided by a dedicated Early Supported Discharge team is associated with significantly reduced lengths of stay, a lower risk of requiring institutional care, and a greater odds ratio of regaining independence in activities of daily living (Early Supported Discharge Trialists, 2009). Discharge to such services is therefore recommended for patients who have a safe

environment to return to, and who are able to transfer safely from bed to chair (National Institute for Health and Care Excellence, 2016, The Royal College of Physicians, 2016). Following discharge long-term rehabilitation should be provided if necessary, to facilitate participation in social roles such as employment, hobbies, and relationships (National Institute for Health and Care Excellence, 2016). Patients' health and social care needs should be assessed at six months and twelve months after stroke, then annually thereafter, to ensure that they are receiving the care they need (National Institute for Health and Care Excellence, 2016)

1.1.3.7. Current status of stroke services in the UK

Although evidence-based guidelines are available to support best practice in stroke care, there has for some time been disquiet that services in the UK compare unfavourably with those of other European nations (Department of Health, 2007), despite the UK having one of the largest total healthcare expenditures for cardiovascular disease of any of the countries studied (Leal et al., 2006). Published in 2007, the National Stroke Strategy (Department of Health, 2007) set out ten key priorities for improving acute and long term care for those with stroke. Central to these recommendations was early access to specialist multi-disciplinary rehabilitation, with intervention beginning in the acute phase on an acute stroke unit then continuing through early supported discharge to the community and maintained thereafter into the long term according to need (Department of Health, 2007). At the time of publication, it was estimated that only around 50% of patients were able to access appropriate rehabilitation within the first six months after discharge (Department of Health, 2007). When progress towards meeting these strategic aims was reviewed in 2010, it was clear that development of rehabilitation services continued to lag behind that of acute care (Acler et al., 2009b). Only 36% of hospitals had early supported discharge teams, with no consensus on how such teams would be funded (Acler et al., 2009b). The National Sentinel Stroke Audit of 2010 found that around a third of patients did not receive therapy of the recommended intensity (Intercollegiate Stroke Working Party, 2011).

1.1.3.8. Stroke services in developing countries

Although stroke services in the UK clearly require improvement, it is important to bear in mind that what is considered to be “gold standard” care in high-income countries is beyond the means of most low- and middle-income nations to deliver. Unfortunately, these nations are the very ones in which the incidence of stroke is rising most rapidly (Feigin et al., 2009, Mendis, 2013). There are significant barriers to implementing intravenous thrombolysis in such countries, including the significant treatment cost (which is usually borne by the patient), delays in accessing medical help (due to lack of awareness of stroke symptoms or prolonged transit times to regional medical centres), and the paucity of scanning facilities (Mendis, 2010, Kamalakannan et al., 2016). There are frequently critical shortages of key health workers, and the costs of setting up organised stroke units may be prohibitive (Mendis, 2010). In such countries public health measures focusing upon the primary prevention of cardiovascular diseases in general may be a more efficient use of limited resources (Garbusinski et al., 2005). For those who sustain a stroke in developing countries, the focus of treatment remains on early initiation of aspirin (World Health Organization, 2009), as well as on the prevention of secondary complications such as infections and pressure ulcers (Mendis, 2010). Multidisciplinary rehabilitation, with its focus upon mitigating disability and maximising independence, may have much to offer stroke survivors in low-income countries. Although data on the availability of rehabilitation services in such countries are sparse, it is probably reasonable to assume that there exists a substantial unmet need (The World Health Organisation, 2011). It is likely, moreover, that many barriers exist to providing rehabilitation services including lack of funding for healthcare, lack of appropriately-trained personnel, and the centralisation of such rehabilitation services as there are in major population centres beyond travelling distance for those in rural communities (The World Health Organisation, 2011). Where rehabilitation services do exist, the cost of accessing them may be prohibitive, and access to appropriate assistive technologies is often limited (Kamalakannan et al., 2016). The implications of unmet rehabilitation needs in developing countries are potentially profound, since a lack of timely and appropriate rehabilitation may contribute to long term limitation of activities and restriction of

participation, and lead to poorer quality of life and dependency upon others for assistance (The World Health Organisation, 2011). Epidemiological data on disability after stroke and on access to rehabilitation in developing countries are urgently needed if appropriate priorities for development of services are to be set (The World Health Organisation, 2011).

1.1.4. What is “rehabilitation”?

1.1.4.1. Definition of “rehabilitation”

The World Health Organisation defines “rehabilitation” as “The use of all means aimed at reducing the impact of disabling and handicapping conditions and at enabling people with disabilities to achieve optimal social integration” (Gutenbrunner et al., 2006). The overall aim of effective rehabilitation is to “enable people with disabilities to lead the life that they would wish, given any inevitable restrictions imposed on their activities by impairments resulting from illness or injury” (Gutenbrunner et al., 2006). The cornerstone of rehabilitation practice is the World Health Organisation International Classification of Functioning, Disability, and Health (ICF; Figure 1.1.) (World Health Organisation, 2001).

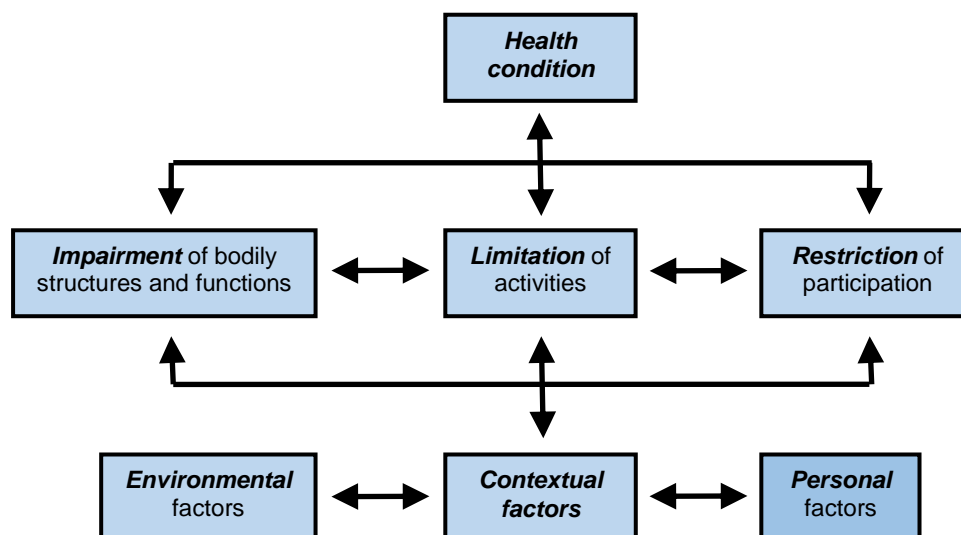


Figure 1.1. The WHO International Classification of Functioning, Disability, and Health (ICF)

1.1.4.2. The WHO International Classification of Functioning, Disability, and Health

The WHO-ICF model was published in 2001, and postulates that an underlying *health condition* may cause *impairment* in bodily structures and function, which lead in turn to *limitation* of activities and *restriction* of participation (The World Health Organisation, 2011). In the context of the ICF, a “health condition” is a generic term used to describe any underlying disease or disorder: including genetic predispositions, or circumstances such as ageing (Gutenbrunner et al., 2006). “Bodily structures” are defined anatomically as organs, limbs, and their components (Gutenbrunner et al., 2006). “Bodily functions” are defined primarily in physiological terms, but also include cognitive, mental and psychological functions (Gutenbrunner et al., 2006). “Activity” is conceptualised as an individual’s ability to execute a tasks such as walking, transferring, or personal care (Gutenbrunner et al., 2006). “Participation” relates to the individual’s ability to fulfil their role in life situations such as employment, leisure pursuits, or relationships (Gutenbrunner et al., 2006). The spectrum of dysfunction after stroke has been defined by expert consensus in a “core set” of ICF codes (Geyh et al., 2004). For example, clinically important *impairments* following stroke include changes in intellectual functions, gait pattern, inability to sequence complex movements, and alterations in muscle tone and power (Geyh et al., 2004). *Activity limitations* and *participation restrictions* may include difficulties in problem solving, transferring oneself, maintaining family and personal relationships, or acquiring and keeping a job (Geyh et al., 2004).

Important though they are, characteristics intrinsic to the individual (physical impairment and limitation of activities) are themselves insufficient to define the construct of “disability”. An individual’s functional abilities are often profoundly influenced by wider *contextual factors* (Gutenbrunner et al., 2006), both *environmental* and *personal*. *Environmental factors* are external to the individual, yet make up the background of their lives (Gutenbrunner et al., 2006). They comprise not only the built environment, but also legislation (such as anti-discrimination laws), and societal attitudes (Gutenbrunner et al., 2006). Depending upon circumstances they may serve as barriers to or as facilitators

of individual function (Gutenbrunner et al., 2006). *Personal factors* are intrinsic to the individual but not related directly to the underlying health condition (Gutenbrunner et al., 2006). They include gender, age, race, and physical fitness (Gutenbrunner et al., 2006).

1.1.4.3. Defining “disability”

In ICF terms, “disability” is therefore constructed as an “*interaction* between an individual (with a health condition) and that individual’s *contextual factors (environmental and personal)*” [author’s italics] (The World Health Organisation, 2011, Dahl, 2002). If the role of rehabilitation professionals is to mitigate disability due to an underlying health condition, then rehabilitation interventions must comprise a package of measures intervening at multiple levels of the ICF. The precise nature of the intervention required, and the ICF level upon which efforts will be focused, will vary depending upon the patient’s goals, and with the passage of time from stroke. Within the first few hours optimum acute care helps to minimise the extent of tissue injury and secondary complications (impairment in structure), and therefore maximise preservation of function (The Royal College of Physicians, 2010). As planning for discharge progresses then environmental assessment and, if necessary, provision of assistive technologies (contextual factors) may help to enhance safety and personal independence on leaving hospital. In the long term, a combination of rehabilitation interventions at impairment (cognition, communication, psychological status), activities, and environmental (environmental modification, assistive technologies) levels may be necessary to address specific goals such as return to work (participation) (National Institute for Health and Care Excellence, 2013). Many stroke survivors require not only intermittent discrete periods of time-limited rehabilitation to address particular functional goals, but also longer term monitoring and support to prevent complications and to mitigate the effects of changing disability (such as deterioration in mobility due to accelerated joint ageing) (The Royal College of Physicians, 2010).

1.1.4.4. Disability on an international scale

Article 26 of the United Nations Convention on the Rights of Persons with Disabilities calls for signatory nations to develop services to support people

with disabilities to attain and maintain maximum independence and full participation in all aspects of life (United Nations, 2009). Despite this, in many low- and middle-income countries people with disabilities continue to experience significant barriers to participating fully in society (The World Health Organisation, 2011). These include the lack of access to rehabilitation services or to appropriate equipment, inaccessibility of the local environment or public amenities including health services, negative attitudes towards people with disabilities, and a lack of involvement of people with disabilities in shaping legislation and policy (The World Health Organisation, 2011). Worldwide, people with disabilities are more likely to be unemployed, and those who are employed generally earn less than non-disabled workers (The World Health Organisation, 2011). Households with a disabled member are more likely to experience poor housing, food insecurity, and a lack of access to sanitation or safe water supplies (The World Health Organisation, 2011). Since disabled people are more likely to live in poverty, they may lack the financial capability to mitigate or overcome barriers to participation in society.

1.1.5. The trajectory of recovery after stroke

1.1.5.1. Early recovery

There is significant heterogeneity in outcomes following stroke: some make a near-complete recovery, whilst others are left with profound residual disability (Kwakkel and Kollen, 2013). Despite this variability, some interesting patterns in recovery may be discerned. The rate of recovery of function tends to be maximal within the first few weeks of stroke. This improvement is thought to be largely spontaneous (Kwakkel and Kollen, 2013).

1.1.5.2. Late recovery

By around eight to ten weeks after stroke the rate of spontaneous recovery begins to plateau (Partridge and Morris, 1993). Thereafter recovery occurs more slowly, but the final level of ability that a patient attains may nevertheless be improved by appropriate rehabilitation (National Institute for Health and Care Excellence, 2013). The biological basis of this process will be discussed below. The sequence of recovery of activities tends to follow a hierarchical pattern: those that may be accomplished by use of compensatory strategies

such as grooming tend to be achieved before more complex tasks such as climbing stairs (Kwakkel and Kollen, 2013).

1.1.5.3. Long-term prognosis

Although considerable effort and money have been invested in early rehabilitation services, there is evidence that gains in function achieved in the first few months after stroke are not sustained in the long term. A recent multi-centre longitudinal cohort study that recruited across several European rehabilitation services found that motor and functional outcomes all improve within the first six months, but this is followed by a progressive decline in function over time (Appelros and Viitanen, 2004). By five years outcome scores do not differ significantly compared with those recorded at two months after stroke (Appelros and Viitanen, 2004). Perhaps it is insufficient to focus research effort upon strategies to achieve short-term gains in function, important though this is. A survey of stroke survivors conducted in 2012 has identified a clear desire to prioritise funding for research in to managing the long-term sequelae of stroke at both impairment (motor, cognitive, and speech functions), and participation (general confidence, and emotional well-being) levels of the ICF (Pollock et al., 2012).

1.1.6. What rehabilitation interventions are effective after stroke?

1.1.6.1. Restorative versus compensatory rehabilitation strategies

Broadly, rehabilitation interventions may be viewed as restorative (for example, encouraging use of a hemiplegic limb in functional tasks or attention to a neglected side), or compensatory (such as learning to complete tasks one handed, or the provision of orthoses, walking aids, or assistive technologies) (National Institute for Health and Care Excellence, 2013).

1.1.6.2. Restorative rehabilitation: motor control approaches

At an impairment level, a variety of physiotherapy approaches have been developed, each based upon differing theories about how patients recover from stroke. Prior to the 1940s, the emphasis of stroke rehabilitation was upon compensating for lost function and maintaining joint range by passive movement (Langhorne et al., 2009). In the 1950s and 1960s, new approaches

were developed that incorporated emerging insights in to the neurophysiology of motor control (Langhorne et al., 2009). Berta and Karel Bobath based their treatment on a detailed assessment of abnormalities in a patient's postural control reflexes, and sought to correct these by intervening upon abnormalities in muscle tone (Semans, 1967).

1.1.6.3. Restorative rehabilitation: feedback on movement performance

The process of rehabilitation depends upon a patient's ability to re-learn motor skills. Learning may be "explicit", in which the components of a task are committed to declarative memory, or "implicit", in which a sequence of movements is learned and optimised without conscious recognition or recall of its component tasks (Subramanian et al., 2010). The biological basis of motor learning is a process termed "neuroplasticity". This will be discussed in detail below, but in the context of motor learning it may be defined as the reorganisation of movement representations in cortical regions (such as the pre-motor cortex and supplementary motor area) resulting from physiological changes to synaptic efficacy and remodelling of dendrite spines (Subramanian et al., 2010).

When re-learning movements after stroke, it is common for patients to develop inefficient or ineffective movement patterns that, if allowed to persist, may compromise their long-term independence (Subramanian et al., 2010). For this reason, therapists commonly offer patients feedback on their performance during practice of a task (van Vliet and Wulf, 2006, Subramanian et al., 2010). Feedback may be *intrinsic* (the sensory and proprioceptive information available to an individual whilst a movement is in progress), or *extrinsic* (an external commentary on the quality of movement performance) (van Vliet and Wulf, 2006, Subramanian et al., 2010). Extrinsic feedback may be given verbally, or augmented with visual information such as video recording of task performance (van Vliet and Wulf, 2006). It may encompass the way in which a movement is executed ("knowledge of performance"), or the outcome of that movement ("knowledge of results") (van Vliet and Wulf, 2006, Subramanian et al., 2010). Since intrinsic feedback mechanisms may be impaired following a stroke, extrinsic feedback may be particularly valuable in improving the quality of movement in stroke survivors (van Vliet and Wulf, 2006). Providing

extrinsic feedback may also allow stroke patients to utilise declarative memory (for facts and events) when learning skills, bypassing impairment in implicit learning mechanisms (the learning of motor skills without conscious awareness; van Vliet and Wulf (2006)).

1.1.6.4. The evidence base for different motor rehabilitation strategies

A recent Cochrane review comparing competing physiotherapy approaches to promoting recovery of lower limb function and postural control following stroke found that there is no clear evidence to favour any one strategy (Pollock et al., 2007). There is, however, some evidence to support specific interventions. Langhorne et al. (2009) found modest effect sizes for constraint-induced movement therapy, electromyographic biofeedback, mental practice with motor imagery, and robotic interventions in promoting recovery of arm function. However, none of these strategies was of proven benefit in enhancing recovery of hand function (Langhorne et al., 2009). Interventions to promote recovery of walking included fitness training (both cardiorespiratory alone, and combined cardiorespiratory and strength training); high-intensity physiotherapy; and repetitive task practice (Langhorne et al., 2009). All showed modest effect sizes, but only in the case of cardiorespiratory training was the evidence of benefit felt to be strong (Langhorne et al., 2009). For standing balance, interventions trialled include biofeedback using a force plate, training on a moving platform, and repetitive task training. Trials were generally small and of poor quality, and evidence to support any one intervention was considered weak (Langhorne et al., 2009). For sit-to-stand transfers, only repetitive task training showed a modest effect across seven trials. However, this review included a total of only 346 patients (Langhorne et al., 2009). A later Cochrane review, of upper limb rehabilitation strategies, identified larger numbers of trials; but, once again, of moderate quality at best (Pollock et al., 2014). Beneficial effects were shown for repetitive task practice (at high doses), mental practice, mirror therapy, constraint-induced movement therapy, and interventions for sensory impairment (Pollock et al., 2014).

1.1.7. The biological basis of rehabilitation

1.1.7.1. Motor learning: the basis of rehabilitation interventions

Although evidence is lacking for particular interventions, participation in a structured multi-disciplinary rehabilitation programme has been shown to improve function and quality of life after stroke, and this complex package of interventions therefore remains the cornerstone of treatment for the majority of patients (Department of Health, 2007, National Institute for Health and Care Excellence, 2013). But what biological processes underpin “rehabilitation”? One might say that strategies that aim to restore (as opposed to compensate for) loss of function depend fundamentally upon a person’s ability to *learn* skills.

1.1.7.2. Motor learning after stroke

The process of learning is itself a complex phenomenon. In healthy individuals, learning a new skill depends upon a close interaction between spatially distributed and functionally disparate areas of the brain (Doyon et al., 2009, Hikosaka et al., 2002, Penhune and Steele, 2012). Where brain injury has occurred, as in stroke, there is evidence to suggest that areas remote to the original injury are activated during task learning: a process that is thought to result in partial reorganisation of brain function and the re-location of motor representations to spared areas of brain (Hodics et al., 2006). At a cellular level, learning is thought to induce the formation of new “hard wired” pathways within the brain through re-modelling of axons, changes in the number and morphology of dendrites (Dimyan and Cohen, 2011, Ward and Cohen, 2004, Gillick and Zirpel, 2012), and long term potentiation or depression of synaptic transmission (Gillick and Zirpel, 2012, Ward and Cohen, 2004). These processes are enhanced by repetitive practice of a task (French et al., 2007).

1.1.7.3. Novel rehabilitation interventions to enhance learning

It is thought that neuroplasticity is enhanced by repetitive practice of a task. Thus, more intensive practice, delivered early after stroke, has the potential to improve rehabilitation outcomes. For this reason, there has been enormous research interest in the development of novel rehabilitation strategies to enhance neuroplasticity. For example, the use of robotic systems to assist a

patient in active movement of a paretic arm might allow a greater duration and intensity of practice than could be provided by a therapist, and may therefore enhance recovery of upper limb function (Pollock et al., 2014, Kwakkel et al., 2008, Sivan et al., 2014). Alternatively, the use of drugs that act to enhance neuroplasticity directly might offer a means of amplifying the effectiveness of traditional rehabilitation interventions (Chollet et al., 2011, Berends et al., 2009, Scheidtmann et al., 2001). An understanding of the processes of learning in stroke might allow the targeting of emergent rehabilitation interventions to those patients most likely to benefit. The biology of learning will thus be considered in Part 1.2, below.

Part 1.2 Motor learning and recovery from stroke

1.2.1. Motor learning after stroke

1.2.1.1. Defining “motor learning”

Motor learning is fundamental not only to rehabilitation but also to daily life (Dayan and Cohen, 2011): from a baby learning to walk to a musician practising a symphony. Becoming skilled in a motor task requires us not only to learn the correct order of movements, but also to develop an awareness of the sensory input that guides decisions such as timing of the movement, trajectory, and what force should be applied when manipulating an object (Penhune and Steele, 2012). Often, acquiring a skill also requires the learner to manipulate or interact with objects in their environment (Doyon et al., 2009). Although there is no universally-accepted definition of “motor learning”, it has been conceptualised as “a change in motor behaviour, specifically referring to the increased use of novel, task-specific joint sequences and combinations, resulting from practice and/or repetition” (Nudo, 2008). It is now recognised that motor learning, and other cognitive functions, are critically dependent upon network interactions between spatially-distributed brain structures. It is interesting to consider how a historical understanding of brain function, revealed by an understanding of the consequences of direct injury to discrete cortical areas, has evolved in to a more modern appreciation of the role of neural networks and the connectional anatomy of the brain.

1.2.1.2. Early theories of motor control

Many of the early endeavours to understand brain function focused upon motor control, rather than cognition. This work has been the subject of a historical review by Gross (2007), which will be summarised here in the next two paragraphs. As long ago as 1664, Thomas Willis suggested that the cortex initiates voluntary movement. Emanuel Schwenbourg (1688-1772) proposed that motor function is localised in the cortex in a somatotopic manner, with cortical neurones projecting down through the white matter to the medulla, and thence to the spinal cord and peripheries. His insight, though astonishing to modern eyes, was well ahead of its time and was largely ignored by the scientific community of the day. Françoise Pourfour du Petit

(1644-1741), a French military surgeon, demonstrated the laterality of motor function in a series of experiments with animals which he correlated with observations in wounded soldiers. Again, these findings were largely ignored, and the prevailing view of the cortex in to the early 18th Century was conveyed by the literal translation of the term from Latin: little more than a protective “rind”.

1.2.1.3. The discovery of the cortical localisation of motor control

Amongst the earliest circumstantial evidence for the localisation of cortical function came in 1870, with the observation by John Hughlings Jackson that his wife’s seizures showed a distinct pattern of progression (Gross, 2007). He recorded twitching that began first in the hand then moved in a stereotyped manner up the arm before the seizure become generalised. From this, he inferred that distinct muscle groups must be controlled by co-located brain areas. He did not, however, directly implicate the cortex as the seat of motor function. The first direct experimental evidence for the existence of a “motor cortex” came at around this time (1870) when Gustav Fritsch and Edvard Hitzig observed reproducible patterns of limb twitching in response to “Galvanic” (electrical) stimulation of the anterior cortex in dogs. Fritsch and Hitzig did not themselves cite Jackson’s work, although Jackson’s findings were certainly known to David Ferrier who successfully reproduced Fritsch and Hitzig’s experiment in 1873. This discovery heralded a growth in interest throughout the 19th Century in determining the localisation of brain functions. Some of this work, such as Carl Wernicke’s seminal 1874 case series of ten patients with the aphasia which now bears his name (Wernicke, 1970), has stood the test of time (de Almeida et al., 2014). Other theories have fallen in to disrepute. Franz Gall (1758-1828) proposed not only that skills and personality traits have their seat in the cortex, but also that the presence of these traits in specific individuals would lead to cortical hypertrophy (de Almeida et al., 2014). This would in turn result in displacement of the overlying skull, and a characteristic pattern of skull prominences from which the presence of defined personality characteristics could be inferred (de Almeida et al., 2014). Although both flawed in its methodology and erroneous in its conclusions, it is worth noting that this theory of “phrenology” was amongst

the first attempts to systematically localise cortical functions (de Almeida et al., 2014).

1.2.1.4. Wernicke, and the discovery of network interactions between brain structures

If stroke were a purely “cortical” phenomenon, then predicting recovery would be straightforward: the spectrum of impairments, and their ultimate outcome, would depend upon the location and extent of the cortical lesion. However, it has long been known that the spectrum of impairment seen following a brain injury of any nature depends not only upon the pattern of cortical injury, but also upon disruption of connections between different brain structures. Wernicke described in 1874 how the production of speech depends upon the integrity of connections between the superior temporal gyrus and Broca’s area in the posterior inferior frontal gyrus (Wernicke, 1970). This was followed in 1885 by Ludwig Lichtheim’s description of what he termed a “reflex arc” between cortical areas responsible for understanding spoken language and those responsible for initiating the motor component of speech (de Almeida et al., 2014).

1.2.1.5. White matter tracts and loop circuits: a contemporary view of brain function

More recently, the existence of extensive networks of white matter projections between spatially-distributed structures (both cortical and sub-cortical) has been recognised. The basal ganglia are key nodes within these circuits. They comprise the striatum (caudate, putamen, and nucleus accumbens), and the globus pallidus (Da Cunha et al., 2009, Bolam et al., 2000). The sub-thalamic nucleus, substantia nigra and ventral tegmental areas are considered associated structures (Da Cunha et al., 2009). Alexander et al. (1986) described five loop circuits between the cortex and basal ganglia: the motor, oculomotor, dorsolateral prefrontal, lateral orbitofrontal, and anterior cingulate. Each arises from different regions of the frontal cortex (Alexander et al., 1986), and sends excitatory inputs to the striatum (McHaffie et al., 2005). Striatal neurones then send a complex web of inhibitory inputs to the substantia nigra and the globus pallidus interna, which project in turn to the thalamus (McHaffie et al., 2005). The primary output of these circuits is

excitatory efferents from the thalamus to cortical areas (McHaffie et al., 2005). (Figure 1.2. adapted from McHaffie et al. (2005)).

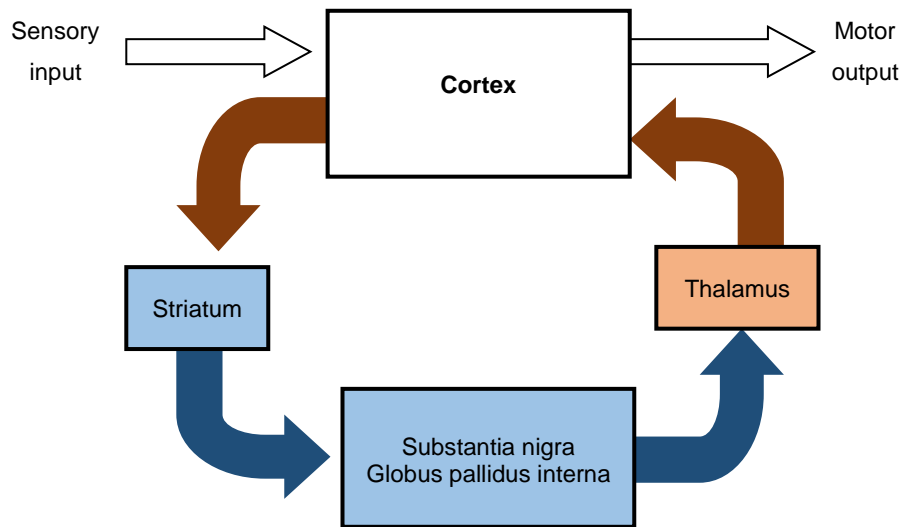


Figure 1.2. Schematic illustration of the architecture of the cortico-basal loop circuits

Predominantly excitatory pathways and structures are in red; those with predominantly inhibitory output are in blue. Figure taken from (McHaffie et al., 2005)

Similar loop circuits are also now known to exist between subcortical structures (McHaffie et al., 2005). In this case, the primary input nucleus is the thalamus, which sends excitatory input to the striatum (McHaffie et al., 2005). This in turn sends inhibitory projections to the substantia nigra and globus pallidus interna, which in turn send inhibitory input back to midbrain and hindbrain structures (McHaffie et al., 2005) (Figure 1.3. (McHaffie et al., 2005)).

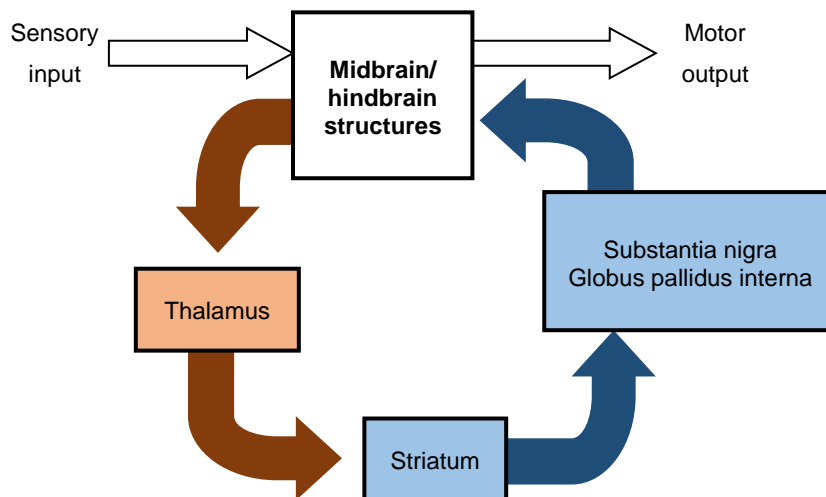


Figure 1.3. Schematic illustration of the architecture of the subcortical loop circuits.

Predominantly excitatory pathways and structures are in red; those with predominantly inhibitory output are in blue. Figure taken from (McHaffie et al., 2005)

1.2.1.6. Neuronal networks and cognitive functioning

Cortico-basal and subcortical circuits are now known to play a role in a variety of cognitive processes. The clinical evidence for this derives in part from conditions other than stroke. Huntington's chorea and Parkinson's disease are both degenerative conditions of the basal ganglia, which have impairment of motor control as their primary manifestation. And yet Huntington (1872) also described a "tendency towards insanity" in advanced cases, including sexual disinhibition. The non-motor manifestations of Parkinson's disease were not at first appreciated: Parkinson (1817) himself noted that the "senses and intellects... [are] uninjured". It was only later that cognitive dysfunction was also recognised in the advanced stages of the illness (Louis, 1997). Impairment in concentration and attention, strategic planning, procedural learning ability, working memory, and verbal fluency are all now recognised features of this condition, as are decreased mental flexibility and difficulty in switching between cognitive tasks (Schmahmann and Pandya, 2008). More recently, studies of discrete stroke lesions in humans have demonstrated a similar pattern of cognitive impairment following injury to the basal ganglia

(Schmahmann and Pandya, 2008) and cerebellum (Schmahmann et al., 2009).

1.2.1.7. Neuronal networks and motor learning

There has recently been considerable interest in how network interactions between disparate brain structures might interact to facilitate the learning of motor skills. Advanced imaging techniques may offer insights into the anatomical basis of learning. At the simplest level, techniques such as Voxel-based morphometry (VBM) or Diffusion Tensor Imaging (DTI) allow detailed analysis of the volume of grey matter structures, or visualisation of the white matter tracts that link them (Thomas and Baker, 2013). Statistical comparison of anatomical differences between trained and untrained individuals, or within the same group before and after learning a task, may allow inferences to be made about the role of these structures in the learning process (Thomas and Baker, 2013). However, VBM and DTI merely provide semi-quantitative estimates of structural change: they do not allow real-time visualisation of the activation of these brain regions as learning takes place.

1.2.1.8. Theories of motor learning: evidence from functional magnetic resonance imaging (fMRI)

In contrast to structural imaging, functional imaging techniques allow exploration of how patterns of metabolic activity within specific brain regions change throughout the learning cycle. Functional MRI (fMRI) relies upon the detection of increased levels of deoxygenated haemoglobin in brain regions of interest: the Blood Oxygen Level-Dependent (BOLD) signal (Arthurs and Boniface, 2002). This is assumed to reflect increased oxygen uptake by metabolically active tissue, and therefore increased neuronal activity in that area (Arthurs and Boniface, 2002). Several studies have used fMRI to explore the process of motor learning. Doyon et al. (2009) suggest that the striatum contributes to consolidation of skills, with activity first predominant in the associative striatum, but a subsequent shift to the sensorimotor striatum. Hikosaka et al. (2002) hypothesised that successful movement requires an initial awareness of the body's spatial position and of the position of environmental objects with which it interacts. This requires integration of any available spatial information, which is thought to be performed by circuits

between the fronto-parietal cortices and the associative striata (Hikosaka et al., 2002). This information is then used to generate a series of motor coordinates for the planned action prior to execution of a movement: a function thought to be performed by loops between the motor cortex, basal ganglia and cerebellum (Hikosaka et al., 2002). When learning a new sequence of movements, initially each component of that sequence is executed individually (Hikosaka et al., 2002). This is an explicit process which requires cognitive effort, and results in slow and deliberate movements (Hikosaka et al., 2002). The sequence of actions are subsequently optimised, in an implicit process requiring no conscious thought (Hikosaka et al., 2002). The end result is fluid effortless movement, that retains spatial accuracy (Hikosaka et al., 2002). Penhune and Steele (2012) believe that the cerebellum is responsible for the construction of an “internal model,” containing the optimum kinematic parameters for a planned movement sequence. This representation is then compared with proprioceptive feedback whilst the movement is in progress, allowing optimisation of movement in real time (Penhune and Steele, 2012). The final anatomical localisation of memory traces for learned action is split, with the motor, pre-motor, and parietal cortex encoding a representation of a learned sequence of movements, and the cerebellum encoding the motor control parameters for that action (Penhune and Steele, 2012). The role of the striatum in the learning process is in the “reward” response when an explicit goal is achieved (Penhune and Steele, 2012). Despite the term “functional” MRI, what this technique actually demonstrates is a signal that is thought to correlate with tissue metabolism: any inferences about the actual *function* of those structures in learning remain speculative.

1.2.2. Cognitive dysfunction after stroke

1.2.2.1. Cognitive impairment after stroke: a “disconnection” phenomenon

As understanding of cognitive function has evolved, it has become apparent that injury to structures such as the basal ganglia, cerebellum, or white matter tracts may give rise to a picture of cognitive dysfunction that mimics a large cortical injury. Such phenomena have been termed “*disconnection syndromes*”, since they represent a failure of the network between brain

structures (Schmahmann and Pandya, 2008). Impairments in key cognitive domains such as memory, executive function, praxis, and visuospatial perception are common after stroke (Barker-Collo et al., 2010). They may occur as a result of a variety of underlying processes, with the common feature being disruption of distributed neural networks and thus failure of interactions between brain structures. Unfortunately, the wide array of pathological lesions that may lead to cognitive failure has led to a bewildering array of terminologies to describe these phenomena (O'Brien et al., 2003). Some imply the presence of specific histological findings: “multi-infarct dementia”, for example, presumes an additive burden of several cortical infarcts, whereas “subcortical dementia” and “subcortical ischaemic vascular dementia” suggest a burden of lacunar infarcts to the basal ganglia. The term “dementia”, common to all of the above, is based largely upon the characteristics of Alzheimer’s disease, and therefore presupposes the presence of *memory* impairment as a key diagnostic feature (O'Brien et al., 2003, Moorhouse and Rockwood, 2008). Other terms, such as “vascular cognitive impairment” seek to define a construct, whilst minimising assumptions about aetiology and pathophysiology (O'Brien et al., 2003).

1.2.2.2. Classifying cognitive impairment after stroke

Perhaps the most straightforward taxonomy is that proposed by O'Brien et al. (2003) and later elaborated by Moorhouse and Rockwood (2008) (Figure 1.4.). The use of “Vascular Cognitive Impairment” (VCI) was initially proposed as an umbrella term for cerebrovascular pathology which results in a specific cognitive profile: preserved memory, with impairment in attentional and executive functioning (O'Brien et al., 2003). It has subsequently been suggested that vascular cognitive impairment which results in memory impairment (thereby meeting diagnostic criteria for “dementia”) be termed “VCI with dementia”. There is, of course, a substantial overlap between VCI and neurodegenerative pathology.

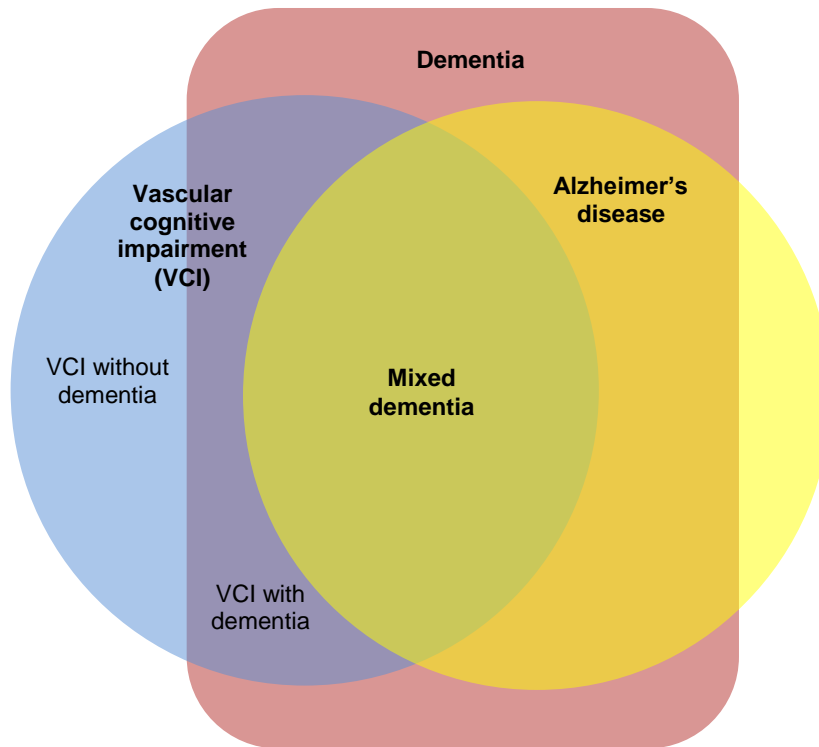


Figure 1.4. schematic representation of the overlap between syndromes of cognitive impairment.

After Moorhouse and Rockwood (2008)

1.2.2.3. Cognitive impairment as a result of a global burden of injury to neuronal networks

If cognitive dysfunction is conceptualised as being a result of disruption to the network anatomy of the brain, then it is clear that this impairment may arise as a result of a variety of pathologies, and as a consequence of disruption to any of the structures or white matter tracts within the network. Some conditions, such as a stroke affecting a large cortical territory or the cerebellum, or a smaller “strategic” lesion to an area critical to cognitive function, may cause a sudden and dramatic deterioration, which may fulfil the criteria for dementia (Iadecola, 2013). However, a more generalised burden of chronic ischaemic injury to the cortico-basal and subcortical loop pathways may result in a subtle and insidious cognitive deterioration, which may even pre-date or occur in the absence of a large-vessel stroke (Iadecola (2013); Figure 1.5).

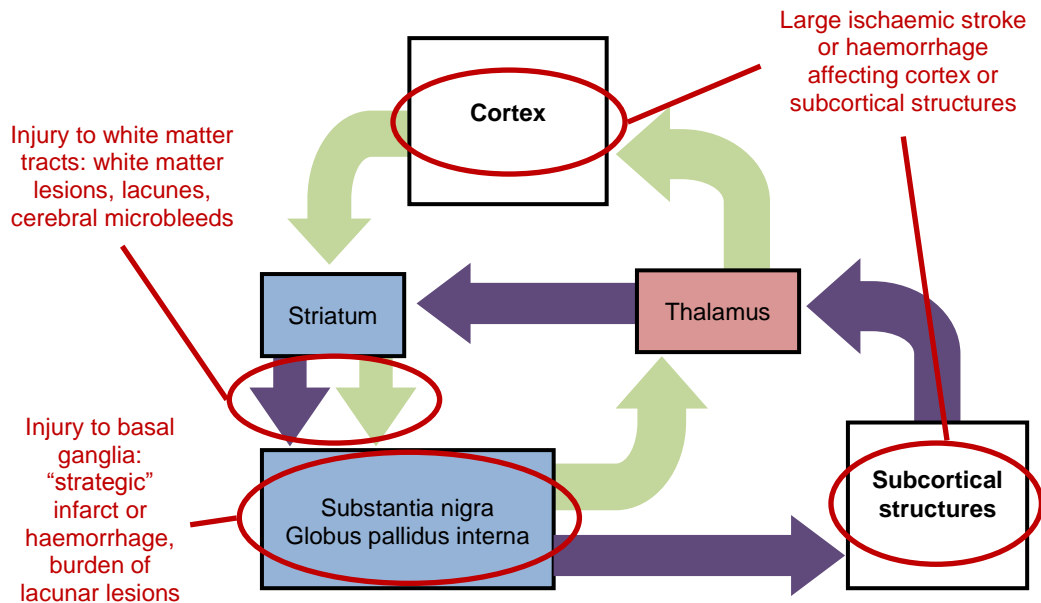


Figure 1.5. Mechanisms of injury to cortico-basal and subcortical loop circuits.

Schematic illustration of the cortico-basal (green) and subcortical (purple) loop circuits. Different components of these pathways may be susceptible to injury by a variety of mechanisms (red). This may manifest as vascular cognitive impairment.

In short, the overall picture of cognitive dysfunction after stroke most likely represents an interaction between a large-vessel lesion (infarct versus ICH) and a more global burden of “small vessel” injury (some of which may be pre-existing). How this overall burden of structural (brain injury) and functional (cognitive) impairment might attenuate a patient’s ability to re-learn motor skills is of particular interest to rehabilitation practice.

1.2.2.4. “Small vessel” injury: an underlying cause of cognitive impairment

“Small vessel” injury is a concept that covers a variety of lesions seen on brain imaging, which may or may not have similar underlying pathological mechanisms. Here again, one encounters the problem of a lack of standardised terminology and definitions for these lesions (Wardlaw et al., 2013b). Often several different terms are used to describe the same phenomenon. Some (such as “white matter hyperintensities”) describe radiological findings (the appearance of these lesions on T2-weighted MRI); others (such as “leukoencephalopathy”) refer to histopathological changes

(white matter necrosis of presumed ischaemic aetiology) (Wardlaw et al., 2013b). From an imaging perspective, a single pathological process may mature to give very different radiological appearances on follow-up scans. For example, a small acute subcortical infarct may leave no visible lesion on a follow-up MRI scan, or it may appear as a cavitating lacunar lesion or white matter hyperintensity (Wardlaw et al., 2013b). In pathological terms, there are a wide range of possible mechanisms by which brain injury may occur; and yet, the repertoire of possible tissue responses to injury (inflammation, necrosis, scarring) are limited (Hachinski, 2007). It therefore cannot be assumed that lesions with similar histological appearances share a common mechanism. With these difficulties in mind, three common lesions implicated in vascular cognitive impairment will be discussed: white matter lesions, lacunar lesions, and microbleeds. For each an attempt will first be made attempt to instil some clarity around definitions, before the underlying pathophysiology of these lesions and their consequences for cognitive function are explored.

1.2.2.5. Imaging correlates of small vessel disease: white matter lesions

There are over 50 synonyms in use to describe white matter lesions: Binswanger's disease, leukoariosis, leukoencephalopathy, white matter hyperintensity, white matter change, and white matter disease are amongst the most common (Wardlaw et al., 2013b). They appear on T2-weighted MRI as areas of hyperintensity in the deep or periventricular white matter, which may be patchy or confluent (Wardlaw et al., 2013b, Wardlaw et al., 2013a). On computerised tomography scanning (CT), they are hypodense, returning an attenuation lower than that of surrounding tissue (although not as low as cerebrospinal fluid (Wardlaw et al., 2013b). They are also seen in other conditions such as multiple sclerosis or leukodystrophies (Wardlaw et al., 2013b). Wardlaw et al. (2013b) therefore proposed the radiological descriptors "white matter hyperintensities of presumed vascular origin" for the MRI appearance, with "white matter hypodensities of presumed vascular origin" endorsed for the equivalent CT finding. Since both magnetic resonance imaging (MRI) and CT findings will be discussed here, the more generic (but

less precise) term “white matter lesions” will be used, with reference to white matter “hyperintensity” or “hypodensity” only in the context of MRI and CT findings respectively. It will be assumed throughout that these lesions are “of presumed vascular origin”, unless otherwise stated. Following the recommendations of Wardlaw et al. (2013b), these terms will not be applied to lesions in the brain stem or deep grey matter. Radiologically, white matter lesions are known to be associated with a number of other findings including lacunes, atrophy, cerebral microbleeds, and prominent perivascular spaces (Wardlaw et al., 2013a). They are strongly associated with cardiovascular risk factors including hypertension, hyperlipidaemia, diabetes, and smoking (Wardlaw et al., 2013b, Iadecola, 2013). Histologically, a number of small vessel changes have been associated with these lesions including atherosclerosis, hyaline deposition in the vessel walls (lipohyalinosis), fibrosis and stiffening of small vessels (arteriosclerosis), and loss of integrity of the vascular basement membrane (fibrinoid necrosis) (Iadecola, 2013). How, or whether, these microvascular changes may give rise to white matter lesions remains opaque, but possible mechanisms include chronic hypoperfusion, and/ or dysfunction of the blood/ brain barrier with extravasation of fluid into white matter tracts (DeBette and Markus, 2010). Histological evidence of white matter injury includes axonal loss, vacuolation, and demyelination (Iadecola, 2013). As they progress lesions tend to expand into adjacent normal white matter, and may eventually become confluent (Iadecola, 2013). White matter lesions are common, with a prevalence of 11%-24% in over-65s, and 94% at age 82 (DeBette and Markus, 2010). They may be asymptomatic, and were once thought to be a benign associate of normal ageing. However, it is now clear that they are associated with an increased risk of stroke, dementia, and death (DeBette and Markus, 2010), a faster rate of decline in global cognitive performance, executive function, and information processing speed (DeBette and Markus, 2010), gait disturbance (de Laat et al., 2011), and an increased risk of transition from independence to disability (Inzitari et al., 2009).

1.2.2.6. Imaging correlates of small vessel disease: lacunar lesions

“Lacunes” were first described in 1838 as cavitating lesions containing cerebrospinal fluid of around 3-20mm in diameter (Potter et al., 2010). They

are more common with age; one large MRI survey of participants aged over 65 found one or more lacunes in around 25% of the sample (Longstreth et al., 1998). On imaging, established lacunar lesions are isointense to cerebrospinal fluid (Roman et al., 2002). They are typically found in the deep white matter, basal ganglia, thalamus, and pons (Wardlaw, 2005). They are often assumed to be ischaemic in origin, although a small deep intracerebral haemorrhage can, when mature, give a radiological appearance that is indistinguishable from an ischaemic lacune (Wardlaw et al., 2013b). Although commonly used as such in the literature, the terms “lacune”, “lacunar stroke”, and “lacunar infarction” are not interchangeable. “Lacune” refers to a radiological or pathological finding of a cavitating lesion. Only a minority of “lacunar” small-vessel infarcts actually go on to cavitate and assume this appearance; the majority take on the appearance of white matter lesions (Potter et al., 2010). Simply counting the numbers of lacunes may therefore underestimate the true burden of ischaemic small vessel disease (Potter et al., 2010). “Lacunar stroke” describes a clinical stroke syndrome consistent with a small subcortical or brainstem lesion (Wardlaw, 2008, Bamford et al., 1991). However, this clinical syndrome may not match radiological findings: in around 10-20% of patients with a clinically-defined “lacunar” syndrome a small cortical infarct is later identified on imaging as the culprit lesion (Mead et al., 1999). Nor do all lacunes give rise to a “lacunar stroke” syndrome. As many as 89% are thought to be clinically silent, or are manifested by more subtle impairments in gait and cognition (Longstreth et al., 1998). “Lacunar infarct” implies a lacunar stroke syndrome for which an underlying ischaemic lacunar lesion is visible on imaging (Wardlaw, 2008). The radiological appearance of “lacunes” may be mimicked by expansion of the perivascular spaces around small perforating vessels (Wardlaw et al., 2013b). These are generally smaller than lacunes (around 3mm), run parallel to the course of vessels, and may be seen most prominently in the basal ganglia (Wardlaw et al., 2013b). Frustratingly, this phenomenon has also spawned its own rash of synonyms including “Virchow-Robin spaces”, “état criblé” (for lesions located predominantly in the basal ganglia) or (confusingly) “Type 3 lacune” (Wardlaw et al., 2013b). Although both give the radiological appearance of fluid-filled cavities, the origins and significance of perivascular spaces cannot be

assumed to be the same as that of lacunes of presumed ischaemic or haemorrhagic origin. It is therefore necessary to distinguish carefully between the two. Wardlaw et al. (2013b) suggest the term “lacune of presumed vascular origin”, since this a) differentiates between vascular and non-vascular causes of cavitation, and b) avoids making assumptions about whether the lesion is a consequence of ischaemia or haemorrhage where initial imaging is not available. For simplicity, the radiological term “lacune” will be used here, leaving implicit that this refers only to lesions “of presumed vascular origin” (ischaemic and haemorrhagic) unless otherwise stated. The clinical syndrome of “lacunar stroke” will be defined according to the Oxford Community Stroke Project classification (OCSP) of Bamford et al. (1991), whilst remaining mindful that this syndrome does not always correlate with imaging findings (Mead et al., 1999).

Although the earliest descriptions of lacunes was of ischaemic necrosis on histology (Fisher, 1965), the presumption that small vessel occlusion is the underlying cause (Fisher, 1968) has been challenged. Common precipitants of ischaemic cortical stroke (carotid stenosis or cardiac emboli) are implicated in only around 10-15% of ischaemic lacunar strokes, and some studies purporting to demonstrate a link between lacunes and risk factors for embolisation actually included mild carotid stenosis (as little as 25%), or cardiac abnormalities not typically associated with emboli (such as left ventricular hypertrophy) (Wardlaw, 2005). In animal models, the majority of particles injected in to the carotid artery embolised to the cortical vasculature rather than the lenticulostriate arteries, suggesting that cardiac embolization is not the primary cause of lacunes in the majority of cases (Wardlaw, 2005). Nor may it be reasonable to assume that all lacunes share a common origin. There have been suggestions that larger lacunar infarcts are a consequence of atheromatous disease in more proximal arterioles, whereas lacunar lesions caused by lipohyalinosis and arteriolosclerosis of the microvasculature tend to coexist with white matter hyperintensities (Wardlaw et al., 2013a). The Leukoariosis and Disability (LADIS) study found that lacunes in the basal ganglia were associated with AF (suggesting an embolic cause), whereas those in the deep white matter were often accompanied by new or expanding

white matter lesions and were associated with a history of hypertension and stroke (Gouw et al., 2008).

Reliable estimates for the incidence and prevalence of cognitive impairment after lacunar stroke are hard to come by. A recent systematic review (Makin et al., 2013) found that studies were generally small, with non-blinded assessment of cognitive function, and did not estimate the prevalence of cognitive dysfunction before the stroke (Makin et al., 2013). None used gold-standard imaging techniques to confirm lacunar infarction. Long-term data are scant, with few studies including follow-up beyond one year (Makin et al., 2013). Within these limitations, the prevalence of cognitive impairment and dementia after lacunar stroke was estimated at 29%: comparable with cortical stroke (24%) (Makin et al., 2013).

The figure quoted in this meta-analysis were heavily influenced by one large study, which accounted for 38% of all patients included (Bejot et al., 2011). In this study, the odds ratio for cognitive impairment with lacunar versus non-lacunar stroke was 3.48 (Bejot et al., 2011); far higher than for pooled estimates derived from all other studies analysed by Makin et al. (2013) (odds ratio 0.67 for cognitive impairment with lacunar versus non-lacunar stroke). One possible reason for the disparity is that Bejot et al. (2011) assessed cognitive function at one month post stroke. Their estimates may not reflect the true prevalence of cognitive impairment in the long term. Secondly, the odds ratio for cognitive impairment after lacunar stroke changed significantly in the 24-year period in which the study was recruiting: from 10.1 in 1991-1996, to 1.51 in 2003-2008 (Bejot et al., 2011). The reasons for this striking observation remain unclear: it is possible that changes in clinical practice over the course of the study period led to an improvement in dementia-free survival from stroke (Bejot et al., 2011). However, the possibility of a change in methodology over the course of that study cannot be discounted (Makin et al., 2013).

Although the limitations of the literature in this area must be acknowledged, it is nevertheless clear that cognitive impairment is common after lacunar stroke: perhaps surprisingly so, given the small size of the lesions concerned (Makin et al., 2013). This implies that the degree of cognitive dysfunction

manifested clinically is not dependent upon the size of the infarct, but rather its impact upon wider network functions, and perhaps an interaction with other markers of small-vessel disease, such as white matter lesions (Makin et al., 2013).

1.2.2.7. Imaging correlates of small vessel disease: microbleeds

What are termed “cerebral microbleeds” are thought to represent small perivascular collections of haemosiderin-laden macrophages (Fazekas et al., 1999) which form as a result of leakage of blood products from small vessels injured by hypertension (lipohyalinosis) or by amyloid deposition (amyloid angiopathy) (Werring et al., 2010). There are many synonyms (including “microhaemorrhage”), but “cerebral microbleed” is the most commonly used and has therefore been proposed as a consensus term (Wardlaw et al., 2013b). The descriptor is primarily radiological (Greenberg et al., 2009): on MRI sequences that are sensitive to magnetic effects (gradient-echo T2*), cerebral microbleeds are visible as small (5-10mm diameter), well-demarcated, hypointense lesions (Werring et al., 2010). The pattern of lesions seen may reflect the underlying pathology: hypertensive vasculopathy generally causes microbleeds in the basal ganglia, thalamus, brainstem, and cerebellum, whereas amyloid angiopathy typically displays a lobar distribution (Greenberg et al., 2009). Cerebral microbleeds are associated with hypertension, and may co-exist with white matter lesions and lacunes (Greenberg et al., 2009). They may also be associated with an increased risk of subsequent large-vessel ICH in patients following a first haemorrhage or infarct, although the evidence for this is based upon small samples (Greenberg et al., 2009). Their significance as a marker of future haemorrhage risk in those who have not already had an overt large-vessel stroke is unclear, and the balance of risks versus benefits in initiating antiplatelet therapy in patients with both cerebral microbleeds and risk factors for ischaemic stroke remains unknown (Greenberg et al., 2009). Several small studies have demonstrated an association between cerebral microbleeds and an increased risk of cognitive impairment, dependency, or death; but this may simply reflect the coexistence of these lesions with white matter lesions and lacunes (Greenberg et al., 2009).

1.2.2.8. “Small vessel disease”: a unifying theory?

White matter lesions, lacunes, and cerebral microbleeds often coexist, and it is by no means clear that atherosclerotic processes analogous to those implicated in large-vessel stroke play a role in these processes (Wardlaw et al., 2013a). Associations between these lesions and “traditional” cardiovascular risk factors (such as hypertension, hyperlipidaemia, smoking, and diabetes) have not been firmly established (Vermeer et al., 2007). Indeed, antihypertensive treatment and lipid-lowering agents are ineffective in preventing the expansion of white matter lesions, and antiplatelet therapy is associated with an increased risk of symptomatic ICH and death after lacunar stroke (Wardlaw et al., 2013a). The hypothesis of endothelial dysfunction has recently been proposed as a common origin for these lesions (Wardlaw et al., 2013a). This theory postulates that disruption of the vascular endothelium leads to localised leakage of tissue fluid into the perivascular space and transepithelial migration of inflammatory cells, leading to localised tissue oedema and the characteristic microvascular changes seen in small vessel disease (fibrinoid necrosis, lipohyalinosis) (Wardlaw et al., 2013a). Over time this process could result in the pattern of demyelination and white matter necrosis seen in white matter lesions (Wardlaw et al., 2013a). The same process may also lead to thickening of arteriolar walls, resulting in luminal narrowing and thrombus formation (Wardlaw et al., 2013a). This could result in tissue ischaemia and “lacunar” infarction (Wardlaw et al., 2013a). How such endothelial dysfunction may arise remains speculative. The permeability of the blood-brain barrier is known to increase with normal aging, but how other stimuli might interact with this process to trigger a pathological cascade has yet to be delineated (Wardlaw et al., 2013a). Amyloid deposition in Alzheimer’s disease is known to enhance blood-brain barrier permeability: but permeability is higher in vascular cognitive impairment with dementia than in Alzheimer’s disease or age-matched healthy controls (Wardlaw et al., 2013a).

1.2.2.9. The role of small vessel disease in cognitive impairment

Clearly further work is needed to understand fully how the lesions that characterise so-called “small-vessel disease” arise, and how they may be prevented. What is clear, however, is that they are far from benign. Although

the role of microhaemorrhages in precipitating cognitive dysfunction is less clear, white matter lesions and lacunes are certainly known to be associated with cognitive decline in people who have not had a stroke. After stroke, they are associated with an increased risk of recurrent stroke, and of transition to dementia and dependency. The underlying mechanism for this is most likely network dysfunction. Pre-existing white matter lesions and lacunes may not cause an overt “stroke” syndrome but could, over time, cause injury and disruption to cortico-basal and subcortical loop circuits: in effect, a “disconnection syndrome”. To this pre-existing burden of brain injury may then be added the further insult of a cortical stroke. Even in the absence of significant pre-existing “small vessel” injury, a large-vessel stroke may cause injury to any one of a number of key “nodes” within these loop circuits: the cortex, white matter tracts, or basal ganglia. Crucially to our purposes, these loop circuits are thought to play a key role in motor learning processes. This may have important consequences in clinical practice, since impairment in learning ability may attenuate a patient’s ability to respond to rehabilitation and thereby act to limit recovery.

1.2.3. Dopamine augmentation of rehabilitation in stroke: a theoretical background

1.2.3.1. How might dopamine enhance rehabilitation interventions?

Although there is considerable uncertainty about precisely how disparate brain structures interact to facilitate motor learning, it is clear that the basal ganglia play a key role in this process. Dopamine is a key modulator of basal ganglia function. It is thought to play a number of important roles in the control of movement and in learning processes, including the selection and termination of motor programmes for skilled movements (Nambu, 2008, Leblois et al., 2006), encoding the “value” of a reward (Wise, 2004), or “stamping in” associations between stimulus and response (Wise, 2004). More recently, it has been proposed that phasic dopamine release acts as an “alerting signal,” prompting the orientation of conscious attention and cognitive processing towards salient environmental cues and increasing general arousal and motivation (Bromberg-Martin et al., 2010). There has therefore

been considerable interest in the possibility of using dopamine agonists as an adjunct to standard rehabilitation interventions in stroke.

Attention has focused in particular upon the use of Levodopa, an orally-administered precursor of dopamine. This crosses the blood-brain barrier before being metabolised to dopamine centrally, resulting in a rise in brain dopamine levels (Berends et al., 2009). Co-careldopa is a combined preparation of levodopa 100mg with a peripheral DOPA-decarboxylase inhibitor, Carbidopa. Carbidopa reduces peripheral levodopa metabolism, thereby maximising the central bioavailability of levodopa (Nutt and Fellman, 1984). Peak plasma levels of levodopa occur between 30minutes to 2hours after a single oral dose of co-careldopa, with a plasma half-life of 1 to 3 hours.

Several trials have evaluated whether administering co-careldopa might enhance the efficacy of standard rehabilitation approaches such as physiotherapy and occupational therapy. A systematic review (Berends et al., 2009) found only two trials of this drug in recovery of walking function after stroke (Sonde and Lökk, 2007, Scheidtmann et al., 2001). Scheidtmann et al. (2001) reported an improvement in mean Rivermead Motor Assessment score of 6.4 points after a three-week course of levodopa, compared with an improvement of 4.1 points with placebo ($p=0.004$). Sonde and Lökk (2007) tested three drug regimens (levodopa, d-amphetamine, or levodopa and d-amphetamine) delivered over ten rehabilitation sessions, and found no improvement in motor function or independence in activities of daily living for any of these treatments when compared with placebo.

A number of smaller studies have also addressed the effect of dopaminergic agents on other aspects of recovery such as independence in activities of daily living and upper limb function (Lökk et al., 2011, Engelter et al., 2010, Rosser et al., 2008b, Restemeyer et al., 2007, Floel et al., 2005, Zorowitz et al., 2005, Acler et al., 2009a). Many were limited by small sample sizes (Restemeyer et al., 2007, Acler et al., 2009a) or comparatively short follow-up (Restemeyer et al., 2007), or administered only single doses of co-careldopa (Restemeyer et al., 2007, Floel et al., 2005). Some recruited patients months or years after stroke (Restemeyer et al., 2007, Acler et al., 2009a).

A wide range of outcome measures has been used to measure response to levodopa treatment, including: the Barthel Index (BI); NIHSS; the Functional Independence Measure (FIM); and walking speed. Unfortunately, the lack of a standardised approach to outcome measurement greatly complicates meta-analysis in this area. Several trials have demonstrated benefit with dopamine on outcome measures including: the BI (Lokk et al., 2011); NIHSS (Lokk et al., 2011); FIM (Engelter et al., 2010); walking speed and manual dexterity (Acler et al., 2009a); procedural motor learning (Rosser et al., 2008b); and motor memory (Floel et al., 2005). However, others have found no improvement in length of hospital stay (Zorowitz et al., 2005), cognitive and motor function (Zorowitz et al., 2005), or upper limb function (Restemeyer et al., 2007). In summary, there is certainly a strong theoretical basis to suggest that taking co-careldopa in conjunction with therapy sessions could enhance the effects of rehabilitation and lead to an improvement in motor recovery: but high-quality evidence to support the implementation of this intervention in routine clinical practice has hitherto been lacking.

1.2.3.2. Dopamine Augmented Rehabilitation in Stroke: the first large-scale trial of levodopa use in stroke rehabilitation

In 2009 the Efficacy and Mechanism Evaluation Programme of the National Institute for Health Research approved funding for a large double-blinded randomised controlled trial to assess the impact on recovery of walking ability of administering co-careldopa or placebo in conjunction with standard multi-disciplinary rehabilitation following stroke. The protocol for this trial, (DARS: ISRCTN99643613), has been published elsewhere (Bhakta et al., 2014) and its methods will be described in detail in the Chapter 2. However, the trial did not show any significant benefit of administering co-careldopa on participants' ability to walk 10m or more independently at up to a year after randomisation.

1.2.3.3. Original aim of this Thesis: exploring factors that influence a patient's response to Levodopa

This Thesis was set in the context of the DARS trial, and utilises data derived from this study. The aims of this work have altered somewhat since the inception of the trial. At the time of trial set-up, it was believed that co-careldopa would show an effect in enhancing motor recovery in the sample as

a whole. However, a fundamental assumption of interventional drug trials is that patients within the sample are homogeneous and that any benefit from treatment will be approximately the same for the whole group. This ignores the possibility that sub-populations within the sample may derive a greater- or less-than-average benefit from treatment (Dorresteijn et al., 2011). Standard trial analyses, which report an effect size for the whole sample, may thus fail to detect clinically-important variations in sub-groups of patients (Young et al., 2005). Since levodopa is known to act upon the basal ganglia, it was felt to be possible that any response to levodopa might be heavily attenuated in patients who had experienced a stroke that included these structures; or conversely, that preservation of the basal ganglia might be a pre-requisite for responding to levodopa-linked rehabilitation. An understanding of how structural brain impairment (as seen on CT imaging) might influence a patient's response to levodopa therefore could have been useful in informing future clinical guidelines about which patients might benefit from this treatment.

The results of the DARS trial are currently in press. Unfortunately, co-careldopa did not prove to be effective in enhancing motor recovery in the moderately-disabled sample of stroke survivors enrolled by DARS. Exploratory analyses showed no evidence of a differential effect of levodopa with stroke type (infarct versus haemorrhage), gender or age. Nor was there evidence of a beneficial effect of levodopa on secondary outcome measures, such as hand function or independence in activities of daily living. Hence, on the basis of the largest randomised controlled trial of levodopa-augmented stroke rehabilitation to date, this intervention cannot be recommended for routine clinical use.

1.2.3.4. Why did Levodopa prove ineffective in enhancing stroke rehabilitation?

Although these results are disappointing, the question of whether levodopa might enhance recovery of walking ability following stroke has at least been answered in a robust clinical trial that is likely to stand as the definitive study of this intervention. However, the mechanism by which levodopa might influence recovery was biologically plausible, and several smaller trials had suggested a positive effect on motor function and other outcomes. It is

therefore worth considering why levodopa did not deliver measureable clinical benefit when applied to a large sample of patients. As we have seen above, motor learning and other cognitive functions are crucially dependent upon interactions between disparate brain structures. Although the putative action of co-careldopa is as a modulator of basal ganglia functions, any effect is likely to be subsumed *in vivo* by the impact of more generalised brain injury both as a result of the stroke itself and as a consequence of a prior burden of microvascular disease. Motor learning ability may be heavily attenuated by both the global burden of brain injury, and by the impairment in cognitive function as a result of this. Furthermore, the efficacy of rehabilitation interventions may be reduced by other impairments in bodily functions that are ostensibly not directly related to the stroke itself, such as musculoskeletal pain, fatigue, and depression.

1.2.4. Other impairments that might influence recovery from stroke

1.2.4.1. Musculoskeletal pain

Musculoskeletal pain is common in the general population (around 15% prevalence; Keenan et al. (2006)), but particularly so amongst stroke survivors (prevalence of up to 47%; Hettiarachchi et al. (2011)). The combined impact of these two pathologies is considerable, and greater than the individual impact of either condition occurring in isolation (Hettiarachchi et al., 2011). For example, having left hip pain without a co-existing stroke confers a ten-fold increase in the odds of reporting difficulty in standing and walking (Hettiarachchi et al., 2011), whereas an isolated right hemiparesis as a result of a stroke confers a five-fold increase in the odds of reporting problems with standing and walking (Hettiarachchi et al., 2011). When both impairments are present simultaneously, the odds for reporting problems in standing and walking is increased by almost fifty-fold (Hettiarachchi et al., 2011). This suggests that the combined effect of both impairments (right hemiparesis and contralateral hip pain) is far more disabling than either impairment alone.

1.2.4.2. Fatigue

Fatigue is a debilitating consequence of stroke, but estimates of prevalence vary enormously (from 23% to 75%) depending on the characteristics of the sample studied (Choi-Kwon and Kim, 2011). Definitions of “post stroke fatigue” (PSF) vary (Lynch et al., 2007, Staub and Bogousslavsky, 2001) and there is even debate about whether PSF represents a single construct or a common manifestation of a variety of underlying processes (Choi-Kwon and Kim, 2011, Wu et al., 2015). A recent systematic review found no evidence of an association between the onset of fatigue and the presence of white matter lesions and brain atrophy, and mixed evidence of an association between fatigue and stroke laterality and location (Kutlubaev et al., 2012). Nevertheless, fatigue may reduce a patient’s ability to participate in rehabilitation.

1.2.4.3. Depression

Depression occurs in up to a third of stroke survivors (Hackett et al., 2005, Kutlubaev and Hackett, 2014) but the interaction between mood and disability is complex. A small study conducted some years ago in an Irish sample found that 20% of patients met the criteria for major depression (Cassidy et al., 2004). However, the presence of depression did not seem to influence final rehabilitation outcomes; the stronger predictor of rehabilitation outcome was the baseline Barthel Disability Score (Cassidy et al., 2004). Depression is associated with stroke severity and disability (Aben et al., 2002, Appelros and Viitanen, 2004, Brown et al., 2012, Desmond et al., 2003, Eriksson et al., 2004, Kotila et al., 1998, Pohjasvaara et al., 1998, Townend et al., 2010, Verdelho et al., 2004, Kutlubaev and Hackett, 2014) but the directionality and causality of this relationship is unclear. It is possible that an extensive stroke leads to severe disability, with depression as a consequence of this; it is equally plausible that depression may attenuate a patient’s motivation to participate in rehabilitation, thereby increasing the risk of an adverse outcome (Kutlubaev and Hackett, 2014).

1.2.5. Might an assessment of brain structure contribute to predicting prognosis in rehabilitation?

Although the initial intention of this Thesis (to explore the impact of structural brain impairment upon a patient's ability to respond to co-careldopa-augmented rehabilitation) was superseded by the negative results of the DARS trial, the data-set that DARS made available represents an important opportunity to explore how patients recover from stroke, and to begin to develop models that might contribute to more accurate prognostication. In particular, the availability of brain imaging for a large cohort of patients allows an exploratory analysis of how impairment in brain structure (as seen on CT) might interact with other impairment in bodily structures (musculoskeletal pain) and functions (cognition, fatigue, and mood) to produce limitation of motor recovery.

1.2.5.1. The current role of brain imaging in rehabilitation assessment

At present, a rehabilitation assessment usually begins with a detailed assessment of a person's physical impairment, and seeks to understand how this interacts with contextual factors (both personal and environmental) to produce limitation of activities and restriction of participation (Gutenbrunner et al., 2006). However, the role of neuroimaging in predicting rehabilitation outcomes is unclear (Stinear and Ward, 2013). As such, although CT scanning remains useful in guiding the acute management of patients presenting with a clinical stroke syndrome, a detailed review of imaging does not at present form a routine part of prognostication in stroke rehabilitation (Stinear and Ward, 2013).

1.2.5.2. The use of advanced imaging techniques in rehabilitation assessment

It has been suggested that MRI scanning may allow impairment in brain structure (injury to the corticospinal tract) to be linked to impairment in bodily function (limb weakness), and thus to limitation of activity (walking) (Lee et al., 2005, Tang et al., 2010). There is, however, there is no consensus on the value of this approach in predicting walking ability (Dawes et al., 2008). More recently, an algorithm comprising a combination of clinical outcome

measures, neurophysiological assessment of corticospinal tract integrity, and MRI assessment of the posterior limb of the internal capsule has been used to predict recovery of upper limb function after stroke (Stinear et al., 2012). However, the value of such an approach in predicting walking ability has not been established. Functional MRI may be helpful in predicting recovery of certain functions such as language, (Saur et al., 2010) but it has not been widely applied in routine clinical practice (Stinear and Ward, 2013).

1.2.5.3. Might CT imaging contribute to prediction of rehabilitation outcomes?

The notion of combining an assessment of impairment in brain structure (on imaging) with impairment in brain function (using clinical assessment) to guide prognostication in rehabilitation is nevertheless appealing. However, if prognostic models are to be useful in clinical practice, then it is advantageous to utilise predictor variables that are easily collected. MRI scanning is more time consuming than CT. It is thus not currently the recommended imaging modality for acute stroke, due to the difficulties of managing safely patients who are unwell whilst images are being acquired, and the requirement for rapid imaging to inform time-sensitive treatment decisions such as thrombolysis. More specialised MRI modalities such as fMRI may have the potential to provide a more detailed assessment of brain structure, but are not yet in routine clinical use. CT scanning, although not without its limitations, is thus the investigation of choice for acute stroke and is routinely performed in all patients presenting with symptoms suggestive of stroke. Thus, if imaging parameters are to be included in a prognostic model, then variables derived from CT imaging are more likely to be routinely available than those that rely upon MRI or fMRI. However, the use of CT imaging in prognostic modelling depends upon a standardised and clinically-meaningful method of coding scan findings. Several such templates exist, for both ischaemic stroke and ICH, and will be discussed below.

Part 1.3 Standardised image analysis instruments

1.3.1. Mechanisms of brain injury in stroke

The use of brain imaging as a predictor of rehabilitation outcomes is not at present widespread. The association between imaging markers of brain injury and motor recovery will be one of the key considerations of this Thesis. However, in order to understand radiological descriptors of tissue injury and how imaging findings evolve over time, it is helpful to first review the cellular mechanisms that underpin these processes.

1.3.2. Mechanisms of brain injury in ischaemic stroke

1.3.2.1. Initial cascades of brain injury following acute ischaemia

In ischaemic stroke the initial cause of tissue injury is hypoxia. This unleashes several cascades of events that lead to neuronal injury and cell death. Some of these pathways cause immediate injury, others result in a delayed insult over subsequent hours and days (Albert-Weisenberger et al., 2013). Acute ischaemia initially results in an inability to maintain transmembrane ion gradients, and water influx causing cell swelling (“cytotoxic oedema”; Meuth et al. (2009)). More prolonged ischemia results in the production of free radicals and changes in intracellular pH (Albert-Weisenberger et al., 2013). Oxidative stress stimulates the release of glutamate, causing the activation of lytic enzymes (Iadecola and Anrather, 2011). Tissue perfusion may be further compromised by microvascular thrombosis as a result of dysfunction of the vascular endothelium (Stoll et al., 2008). In the most severe cases disruption of the blood/ brain barrier may lead to vasogenic oedema, with a life-threatening rise in intracranial pressure (Ayata and Ropper, 2002). In the post-acute phase, a variety of immune-cell-based mechanisms also contribute to tissue injury (Albert-Weisenberger et al., 2013, Iadecola and Anrather, 2011).

1.3.2.2. Evolution of CT findings after acute ischaemic stroke

Cellular injury following stroke is a dynamic process that evolves over time. The same is true of the imaging findings that result from brain ischaemia. In the early stages of an infarct, swelling as a result of cytotoxic oedema may be seen on plain CT as effacement of the sulci overlying the infarct (Wardlaw and

Mielke, 2005). The increased water content of ischaemic brain tissue causes reduced attenuation of the x-ray beam, and therefore an apparent reduction in density of ischaemic tissue relative to surrounding brain. In its earliest stages this may be visible as loss of the normal differentiation at interfaces between grey and white matter (the basal ganglia, frontoparietal cortex, and insular ribbon) (Grotta et al., 1999, Wardlaw and Mielke, 2005), or later as an area of “hypodensity”. The presence of an acute clot within a vessel may cause it to assume a hyperattenuated appearance (Wardlaw and Mielke, 2005), which, if visible, may suggest an infarct involving that vascular territory even if other signs of parenchymal ischaemia are not yet obvious. As an infarct matures, more marked hypodensity develops, with both the cortex and white matter of the infarcted zone appearing darker than normal on CT (Grotta et al., 1999). Early signs of ischaemia are often subtle; a systematic review found inter-observer agreement for these signs to be poor to moderate (Wardlaw and Mielke, 2005). However, these findings were based on evidence that in some cases dates back to the early 1990s (Wardlaw and Mielke, 2005). The technology for both acquiring and viewing imaging has changed considerably since then. For example, only one study reviewed images in digital format on computer workstations (as is current practice); the remainder using lightboxes to view images printed on to films (Wardlaw and Mielke, 2005). Whether subsequent advances in technology have improved detection rates for early ischaemic changes remains unknown (Wardlaw and Mielke, 2005). Interestingly, blinding observers to patients’ clinical symptoms did not affect detection rates for early ischaemic change (Wardlaw and Mielke, 2005).

1.3.2.3. Brain imaging findings and prognosis after stroke

Although scanning provides at best a proxy assessment of clinical impairment after stroke (Kobayashi et al., 2009), the presence of *any* visible infarct on a baseline CT scan is associated with an increased risk of dependency (Wardlaw et al., 1998). It is also positively correlated with adverse functional outcome on the BI, Glasgow Outcome Scale (GOS), and NIHSS (Saver JL et al., 1999), although other studies have found that models incorporating imaging variables are not superior to those incorporating clinical variables alone (Reid et al., 2010). The initial volume of an infarct or haemorrhage

significantly predicts outcome measured by the mRS at 90 days, adding 15% to the explicable variance compared with a model comprising age and NIHSS score alone (Vogt et al., 2012).

1.3.3. Mechanisms of brain injury in ICH

1.3.3.1. Causes of ICH

ICH is typically the result of spontaneous rupture of arterioles that have sustained long-term injury to the intima and sub-endothelium due to hypertension, atherosclerosis, or beta amyloid deposition (Qureshi et al., 2009). More rarely, a secondary cause will be found such as a vascular malformation, coagulopathy, or an underlying neoplasm (Macellari et al., 2014). The sites most commonly affected are the cerebral hemispheres, basal ganglia, thalamus, pons, and cerebellum (Qureshi et al., 2009). Up to 60% of those experiencing an ICH will be dead within a year, with half of these fatalities occurring within the first seven days (Sacco et al., 2009).

1.3.3.2. Mechanisms of brain injury in ICH

Tissue injury and cell death following intracerebral haemorrhage occur via a variety of mechanisms, which will be reviewed in detail below. The initial injury is usually a direct mechanical disruption of tissue as a result of the expanding haematoma (Xue and Yong, 2008). Subsequently the release of blood products, inflammatory cells, and proteases into the brain parenchyma causes a widespread inflammatory response (Xue and Yong, 2008). Rupture of a haematoma into the ventricular system with subsequent formation of intraventricular thrombus may also obstruct the normal circulation of cerebrospinal fluid, causing secondary brain injury due to acute obstructive hydrocephalus (Balami and Buchan, 2012).

1.3.3.3. Causes of early deterioration in ICH: haematoma expansion

Deterioration in neurological status is common within the first 48 hours following ICH (Mayer et al., 1994). The most common cause is early expansion of the haematoma. Radiologically, there is no consensus on what constitutes haematoma “expansion”: various definitions have been proposed including absolute increase in volume or percentage change in haematoma

volume relative to initial lesion size (Dowlatshahi et al., 2011). Various cut-offs have been proposed, including absolute volume change by >3ml, >6ml, or >12.5ml (Dowlatshahi et al., 2011), and relative increase by >33% or >40% (Dowlatshahi et al., 2011). Although haematoma expansion is associated with adverse functional outcome (defined using the mRS), the sensitivity and specificity of this predictor for unfavourable outcome change depending upon how “expansion” is defined (Dowlatshahi et al., 2011, Davis et al., 2006). Furthermore, detecting haematoma expansion relies upon serial imaging at two or more time-points. This necessitates additional radiation exposure for the patient, and is thus not routine practice unless there is strong clinical evidence to suggest expansion. Since serial imaging is not universally performed in stroke care, the models presented here will be constructed using only findings from the first available CT scan. For this reason, haematoma expansion will not be included as a possible predictor variable.

The initial volume of a haematoma, however, may be obtained from first imaging and appears to be important in predicting the risk of early expansion. Data from several trials in the Virtual International Stroke Trials Archive (VISTA: Ali et al. (2007)) have shown that a smaller initial haematoma volume is not only independently associated with lower mortality and favourable functional outcomes on the mRS at 90 days, but also with a decreased risk of subsequent haematoma expansion (Dowlatshahi et al., 2010). No haematomas with an initial size of <3ml went on to expand by >6ml, whereas 7% of haematomas of 3-10ml, and 30.4% of those >10ml did so (Dowlatshahi et al., 2010). For an expansion of >12.5ml, the percentages of haematomas expanding were 0% (initial volume <3ml), 1.4% (initial volume 3-10ml), and 16.9% (initial volume >10ml). For expansion of >33% above baseline, 17.6% of haematomas <3ml expanded, versus 22.4% of those 3-10ml and 27.6% of those >10ml (Dowlatshahi et al., 2010).

1.3.3.4. Causes of early deterioration in ICH: rupture of the haematoma in to the ventricular system

Aside from direct pressure effects of the initial lesion, a further burden of brain injury may be imposed by rupture of the haematoma into the ventricular system. This results initially in a widespread ventriculitis. Subsequently, clot

formation within the ventricular system may cause impedance of the normal circulation of cerebrospinal fluid, leading to an obstructive hydrocephalus in up to 50% of patients (Balami and Buchan, 2012). This may contribute further to rising intracranial pressure, causing secondary brain injury (Balami and Buchan, 2012). The mortality attendant with intraventricular extension is greater than that of intracerebral haemorrhage alone, and is estimated at 50-75% (Balami and Buchan, 2012). Thalamic and caudate haematomas have the highest incidence of Intraventricular rupture due to their anatomical proximity to the ventricular system (Hallevi et al., 2008). Although larger haematoma volumes are generally correlated with the risk of intraventricular rupture (Steiner et al., 2006), the actual volume of blood required to precipitate this is lower for thalamic and pontine haematomas versus lobar lesions (Hallevi et al., 2008).

The volume of blood in the ventricles is of prognostic importance, with 20ml or more being associated with poor outcome (Young et al., 1990). A combined volume of parenchymal and intraventricular blood of >40ml also predicts poor outcome on the mRS, with combined volumes of >60ml being predictive of mortality (Hallevi et al., 2009). With a combined haemorrhage volume above 50ml, poor outcome (death or dependency) is universal (Hallevi et al., 2009). The presence of *any* intraventricular extension is also indicative of a poor prognosis, irrespective of volume. Hallevi et al. (2008) found that patients with intraventricular haemorrhage were twice as likely to be dead or dependent (mRS 4-6), and three times more likely to die than those with an isolated parenchymal haemorrhage. In the one large study of recombinant clotting Factor VII for intracerebral haemorrhage, only 20% of patients with intraventricular haemorrhage achieved a good outcome (mRS of 1-3), versus 43% of those with unruptured parenchymal haematomas (Steiner et al., 2006). The presence of intraventricular rupture at any time-point independently predicted death or dependency (mRS of 4-6) at 90 days (Steiner et al., 2006). The Surgical Trial in Intracerebral Haemorrhage (STICH) likewise demonstrated that 31.4% of those without intraventricular extension achieved a good outcome (defined using the GOS), versus 15.1% of those with haematomas that ruptured (Bhattathiri et al., 2006).

1.3.3.5. Evolution of CT findings after acute ICH

The appearance of a haematoma on CT changes as it matures, and may provide a rough estimate of its age (Macellari et al., 2014). An acute haematoma returns a homogenous, hyperdense appearance (Macellari et al., 2014). Over the first 48 hours fluid levels may become apparent within the haematoma, reflecting different rates of coagulation or new bleeding within the haematoma (Macellari et al., 2014). The haematoma itself is typically surrounded by oedematous tissue (Balami and Buchan, 2012). Initially this is vasogenic oedema: a consequence of the leakage of osmotically active proteins from injured vessels in to the brain parenchyma (Balami and Buchan, 2012). There is often an initial rapid expansion of oedema volume within the first 72 hours after haemorrhage (Arima et al., 2009). This may be seen on CT as a rim of hypodensity surrounding the haematoma itself (Macellari et al., 2014).

Subsequently, the presence of highly irritant blood products within the extravascular space stimulates activation of the inflammatory cascade, leading to a second phase of oedema accumulation which may persist over the first few days and weeks after ictus (Balami and Buchan, 2012). Over the next three weeks the appearance of the haematoma on CT imaging slowly becomes less intense, and when intravenous contrast medium is administered the periphery of the lesion may enhance to give a ring-like appearance mimicking an abscess (Macellari et al., 2014). By around nine weeks the lesion assumes its final hypodense appearance on CT (Macellari et al., 2014). Whilst plain CT is used primarily for confirming the diagnosis of ICH, contrast-enhanced CT (CE-CT) may be helpful in judging the risk of further haematoma expansion (Macellari et al., 2014). Leakage of contrast medium from injured vessels in to the haematoma cavity (the “spot sign”) has been shown to predict haematoma expansion and clinical deterioration (Macellari et al., 2014). CE-CT may also be used to detect underlying vascular abnormalities (Macellari et al., 2014). However, CE-CT carries a greater radiation exposure than plain CT, and in some patients a risk of precipitating renal failure or allergic reaction to the contrast medium (Macellari et al., 2014).

In practice, the use of this modality in evaluating acute ICH is therefore limited (Macellari et al., 2014).

1.3.3.6. Models to predict prognosis after acute ICH

Several models have been developed in order to predict which patients might be at risk of adverse outcomes following ICH. These are referenced and summarised in Appendix A. Many include radiological, as well as clinical, predictors. The aims of such models range from informing initial management decisions (such as who might benefit from surgery or intensive care) to providing more accurate prognostic information to patients and their families, or stratifying patients for entry in to research trials. For this reason, all models incorporated predictors that could be collected shortly after admission: typically a combination of radiological and clinical variables. Early mortality remains a key outcome with predictors including age, disordered consciousness on admission, admission GCS, haematoma diameter, volume of haematoma, the presence or absence of intraventricular extension, the presence of hydrocephalus, haemorrhage location (supratentorial versus infratentorial), history of hypertension, pulse pressure, subarachnoid extension, NIHSS score, blood glucose level, displacement of midline structures, mean arterial pressure, vomiting on admission, and the presence of signs of ischaemia on CT. Many of these variables (age, impairment of consciousness, baseline GCS, volume of haematoma, intraventricular extension, presence of hydrocephalus, supratentorial versus infratentorial location, midline shift, pulse pressure, admission temperature, and NIHSS score) are also predictors of functional outcomes measured using the mRS, GOS, or BI. Models therefore tend to include combinations of the same predictors, but with differing categorisation of predictor variables, cut-off scores, or score weightings used to provide additional sensitivity or specificity for mortality or for functional prognosis as required.

1.3.4. The need for a standardised system to code scan findings

As has been shown above, CT findings may offer important prognostic information; at least in terms of predicting the risk of death or dependency in activities of daily living (Wardlaw et al., 1998, Saver JL et al., 1999, Vogt et

al., 2012, Cheung and Zou, 2003, Cho et al., 2008, Lisk et al., 1994, Ruiz-Sandoval et al., 2007, Tuhim et al., 1991). However, the value of early CT scanning as a predictor of walking ability is less certain. If the question of whether or not baseline imaging is a useful predictor of rehabilitation outcome is to be addressed, then a standardised system for analysing scans and coding findings is required. The use of standardised templates to aid scan reporting is thought to improve the accuracy of scan reporting by non-radiologists, since having a systematic approach to reviewing images helps to ensure that subtle signs are not missed (Wardlaw et al., 2007, Wardlaw et al., 2010),

1.3.5. Standardised systems for coding ischaemic stroke

1.3.5.1. Why were standardised coding systems for CT findings in acute ischaemic stroke first developed?

Several standardised templates for reporting scan findings in ischaemic stroke exist: but they were not developed with prediction of functional prognosis in mind. Their development can be traced back to early trials of thrombolysis, which showed that patients with more extensive infarcts carry a greater risk of subsequent haemorrhagic transformation (Kaste et al., 1995, Anonymous, 1995, Anonymous, 1996, Hacke et al., 1995). The ability to determine the extent of an infarct on early imaging was therefore felt to be crucial to the safe implementation of this intervention. This prompted the development of standardised methods for quantifying the extent of ischaemic change on CT scans. Such systems required good inter-observer agreement: not only amongst neuro-radiologists, but also amongst the stroke physicians, neurologists, and acute physicians who it was felt would most likely be involved in making a decision to proceed with thrombolysis (Grotta et al., 1999).

1.3.5.2. The “ $\frac{1}{3}$ MCA rule”

The earliest and simplest system to be developed was the so-called “ $\frac{1}{3}$ MCA rule” (Kaste et al., 1995). This divided the volume of the middle cerebral artery (MCA) territory into 3, with ischaemic change in $>\frac{1}{3}$ thought to pose a higher risk for thrombolysis (Kaste et al., 1995, Hacke et al., 1998). Unfortunately this

system yields only moderate inter-observer agreement (Grotta et al., 1999). When used to determine eligibility for a hypothetical thrombolysis trial, even experienced observers correctly classified only 45% of patients as eligible or ineligible, although prior training did improve the rate of correct classification (von Kummer, 1998). This system does however have reasonable test-retest reliability, with 80% of scans being coded identically on first and second reading (Wardlaw et al., 1999). The utility of the “ $\frac{1}{3}$ MCA rule” in predicting recovery of walking has not been established.

1.3.5.3. The Alberta Stroke Program Early CT Score (ASPECTS)

The difficulty in estimating the extent of ischaemic change using the $\frac{1}{3}$ MCA rule prompted the development of more systematic methods for quantifying MCA ischaemia, in the hope of improving inter-observer agreement (Barber et al., 2000). The Alberta Stroke Program Early CT Score (ASPECTS; Barber et al. (2000)) defines ten regions within the MCA territory: six cortical (numbered M1-M6), and four subcortical (caudate, lentiform nucleus, insular cortex, and internal capsule). From an initial score of ten points, one point is deducted for each region in which signs of acute ischaemia (swelling, hypoattenuation, or loss of grey/ white matter definition) are seen (Barber et al., 2000). A score of 10 therefore denotes no visible ischaemic change, whilst a score of zero indicates ischaemia involving the entire MCA territory (Barber et al., 2000).

1.3.5.4. ASPECTS as a predictor of prognosis

In contrast to the “ $\frac{1}{3}$ MCA rule”, the ASPECTS was intended from the outset to predict functional outcome: albeit dichotomised as “independent” versus “dead or dependent” (Barber et al., 2000). It was anticipated that it would be helpful not only as a decision aid for thrombolysis but also in the selection of patients for trials of future neuroprotective drugs (Barber et al., 2000). Dichotomising baseline ASPECTS (scored on the CT scan performed at presentation) into ≤ 7 or > 7 discriminated independence (mRS 0-2) from dependence and death (mRS 3-6) at 3 months (Barber et al., 2000). Dichotomised ASPECTS was also predictive of symptomatic haemorrhage after thrombolysis ($p=0.012$; Barber et al. (2000)). With knowledge of the affected side, the inter-observer agreement for dichotomised ASPECTS

ranged from $k=0.71$ for radiology trainees to $k=0.86$ for stroke neurologists and $k=0.89$ for neuro-radiologists (Barber et al., 2000). This compared favourably with the $\frac{1}{3}$ MCA rule ($k=0.64$ for radiology trainees, 0.61 for stroke neurologists, and 0.52 for experienced neuro-radiologists). Without knowledge of the affected side agreement was more modest ($k=0.39$ for radiology trainees, 0.69 for stroke neurologists, and 0.47 for neuro-radiologists), but agreement remained higher than that achieved using the $\frac{1}{3}$ MCA rule ($k=0.14$ for trainees, 0.49 for neurologists, and 0.37 for neuroradiologists; Barber et al. (2000)). The sensitivity of dichotomised ASPECTS for good versus poor functional outcome was 0.78 , with a specificity of 0.96 (Barber et al., 2000). Comparable figures for the $\frac{1}{3}$ MCA rule are sensitivity 0.73 , specificity 0.91 .

As a predictor of good/poor functional outcome, the ASPECTS score thus compares favourably with baseline clinical findings using the NIHSS. When dichotomised as ≤ 15 or >15 , the NIHSS has sensitivity 0.69 , and specificity 0.76 for good versus poor outcome (Barber et al., 2000). Since its inception, ASPECTS has been widely used and evaluated. When comparing ASPECTS scored in “real time” clinical practice (i.e. at the point of treatment) by stroke physicians with the score subsequently allocated by a radiologist on later review, there was substantial agreement (weighted $k=0.69$), although the stroke physicians tended to under-estimate the extent of ischaemic change when ASPECTS was >7 , and over-estimate the extent of changes by nearly two points for the worst-affected scans (ASPECTS <3) (Coutts et al., 2004). Although useful in predicting “good/ poor” functional outcome, the ASPECTS remains a comparatively crude assessment of the extent of an infarct. For example, an ASPECTS of 8 indicates ischaemia in two of the ten areas of interest, but does not specify which two (Wardlaw et al., 2010). Also, ASPECTS only codes infarction in the MCA territory. It offers no information on other vascular territories, and nor does it provide accurate coding of features such as oedema, mass effect, lacunar lesions, white matter lesions, and haemorrhagic transformation of infarcts. It does not classify old infarcts and non-stroke lesions.

1.3.5.5. The Acute Ischaemic Stroke Classification Template (AISCT)

The Acute Ischaemic Stroke Classification Template (AISCT), was developed by Professor J.M. Wardlaw et al, of the Brain Research Imaging Centre, Neuroimaging Sciences, Edinburgh (www.bric.ed.ac.uk; Wardlaw and Sellar (1994)). It comprises a series of templates for coding the location and extent of ischaemic change and tissue swelling, which were constructed following a review of over 100 scans (Appendix B; Wardlaw and Sellar (1994)). The AISCT addresses many of the deficiencies of the “ $\frac{1}{3}$ MCA rule” and ASPECTS, and was thus used in DARS with the agreement of Professor Wardlaw.

1.3.5.6. Inter-observer agreement of the AISCT

Amongst experienced neuroradiologists, the AISCT has good inter-observer agreement for infarct size and type ($K=0.78$), excellent agreement for infarct swelling ($K=0.80$), and moderate agreement for haemorrhagic transformation of the infarct (Wardlaw and Sellar, 1994). Amongst radiology trainees, it has moderate to good agreement for infarct size and site, fair to moderate agreement for infarct swelling, and poor to fair agreement for haemorrhagic transformation (Wardlaw and Sellar, 1994). The inter-observer agreement of AISCT system has also been evaluated in a large sample of non-radiologists from a variety of specialties and with prior experience of CT interpretation ranging from <5 years to >10 years (Wardlaw et al., 2007, Wardlaw et al., 2010). The inter-observer agreement of the AISCT is comparable to that of the $\frac{1}{3}$ MCA rule, ASPECTS (area under receiver operator characteristics curve 0.602-0.604 for all scales; Wardlaw et al. (2010)). When the performance of neuroradiologists was compared to that of non-radiologists, the expert observers tended to spot more subtle signs of ischaemia compared with non-experts (Wardlaw et al., 2007), and took longer to read each scan (Wardlaw et al., 2010). More severe ischaemia (hypodensity and swelling) was more reliably detected than subtle signs, and a longer time from presentation to scan also improved detection rates for ischaemic change (Wardlaw et al., 2010). Detection of acute ischaemia was not influenced by scan quality, or by the presence of an old ischaemic lesion (Wardlaw et al., 2010). The AISCT has been used in the Third International Stroke Trial, IST-3 (Sandercock et al., 2012, The IST collaborative group, 2015), and was also adapted for a

subsidiary study to determine if CT or MRI angiography might be used to guide administration of tissue plasminogen activator at up to six hours after stroke (Wardlaw et al., 2014a). The relationship between infarct classification using this system and prognosis for return to walking has not yet been established.

1.3.5.7. How well does the AISCT correlate with clinical impairment?

In DARS, the diagnosis of “stroke” was *clinically* defined (Aho et al., 1980, Bhakta et al., 2014). Patients were not required to have a compatible stroke lesion visible on imaging in order to be eligible to participate. However, the AISCT classification of ischaemic lesions has shown a good correlation with the clinical classification of ischaemic stroke described by Bamford et al. (1991). Radiological lesions that are compatible with a total anterior circulation stroke syndrome (TACS) include infarcts involving: the whole of the cortical MCA territory (Kobayashi et al., 2009); the whole of the cortical MCA territory plus the lateral basal ganglia (Kobayashi et al., 2009); the whole MCA territory (Kobayashi et al., 2009); or more than 50% of the MCA territory plus the anterior cerebral artery (ACA) or posterior cerebral artery (PCA) territory (Mead et al., 2000). Infarcts compatible with a partial anterior circulation stroke syndrome (PACS) include: small cortical infarcts; infarcts of the ACA territory; or a border-zone infarct between the MCA and ACA or MCA and PCA territories (Mead et al., 2000). Lesions compatible with either a TACS or a PACS syndrome include: medium-sized cortical infarcts of about half the MCA territory; or large (>1.5cm) striatocapsular infarcts (Mead et al., 2000). A posterior circulation stroke syndrome (POCS) is compatible with a cortical infarct in the posterior cerebral artery (PCA) territory, or a lesion in the brain stem or cerebellum, including small infarcts to the pons (Mead et al., 2000). A lacunar stroke syndrome (LACS) is compatible with any subcortical lacunar infarct, including in the centrum semiovale (Mead et al., 2000).

With this in mind, it is possible to determine the proportion of patients in a sample who manifest imaging findings that are compatible with their clinical stroke syndrome. Mead et al. (2000) reported a series of 1012 patients with ischaemic stroke, of whom 655 had a recent infarct visible on CT or MRI. Of those with a visible recent infarct, 75% had a lesion compatible with their clinical syndrome (Mead et al., 2000). Breaking down by syndrome, 79% of

patients with a TACS syndrome had a compatible lesion, as did 71% of patients with a PACS syndrome, 83% of those with a POCS, and 73% of those with a LACS (Mead et al., 2000). Figures for the whole sample can be estimated by assuming a “best case scenario (in which all patients with no visible acute infarct at baseline later develop a lesion compatible with their clinical syndrome), and a “worst case” scenario (under which no patients without a visible infarct initially later develop a lesion compatible with their clinical syndrome). Assuming the “best case” scenario the total percentage of patients who would have a compatible lesion would be 84%: 81% for TACS; 81% for PACS; 90% for POCS; and 85% for LACS (Mead et al., 2000). Assuming the “worst case” scenario, the total percentage with compatible scan findings would be 49% (73% for TACS; 48% for PACS; 50% for POCS; and 40% for LACS; Mead et al. (2000)). The positive predictive values of each clinical syndrome for a compatible CT lesion were: TACS 79% (95% CI 68%-87%); PACS 71% (66%-75%); POCS 83% (75%-89%); and LACS 73% (63%-80%) (Mead et al., 2000).

Kobayashi et al. (2009) reported findings from the first 510 patients presenting to the IST-3 trial. Of those with abnormal scans (329/510), figures for those presenting with a compatible clinical syndrome and radiological lesion were: 79% for TACS; 57% for PACS; 33% for POCS; and 40% for LACS. Assuming a “best case” scenario in the whole sample gives: 100% TACS; 80% PACS; 81% POCS; and 62% LACS (Kobayashi et al., 2009). Assuming a “worst case” scenario in the whole sample, the estimates are: 79% TACS; 37% PACS; 14% POCS; and 2% LACS (Kobayashi et al., 2009). This indicates that, if an infarct is not seen on imaging, the clinical syndrome as defined by the Oxford Community Stroke Project classification may provide a reasonably reliable surrogate estimate of its location and size (Mead et al., 2000): at least for TACS and PACS syndromes (Kobayashi et al., 2009). Findings for POCS and LACS were less consistent. This may in part reflect the lower sensitivity of acute CT imaging for lacunar infarcts (Kobayashi et al., 2009). It may also reflect the mis-assignment of patients to the LACS category by the computer algorithm used by IST-3: in the sample reported by Kobayashi et al, around 20% of those judged to have a “LACS” syndrome in fact had a small cortical

or striato-capsular infarct that would be more consistent with a PACS syndrome (Kobayashi et al., 2009).

1.3.6. Standardised systems for coding ICH

1.3.6.1. What imaging variables are of prognostic importance after ICH?

Several imaging variables are associated with poor functional outcome from intracerebral haemorrhage. These include: volume of intracerebral haemorrhage (Cho et al., 2008, Godoy et al., 2006, Ruiz-Sandoval et al., 2007, Tuhim et al., 1991, Lisk et al., 1994, Mase et al., 1995, Schwarz et al., 2000, Shaya et al., 2005); intraventricular extension (Godoy et al., 2006, Ruiz-Sandoval et al., 2007, Tuhim et al., 1991, Portenoy et al., 1987, Mase et al., 1995, Schwarz et al., 2000, Hallevy et al., 2002, Bhattathiri et al., 2006); presence of hydrocephalus (Shaya et al., 2005); supra-tentorial versus infra-tentorial location (Ruiz-Sandoval et al., 2007); and midline shift (Hallevy et al., 2002). As discussed previously the utility of these variables in predicting other outcomes such as recovery of walking ability remains unknown. The classification systems for ischaemic stroke that were described above either do not include detailed coding of ICH (Barber et al., 2000, Kaste et al., 1995), or code ICH primarily in the context of haemorrhagic transformation of an underlying infarct (Wardlaw and Sellar, 1994). Additional coding for specific ICH variables that may be of prognostic importance is therefore required. A variety of classification systems have been developed, for purposes as diverse as enrolling patients in to clinical trials of surgical haematoma evacuation (Bhattathiri et al., 2003, Mendelow et al., 2005), judging the effects of interventions such as blood pressure reduction (Delcourt et al., 2012, Delcourt et al., 2010) or recombinant factor VII for intracerebral haemorrhage (Mayer et al., 2005a, Mayer et al., 2005b), or the recording and reporting of intracerebral haemorrhage as an adverse event during trials of thrombolysis for acute myocardial infarction (Gebel et al., 1998).

1.3.6.2. Assessment of haematoma volume

All of the standardised systems for coding ICH that were outlined above incorporate an assessment of haematoma volume (Bhattathiri et al., 2003, Mendelow et al., 2005, Delcourt et al., 2012, Delcourt et al., 2010, Mayer et

al., 2005a, Mayer et al., 2005b, Gebel et al., 1998), but methods for deriving this vary. Computer-assisted planimetry uses computer software to estimate the surface area of each axial slice, which is then multiplied by slice thickness and summed to provide an estimate of volume (Gebel et al., 1998, Delcourt et al., 2012, Delcourt et al., 2010, Mayer et al., 2005b, Mayer et al., 2005a, Kosior et al., 2011). These software packages are rarely available in routine clinical practice (Divani et al., 2011). A more straightforward, albeit less accurate, method for estimating the volume of a haematoma utilises the formula:

$$\text{Volume} = \frac{A \times B \times C}{2}$$

in which *A* is the greatest dimension of the haemorrhage on axial imaging; *B* is the greatest diameter at 90° to *A*; and *C* is the approximate number of axial imaging slices on which the haematoma is visible, multiplied by slice thickness (Kwak et al., 1983, Kothari et al., 1996). This “ABC/2 method” has been widely used (Bhattathiri et al., 2003, Mendelow et al., 2005, Mendelow et al., 2011, Hemphill et al., 2001, Ruiz-Sandoval et al., 2007, Vogt et al., 2012, Sloan et al., 1995, Claassen et al., 2001, Leira et al., 2004, Lisk et al., 1994), but it assumes that the 3-dimensional shape of a haematoma approximates to an ellipsoid (Divani et al., 2011). It therefore provides a less accurate estimation of the volume of irregularly-shaped haematomas than computer-assisted planimetry (Kothari et al., 1996, Divani et al., 2011), with reports of an error of up to 8% in such lesions (Divani et al., 2011). The measurement error for the “*C*” term has also been shown to increase with each millimetre of axial slice thickness (Divani et al., 2011). These sources of error may be problematic in studies that rely upon detecting *change* in haematoma volume over two or more imaging series acquired at different time points (Divani et al., 2011): but in this Thesis only the initial haematoma volume will be considered. Furthermore, the ABC/2 method can be used with little training and without the need for computerised post-acquisition image processing. For these reasons, this was the method chosen for estimating haematoma volume in DARS.

1.3.6.3. Systems for coding other ICH characteristics

Numerous systems have been proposed to classify other lesion characteristics such as haematoma location, and the presence of intraventricular extension, hydrocephalus, or mass effect (Bhattathiri et al., 2003, Gebel et al., 1998, Lisk et al., 1994).

Gebel et al. (1998) classified haemorrhage location as parenchymal, subdural, intraventricular, or subarachnoid. The site was coded as frontal, parietal, occipital, thalamus, basal ganglia/ internal capsule, cerebellar, and brainstem (Gebel et al., 1998). They noted haematoma characteristics such as the presence of a blood/ fluid level, and the haematoma appearance (confluent, mottled, or hypodense) (Gebel et al., 1998). Signs of mass effect were noted including: ventricular effacement; horizontal or vertical shift of the pineal gland; effacement of the cisterns; and subfalcine or transtentorial herniation of brain structures (Gebel et al., 1998). Hydrocephalus and intraventricular extension were classified as mild, moderate, or severe, with descriptors of each provided (Gebel et al., 1998). The presence of any other brain lesion (cortical or lacunar infarct, atrophy, periventricular leucomalacia, underlying mass lesion) were noted (Gebel et al., 1998). The volume of the largest haematoma was calculated using computer-assisted planimetry: and thus relied upon software that would not be available for use in routine clinical practice (Gebel et al., 1998).

Lisk et al (Lisk et al., 1994) classified the site of the haemorrhage as putaminal, thalamic, or lobar; ventricular enlargement, ventricular extension, and mass effect were graded as “none”, “slight”, “moderate”, or “severe”; subarachnoid blood was coded as present or absent; the volume of haemorrhage was calculated using the ABC/2 method; and the largest diameter of the haemorrhage was noted (Lisk et al., 1994).

Although comprehensive, the methods of Gebel et al. (1998) and Lisk et al. (1994) have not been validated for inter-observer agreement. The system proposed by Bhattathiri et al. (2003), however, has a published assessment of inter-observer agreement. This system calculates volume (using the ABC/2 method), records the subjective presence or absence of hydrocephalus, a

measurement of midline shift, depth of the haematoma from the cortical surface, and haematoma location (classified as frontal, temporal, parietal, occipital, basal ganglia, internal capsule, and thalamus) . When inter-observer agreement for this system was evaluated (in two consultant neurosurgeons, two neurosurgical trainees, one consultant neuroradiologist, and one radiology trainee), there was good agreement for side of haemorrhage ($k=0.87$), lobar (versus deep) origin ($k=0.78$), involvement of the basal ganglia and thalamus ($k=0.85$), and the presence of intraventricular extension ($k=0.82$). However, agreement was moderate ($k=0.44$) for the presence of hydrocephalus (Bhattathiri et al., 2003). The system developed by Bhattathiri et al. (2003) will therefore form the basis of the haemorrhage classification system used for DARS, albeit with one minor addition not included in the original: the presence or absence of intraventricular extension as a dichotomous variable.

Part 1.4 Developing a prognostic model to predict recovery of mobility after stroke

1.4.1. The need to predict specific rehabilitation outcomes

1.4.1.1. What is a “prognostic model”?

A prognostic model is a means of combining a series of predictor variables in order to generate an estimate of the risk of a specified outcome for an individual patient (Steyerberg et al., 2013). As detailed above, there has been substantial research interest in predicting outcome after both ischaemic stroke and ICH. However, many existing models focus upon predicting the risk of death (Broderick et al., 1993, Hemphill et al., 2001, Cheung and Zou, 2003, Godoy et al., 2006, Tuhim et al., 1991, Lisk et al., 1994, Cho et al., 2008, Ruiz-Sandoval et al., 2007, Bhattathiri et al., 2006), the risk of adverse events such as post-thrombolysis haemorrhage (Kaste et al., 1995, Hacke et al., 1998), or crude “dependent/ independent” functional outcomes such as the mRS (Cheung and Zou, 2003, Lisk et al., 1994, Hallevy et al., 2002, Portenoy et al., 1987, Schwarz et al., 2000).

1.4.1.2. How useful are existing prognostic models in rehabilitation practice?

Whilst such models might be of value in guiding treatment decisions in the acute stages of care, their value in the rehabilitation phase is limited. In rehabilitation practice, patients and their families are usually concerned about the recovery of specific activities (Craig et al., 2011). Although measures such as the NIHSS have some relationship to broad functional outcome, they are of limited value in predicting whether an individual patient might be able to achieve goals such as walking, using a computer, or holding a conversation (Stinear and Ward, 2013).

1.4.1.3. The potential uses of models to predict specific rehabilitation outcomes

The ability to offer a more detailed rehabilitation prognosis than “death or dependency” would be enormously useful for rehabilitation professionals and patients and their families. In particular, the ability to walk independently is a

crucial determinant of whether or not a patient is likely to be able to return to independent living or require institutional care (Craig et al., 2011). Being able to predict reliably who might walk again would therefore allow the rehabilitation team to deliver a more accurate prognosis when speaking to patients and their families (Craig et al., 2011). This would in turn allow both patients and the team to set achievable rehabilitation goals, and may also facilitate early planning for discharge thereby minimising the length of a patient's hospital stay (Kwakkel and Kollen, 2013). Accurate prognostication may also allow rehabilitation interventions to be selected and tailored towards the needs of individual patients. For example, those who have the highest probability of regaining independent walking may be more likely to benefit from restorative approaches, whereas for those with a low probability of walking again the focus of rehabilitation would be to teach the use of adaptive and compensatory strategies. Perhaps most interestingly, prognostic models may allow emergent rehabilitation interventions such as robotics to be targeted towards those who stand the greatest chance of responding to these treatments (Kwakkel and Kollen, 2013). However, the derivation and introduction in to clinical practice of a prognostic model poses several challenges.

1.4.2. The challenges of developing prognostic models

1.4.2.1. Clinical versus statistical validity

Of course, fully realising the potential benefits of prognostic modelling will depend upon constructing valid models to predict outcomes that are important in rehabilitation practice. It is therefore important to consider what is meant when one refers to the "validity" of a model. In modelling, "validity" may be viewed as both a clinical and a statistical concept. The clinical validity of a model refers to its ability to predict satisfactorily an outcome that is of clinical importance (Altman and Royston, 2000). This is determined in part by the choice of predictor variables that are fitted to the model. If variables that are intrinsically weak predictors are used, then its predictive ability in practice will be limited. The concept of statistical validity refers to the goodness-of-fit of a model to both the data-set in which it was derived, and to the population of patients to which it is ultimately applied. The two concepts are not synonymous. For example, it is possible to derive a model that is an excellent

fit to data, but which predicts an outcome that is of no clinical interest whatsoever (Vickers and Cronin, 2010). Such a model might be said to have good statistical validity, but limited clinical validity. Conversely, the Framingham model to predict mortality from cardiovascular disease has high clinical validity and is widely used in practice, despite having only modest discriminatory power in some sub-populations (Moons et al., 2009a). Some of the difficulties associated with the development and introduction in to clinical use of prognostic models are well-illustrated by considering the myriad of models developed to predict outcomes in ICH that were discussed above.

1.4.2.2. What is meant by the “clinical validity” of a model?

If the clinical validity of a model depends upon its ability to predict an outcome reliably, then the choice of predictor variables included is a crucial and often difficult judgement when constructing a model. Often this decision is based upon what has already been demonstrated to be of prognostic importance, with variables being entered in to or excluded from the model in order of the strength of their individual association with the outcome of interest (Altman and Royston, 2000). Statistical software packages may allow this process to be semi-automated, with little or no intellectual input in to selection of predictor variables (Altman and Royston, 2000). Unfortunately it does not always follow that predictors selected by a computer algorithm are the most clinically valid, or that they will yield the strongest prediction (Altman and Royston, 2000). Furthermore, choosing candidate predictors based upon a review of previous literature may present a problem if the outcome of interest has not been explored previously. For example, although several models predict survival with disability following ICH (Cheung and Zou, 2003, Lisk et al., 1994, Hallevey et al., 2002, Portenoy et al., 1987, Schwarz et al., 2000, Ruiz-Sandoval et al., 2007, Godoy et al., 2006, Cho et al., 2008, Shaya et al., 2005, Tuhrim et al., 1991, Bhattathiri et al., 2006), the outcome measures they adopted (Jennett and Bond, 1975, Bonita and Beaglehole, 1988) are fairly crude categorical scales that classify disability as mild, moderate, or severe. It is possible that variables that are predictive of dichotomised “death or dependency” outcomes will have little or no validity in predicting specific rehabilitation outcomes such as recovery of walking ability.

1.4.2.3. What is meant by the “statistical validity” of a model?

The construction of a statistical model is simply the first step in the process of development and introduction into clinical use. Many of the models developed to predict prognosis in ICH appeared to perform reasonably well in predicting outcomes in the samples in which they were derived (Ariesen et al., 2005). However, their adoption into routine clinical practice has been limited (Ariesen et al., 2005). It is worth considering why this might be so. Firstly, fitting a model to a single sample without subsequent validation risks over-estimating the predictive value of variables included: so-called “over-fitting” (Altman and Royston, 2000). Models that are over-fitted tend to be unduly complex and may contain predictors that, whilst statistically significant in the derivation cohort, add little or no predictive value when applied to other samples (Altman and Royston, 2000). This may explain why an evaluation of three prognostic models for intracerebral haemorrhage (Hemphill et al., 2001, Ruiz-Sandoval et al., 2007, Cho et al., 2008) found that the predictive power of the GCS alone was similar to that of all three more complex models (Parry-Jones et al., 2013). Furthermore, the properties of an over-fitted model may change when it is applied to a sample of patients with different characteristics to the derivation cohort. This instability may engender mistrust amongst clinicians (Ariesen et al., 2005). For example, models to predict outcome in intracerebral haemorrhage tend to be most reliable in identifying patients with the highest probabilities of death or poor outcomes (Ariesen et al., 2005). In clinical practice however, the majority of patients have a somewhat lower risk of death or disability (Ariesen et al., 2005). The accuracy of models in discriminating moderate from high risk of death or poor outcome is questionable, and raises concerns that patients may be inappropriately assigned to the “poor prognosis” group and thus denied aggressive care (Ariesen et al., 2005).

Statistical models cannot therefore be assumed to be generalisable beyond the data-set in which they were originally derived (Altman and Royston, 2000). For this reason, formal validation is required for all models prior to adoption in to clinical practice (Altman and Royston, 2000). Moons et al. (2015) have proposed a hierarchy of validation, ranging from development only (in which

a model is derived but not validated) to the evaluation of a published model using an independent data-set.

1.4.2.4. Other considerations when developing a prognostic model

Aside from the requirement for both statistical and face validity, the ease with which a model may be used in practice is an important determinant of how widely it is ultimately adopted. The use of predictors that are time-consuming to collect or that are not collected as part of routine practice may severely limit a model's clinical utility. A second consideration is how the model is used to generate a summary risk score. Using a single predictor variable (a *univariate* model) is most straightforward, but in practice univariate models rarely provide sufficient predictive ability (Moons et al., 2015). Combining several predictors in a multivariate model may improve the accuracy of predictions, but the complex scoring systems required to assimilate data from several variables into a summary statistic may be difficult to apply and interpret in practice (Moons et al., 2015). This was certainly the case for several early prognostic models for intracerebral haemorrhage, which derived their output scores using complex algebraic equations (Ariesen et al., 2005).

1.4.3. Aims of this Thesis

The development and introduction into clinical practice of a prognostic model is thus no easy matter. However the rich DARS data-set presents an opportunity for an exploratory analysis of variables that influence recovery of walking ability after stroke, whilst remaining mindful of the above considerations. This Thesis aims to develop a series of models to predict who might recover the ability to walk 10m or more after stroke. It will utilise a combination of clinical predictor variables and markers of impairment in brain structure derived from CT to predict the dichotomous outcome "able to walk independently for 10m or more, with aid if necessary but no standby help" at two time points: eight weeks and six months after randomisation in to the DARS trial.

This work will focus upon the initial development of the models, but not their validation. Nor will their utility in a clinical setting be assessed. It must be stressed that these models are derived "post-hoc" from an existing data-set.

The limitations of this approach are well recognised, and will be discussed in detail in the concluding chapter. The intention of this work is to generate hypotheses that might be explored subsequently in a prospective cohort.

Chapter 2. Methods

Part 2.1 Methods for the DARS trial

2.1.1. What was the DARS trial?

2.1.1.1. Primary objective of the DARS trial

DARS was a multi-centre double-blinded randomised controlled trial of co-careldopa or placebo in conjunction with standard NHS physiotherapy and occupational therapy in patients presenting with a new clinical diagnosis of stroke (Bhakta et al., 2014). The primary objective of the trial was to compare the proportion of patients in the active treatment and placebo arms who were walking independently by eight weeks after randomisation (Bhakta et al., 2014).

2.1.1.2. Primary outcome measure

The primary outcome measure was the self-reported ability of patients to walk ten metres or more (with an aid if necessary) at eight weeks after randomisation (Bhakta et al., 2014). This was defined using the Rivermead Mobility Index (RMI; Collen et al. (1991)): specifically a RMI score of seven or greater, or answer “yes” to question seven. Since co-careldopa proved ineffective in enhancing recovery from stroke, the DARS data-set has been treated as a large observational cohort for the purposes of this Thesis. The process of randomisation and the medication regimen administered are thus of little relevance to this work: they will be outlined here in sufficient detail to provide context, but the full protocol for DARS has been published elsewhere (Bhakta et al., 2014).

2.1.2. Trial setup

2.1.2.1. Sponsorship and sources of funding

The DARS trial was funded by the National Institute for Health Research (Grant reference number 08/43/61; Bhakta et al. (2014)). It was sponsored by

the University of Leeds, and coordinated from the University of Leeds Clinical Trials Research Unit (CTRU; Bhakta et al. (2014)).

2.1.2.2. Trial management and oversight

The management of the trial was the responsibility of the Trial Management Group, with support from the Trial Steering Committee (TSC). An independent Data Monitoring and Ethics Committee (DMEC) was convened to provide oversight as the trial progressed (Bhakta et al., 2014). The DMEC had full access to unblinded trial data, including all end-points and reports of adverse events and serious adverse events. They had the authority to terminate the trial at any point in the light of concerns about safety, or in the event of new evidence emerging that called in to question the clinical equipoise of the trial. The design and conduct of the trial, data analysis, and authorship of any papers that arise from this work are at the sole discretion of the TSC, without influence from the Sponsor or funding body.

2.1.2.3. Trial registration

DARS was registered with the International Standard Randomised Controlled Trial Number (ISRCTN) database (Faure and Hrynaskiewicz (2011); www.isrctn.com, registration number ISRCTN99643613) and the European Union Drug Regulating Authorities Clinical Trial Register (<https://eudract.ema.europa.eu/>: registration number 2009-017925-20) (Bhakta et al., 2014).

2.1.2.4. Legal framework and ethical approval

The Trial was approved by the Medicines and Healthcare Products Regulatory Agency, and was conducted in accordance with: applicable legislation including the European Union Clinical Trials Directive 2001/20/EC (The European Parliament and Council, 2001) and the UK Medicines for Human Use (Clinical Trials) Amendment Regulations, 2006 (The Stationery Office, 2006); the principles of Good Clinical Practice in research as laid down in the Declaration of Helsinki (The World Medical Association, 2013); and the NHS Research Governance Framework for Health and Social Care, 2005 (Department of Health, 2005). All data received, including digital copies of scans, were collected, handled, and stored securely in accordance with the

requirements of the Data Protection Act, 1998 (The Stationary Office, 1998). Ethical approval was granted by the UK National Research Ethics Service (reference 10/H1005/6; Bhakta et al. (2014)). Initial ethical approval was followed by approval of an amendment to the protocol and patient information sheet, to make explicit the intention of the trial team to collect copies of any available neuroimaging for the index admission.

2.1.2.5. Recruiting centres

Participants were recruited from NHS stroke services across all four nations of the United Kingdom. Coordination of site setup, recruitment, and randomisation was managed by The University of Leeds CTRU. The trial opened to recruitment in May 2011; the final participant was recruited in March 2014. The large number of centres that were recruiting to DARS necessitated the establishment of a network of local Principal Investigators (PIs), each of whom was responsible for the conduct of the trial at their site. PIs were consultant physicians, usually in stroke medicine, neurology, or rehabilitation medicine. Their duties included the recruitment and consenting of participants, overseeing the clinical care of patients enrolled in the trial, and the reporting and medical management of adverse events. Screening, consent, randomisation, and follow-up of participants were performed by researchers employed by the Comprehensive Local Research Network (CLRN). However, the DARS trial also necessitated considerable involvement from NHS staff such as ward nurses and therapists: many of whom had had no previous involvement in research. For this reason, a series of face-to-face site initiation visits were arranged, with training in trial procedures provided for key personnel. The training package was developed and delivered by Dr John Pearn (JP: Clinical Research Fellow to the DARS trial) and Ms Lorna Barnard (CTRU Trial Monitor). Ongoing support and advice to sites was also provided throughout the recruitment period by CTRU.

2.1.3. Participants

2.1.3.1. Sample size and power calculation

An initial target of 572 participants was set, providing a 90% power to detect a 50% difference in the proportion of patients walking independently in the

intervention and control arms at eight weeks, using a 5% significance level. This sample size also provided an 80% power to detect a small to moderate effect size (0.3) in key secondary outcomes (Bhakta et al., 2014). However, a slightly higher than expected rate of loss to follow-up at the eight-week time point necessitated continuation of recruitment to a larger sample size of 592 participants in order to retain power for the primary outcome at the specified level (*The DARS Collaborators: data not yet published*).

2.1.3.2. Inclusion criteria

Patients were eligible to participate if they: had sustained a new or recurrent clinically diagnosed ischaemic stroke or ICH (excluding subarachnoid haemorrhage) within 5 to 42 days prior to randomisation; were unable to walk ten metres or more indoors independently (with an aid if necessary, but no physical assistance); had a therapist-completed RMI score of <7; were expected to require rehabilitation treatment; were aged 18 years or above; were able to give informed consent; were able to access continuity of rehabilitation treatment following discharge from hospital (defined as being able to access community-based rehabilitation services within five days of hospital discharge); were expected to be able to comply with the treatment schedule (for example, were able to swallow whole tablets); and were expected to be in hospital for at least the first two doses of trial medication (Bhakta et al., 2014).

2.1.3.3. Exclusion criteria

Patients were deemed ineligible if they: were not expected to survive for two months following stroke; had a diagnosis of Parkinson's disease, severe medical or surgical illness, or severe psychosis; had a known hypersensitivity or contraindication to co-careldopa treatment; had symptomatic orthostatic hypotension; needed physical assistance of at least one person to walk prior to their index stroke due to pre-existing co-morbidities (for example, heart failure, or osteoarthritis); were pregnant, lactating, or a women of child-bearing potential (unless they were willing to use medically approved contraception whilst receiving treatment and for one month after the cessation of treatment); could not walk 10 metres or more indoors prior to their stroke (Bhakta et al., 2014).

2.1.3.4. Patient consent

Written informed consent was obtained from all participants by local stroke physicians, all of whom had received training in trial procedures and in Good Clinical Practice in Research (Bhakta et al., 2014). Provisions were made to allow witnessed consent to be obtained in patients who had capacity to consent but were unable to sign the consent form (for example, due to weakness of the dominant hand).

Part 2.2 Treatment and follow-up

All DARS participants were randomised to receive up to eight weeks of co-careldopa or placebo, in conjunction with usual NHS physiotherapy and occupational therapy sessions (Bhakta et al., 2014). There were four points at which data were collected: at entry in to the trial; at eight weeks after randomisation; at six months after randomisation; and at twelve months after randomisation. The process of data collection and the treatment regimen will be described below.

2.2.1. Timing of assessments

2.2.1.1. Protocol-specified timing of assessments

The timing of assessments was specified in the DARS protocol (Bhakta et al., 2014), and is summarised in Figure 2.1. A baseline assessment was performed on the day of randomisation. Follow-up assessments were then conducted at “eight weeks” and “six months” after randomisation. However, since patients were eligible for enrolment in to the trial between 5 and 42 days after stroke, the actual timing of assessments in relation to the date of the stroke was heterogeneous. For example, when allowance was made for the recruiting window, the “eight week” assessment could be conducted within a range of 61-96 days. In addition, a grace period of ± 5 days was permitted within which visits could be arranged. Hence, the window within which “eight week” assessments could be conducted was 56-101 days after stroke. Similarly, the window for performing “six month” visits (accounting for the recruitment period and a “grace period” of ± 7 days) was 178-229 days after stroke.

2.2.1.2. Definition of data collection points used in this Thesis

This Thesis will refer to the timing of assessments as follows: “baseline” assessment as “T₀”; “eight week” assessment as “T₁”; and “six month” assessment as “T₂”. The protocol also specified a final follow-up visit at up to one year after randomisation (Bhakta et al., 2014). These data are not available at the time of writing, so the latest end-point which will be considered

here is T₂. For consistency, the “twelve month” follow-up will where necessary be referred to as “T₃” (Figure 2.1.).

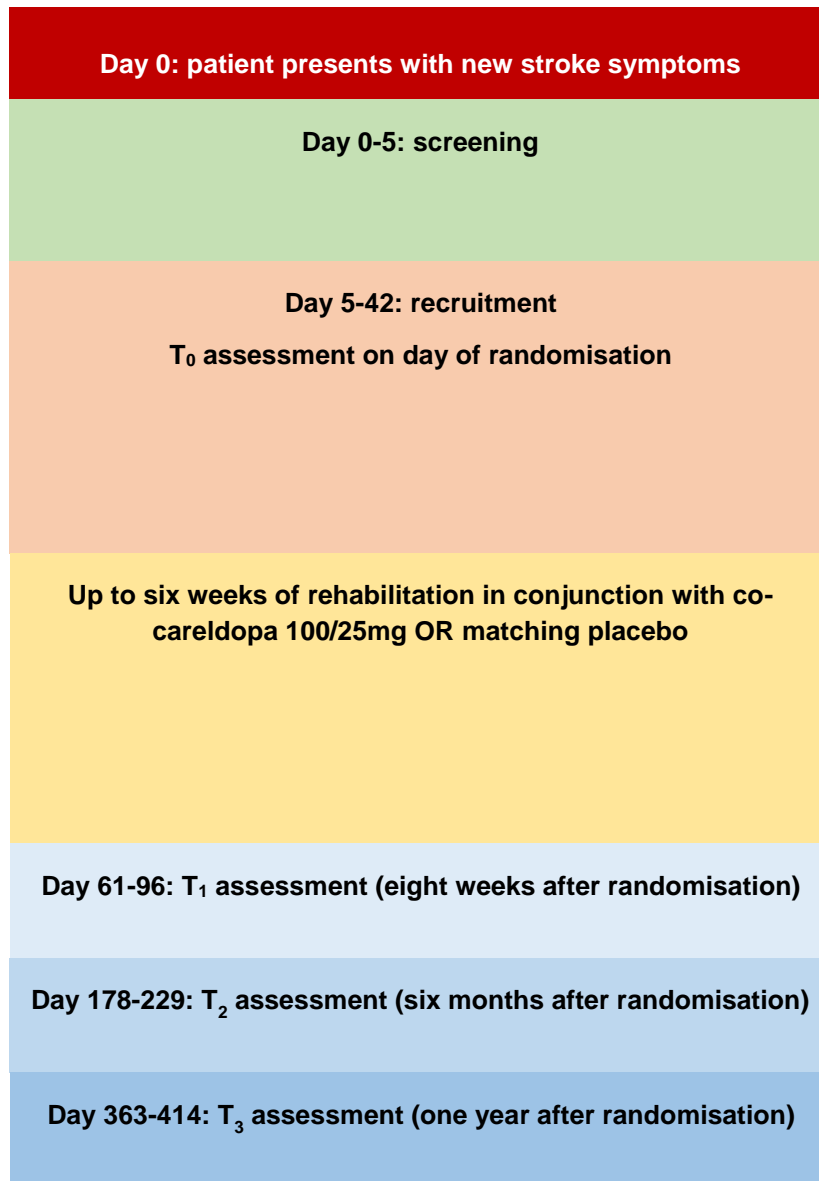


Figure 2.1. Summary of timeline for screening, recruitment, treatment, and follow-up of DARS participants.

2.2.2. Summary of outcome measures used in the DARS trial

2.2.2.1. Overview of data collection

The outcome measures used in DARS, and the timing of their collection, are summarised Appendix C. The primary outcome measure was the RMI (Collen et al. (1991); see Appendix D) This was collected at T₀, T₁, T₂, and T₃. Data

were also collected on possible modifiers of levodopa effect, including psychological morbidity (the 12-item version of the General Health Questionnaire: GHQ-12) (Goldberg and Blackwell, 1970, Goldberg and Hillier, 1979); musculoskeletal pain (the Musculoskeletal Signs, Symptoms, and Pain Manikin: MSK-SSP, Keenan et al. (2006)); fatigue (the Fatigue Assessment Scale: FAS, Michielsen et al. (2003)); and cognitive impairment (the Montreal Cognitive Assessment: MoCA, Nasreddine et al. (2005)). Each will be described below. A range of secondary outcome measures were also collected, encompassing other aspects of physical function (the Abilhand scale; Penta et al. (1998)), independence in activities of daily living (a postal version of the BI (pBI) (Gompertz et al., 1994, Mahoney and Barthel, 1965); and the Nottingham Extended Activities of Daily Living scale (NEADL; Nouri and Lincoln (1987)); health-related quality of life (the EuroQol EQ-5D; Rabin and de Charro (2001)); and global disability (the mRS; Bonita and Beaglehole (1988)). These secondary outcome measures did not form part of the analysis presented here, and thus will not be described in detail.

2.2.2.2. Clinician-reported and self-reported Rivermead mobility Index

The RMI score at T_0 was crucial both for confirming eligibility and for the stratification of participants during randomisation. For this reason, a clinician-completed RMI was obtained at T_0 . This was filled in by the patient's therapist, based upon their knowledge of the patient's capabilities. The RMI was also the primary outcome measure at T_1 , T_2 , and T_3 . However, it was deemed impractical to collect therapist-completed RMI scores at these times, since a proportion of patients (particularly at T_2 , and T_3) would have been discharged from therapy by the time of follow-up. A patient self-reported RMI was therefore collected at T_1 , T_2 , and T_3 as an outcome measure, and also at T_0 to facilitate comparison with the self-reported RMI. Clinician-reported and self-reported versions of the RMI were identical in all respects, with no difference in the wording of items or in their scoring. Nevertheless, the abbreviations C-RMI (for the clinician-scored RMI administered only at T_0) and SR-RMI (self-reported RMI collected at T_0 , T_1 , T_2 , and T_3) will be used in order to make explicit how they were completed. The abbreviation "RMI" will also be used,

when referring to the Rivermead Mobility Index in a context in which it is not necessary to specify the manner of its completion.

2.2.3. Assessment at T₀ and treatment regimen

2.2.3.1. Assessment at T₀

The assessment at T₀ was conducted after consent had been obtained, but prior to randomisation. It included the administration of a basket of standardised outcome measures and basic demographic and clinical information. The demographic information collected included: initials; date of birth; and number of years of formal education. Clinical history included: the date and type (ischaemic or haemorrhagic) of stroke; the date and type (CT or MRI) of any available neuroimaging; whether the patient had received intravenous thrombolysis; and details of the patient's past medical history. Those with ischaemic stroke were prospectively categorised by treating clinicians according to the Oxford Community Stroke Project classification (Bamford et al., 1991). This categorises the patient's clinical stroke syndrome on the basis of their presenting features as TACS, PACS, POCS, or LACS syndromes (Bamford et al., 1991).

In addition to baseline demographic and clinical information, each patient was asked to complete a range of self-reported outcome measures including: the SR-RMI; the pBI; the NEADL; the Abilhand, the EQ-5D; the GHQ-12; and the MSK-SSP (Bhakta et al., 2014). The NEADL, Abilhand, and MSK-SSP were answered with respect to the patient's pre-stroke status: all other self-reported outcome measures were answered in relation to the patient's current abilities (Bhakta et al., 2014). All of these questionnaires could either be self-completed by the patient or directly administered by the researchers if the patient was physically incapable of marking responses. The MoCA was supplied separately, and was administered by the DARS researchers during a face-to-face visit (Bhakta et al., 2014).

2.2.3.2. Randomisation and blinding

Following the T₀ assessment, patients were randomised on a 1:1 basis to receive either co-careldopa 125mg (levodopa 100mg, and a peripheral DOPA decarboxylase inhibitor, carbidopa 25mg) or matching placebo. To ensure that

treatment groups were well balanced patients were stratified (Bhakta et al., 2014) by: recruiting centre; type of stroke (infarct versus haemorrhage); and C-RMI score (0-3, 4-6).

2.2.3.3. Medication regimen

The medication regimen for DARS has been described elsewhere (Bhakta et al., 2014). In brief, one oral tablet of either co-careldopa 125mg or matching placebo was taken 45-60minutes before the start of physiotherapy or occupational therapy sessions. The medication was administered only prior to therapy sessions in which the focus was on motor activity (Bhakta et al., 2014). The timing of dosing reflected the absorption kinetics of levodopa (which reaches peak plasma levels at 30-120minutes after a dose), as well as evidence from previous trials of co-careldopa in stroke (Rosser et al., 2008a, Scheidtmann et al., 2001). Although the protocol stipulated that trial medication should be administered within this optimum time window, it was appreciated from the outset that circumstances may not always permit this. A pragmatic approach was therefore adopted, under which it was deemed permissible to administer the drug between at any time prior to the start of a therapy session if necessary (Bhakta et al., 2014). If a patient had two therapy sessions within three hours of a first dose, then an additional dose of the drug was not given prior to the second session (Bhakta et al., 2014). A maximum of two doses were administered in any 24-hour period (Bhakta et al., 2014). Sufficient medication was supplied to each patient for a maximum course of six weeks of treatment: a duration which reflects that used in previous trials of co-careldopa-augmented stroke rehabilitation (Acler et al., 2009a). If a clinical decision was taken by the treating team that further rehabilitation intervention was no longer required before the conclusion of the six week drug treatment period, then the drug was discontinued at the point of discharge from rehabilitation services. When patients were discharged to the care of a community stroke team, the treating therapist was asked to provide a medication prompt by telephone up to an hour before their planned session. Adverse events were classified in accordance with definitions laid down by the European Union Directive 2001/20/EC (The European Parliament and Council, 2001). Responsibility for detecting, managing, and reporting these

events rested with local PIs. Adherence to the medication regimen was estimated by performing a count of the number of tablets remaining in the medication kit at the end of the treatment period, and reconciling this count with records of the number of drug-linked therapy sessions the patient had received. This was perhaps an imperfect measure of pharmacoadherence, since it relied upon a retrospective review of records which were in some cases incomplete. Pill counts are also time consuming, and may over-estimate adherence since they assume that all tablets that are not present have actually been taken (Chisholm-Burns and Spivey, 2008). However, these methods were felt on balance to provide a more reliable estimate than direct questioning of the patients.

2.2.3.4. Content of therapy sessions

The content of drug-linked therapy sessions was entirely at the discretion of the treating teams according to their assessment of the patients' needs and their routine clinical practice (Bhakta et al., 2014). Therapists were blinded to treatment allocation, and instructions given to ward nurses and therapy staff concerning drug administration were the same for both treatment arms (Bhakta et al., 2014). After each drug-linked session, therapists were asked to complete a short case report form. This detailed the time at which the drug was taken (or the reason for omission of the drug), and an approximate breakdown of the amount of time devoted to upper limb motor therapy, lower limb motor activity, and non-motor work. The purpose of this was to allow the amount of therapy intervention delivered to be ascertained, and to verify parity between intervention and control arms.

2.2.4. Follow-up of participants

2.2.4.1. Follow-up assessment at T₁

Follow-up assessments were conducted by a network of CLRN researchers at T₁, T₂, and T₃. Visits at T₁ were sometimes conducted whilst the participant was still an inpatient. Visits at T₂ and T₃ were usually conducted in the patient's usual place of residence, but allowance was made for participants to be recalled to research clinics where this was local practice. Telephone follow-up was permitted only under exceptional circumstances. A booklet of

questionnaires for the T₁ follow-up was supplied to the patient at the point of randomisation, as part of a patient information pack which also contained a patient information booklet, DVD, and CRFs for therapists to complete in their sessions. The self-reported outcome measures contained in this booklet included: the SR-RMI; the pBI; the NEADL; the ABILHAND; the EQ-5D; the GHQ-12; the MSK-SSP; and the FAS (Bhakta et al., 2014). All were recorded in relation to the patient's current status. The patient was contacted by the researcher prior to the visit, and asked to complete them in advance. If this was not possible, they were administered directly by the researcher during their visit. The Modified Rankin Scale (mRS) (Bonita and Beaglehole, 1988) and the MoCA (Nasreddine et al., 2005) were administered directly (Bhakta et al., 2014). Patients were also asked about any adverse events associated with treatment.

2.2.4.2. Follow-up assessment at T₂ and T₃

The visits at T₂, and T₃ were similar in nature. A booklet of patient self-completed questionnaires was posted out to the patient in advance. As for the follow-up visit at T₁, these comprised: the SR-RMI; the pBI; the NEADL; the ABILHAND; the EQ-5D; the GHQ-12; the MSK-SSP; and the FAS (Bhakta et al., 2014). Once again questions were answered with respect to the patient's current status. The patient was asked to complete them within the week before the researcher's visit, but direct administration by the researcher was permitted if necessary. The mRS (Bonita and Beaglehole, 1988) and the MoCA (Nasreddine et al., 2005) were administered directly by the researcher (Bhakta et al., 2014). Patients were also asked if they had experienced any new and significant medical or surgical illnesses since the last follow-up visit. At the T₃ visit an exit poll was also conducted, asking the patient if they thought they had been in the active or placebo arm of the trial: the purpose of this was to assess the security of blinding procedures.

The DARS data-set thus comprises an extensive set of outcome measures, taken at four time-points. The purpose of the models derived here is to establish whether impairment-level variables (depression, fatigue, musculoskeletal pain, and cognitive dysfunction) may serve as useful predictors of walking ability at T₁ and T₂. However, in order to understand how

these variables were used in analysis, it is important to consider the properties of the scales used to measure them.

Part 2.3 Outcome measurement in the DARS trial

2.3.1. The principles of measurement

2.3.1.1. What is “measurement”?

It is common to make measurements in medicine. For example, when predicting the future risk of stroke in a person with AF, we might wish to measure certain biometric parameters such as their SBP, or serum glucose levels (Wang et al., 2003). These are all easily observable and can be quantified directly by measuring devices, for example a sphygmomanometer, or laboratory tests (Hobart and Cano, 2009). The units in which they are described have precisely the same interval between each gradation. One can therefore be confident that, within the limits of measurement error, a blood glucose level of 10mmol/L is twice as great as a level of 5mmol/L. When intervening to control risk factors for stroke, one might also define success in terms of change in these parameters over time: for example, in reducing a patient’s SBP from 180mmHg to 130mmHg (a fall with a magnitude of 50mmHg). Clinical trials commonly report mean reductions in SBP for different treatments (Rashid et al., 2003). A key property of such measurements is that the interval between each unit change is a fixed, and known, value. Such measurements are thus termed “*interval-level*”, and can be subjected quite legitimately to mathematical operations such as calculation of mean values or the change in value between from one measurement to another.

2.3.1.2. Measuring rehabilitation outcomes

Unfortunately measuring rehabilitation outcomes is not as straightforward as measuring physiological variables, since the property that is under consideration is frequently some aspect of a person (or their physical performance) that cannot be observed or measured directly: for example “mobility,” “fatigue,” “depression,” or “independence” (Hobart and Cano, 2009). These are termed *latent* (hidden) *traits*. Although they are not directly measurable, it is often possible to draw inferences about the degree to which they are present in an individual or a group (Hobart and Cano, 2009). To do this one must first define the property (or “construct”) to be measured, then devise a range of questions (“items”) that map out its possible range on a

continuum or “lesser” to “greater” (Hobart and Cano, 2009). This is what outcome measures such as the RMI, GHQ-12, or FAS seek to do for the constructs of “mobility”, “depression”, and “fatigue” respectively. The properties of the RMI, GHQ-12, mRS, MSK-SSP, FAS, MoCA will be examined below.

2.3.2. Outcome measures used in the DARS trial

2.3.2.1. Primary outcome measure: the Rivermead Mobility Index

The RMI (Collen et al. (1991); Appendix C) was the primary outcome measure selected for DARS. It is a fifteen-item scale, with each item having two possible responses: “yes” (able to do the task) or “no” (unable to do the task) (Collen et al., 1991, Hobart and Cano, 2009). Items are presented in order of increasing difficulty, with the first item being turning over from the back to the side in bed, and the final item running ten meters in four seconds without limping (Collen et al., 1991). Fourteen items rely upon the patient’s self-report; one (question 5) requires the interviewer to observe the patient standing unsupported for ten seconds. In the DARS trial question 5 was operationalised by simply relying upon the patient’s self-report of whether they could achieve this. Each “yes” response scores one point, which are then summed to give a total score out of 15.

2.3.2.2. The Modified Rankin Scale (mRS)

The Rankin Scale is an ordinal hierarchical scale that ranks different levels of disability (Rankin, 1957). The original measure was developed in 1957, and mapped a continuum from Grade I (“No significant disability: able to carry out all usual duties”) to Grade V (“Severe disability: bedridden, incontinent and requiring constant nursing care and attention”; Rankin (1957)). This original metric has since been modified several times. A Grade 0 (“no symptoms at all”) was proposed in 1988 (van Swieten et al., 1988). A Grade VI (denoting death) is also sometimes applied (O’Connor et al., 2011, Uyttenboogaart et al., 2005), although it is unclear when this modification was made or by whom. Here we will use the terms “Rankin Scale” (RS) to denote the original five-level version described by Rankin (1957), and “Modified Rankin Scale” (mRS)

to denote all subsequent variations (Farrell et al., 1991, van Swieten et al., 1988).

Although a commonly used outcome measure in stroke research (Banks and Marotta, 2007, Quinn et al., 2009b), the original descriptions of the RS and mRS did not clearly define parameters for each grade (O'Connor et al., 2011). This led to significant inter-observer variability in how the RS/mRS are scored (Quinn et al., 2009a, O'Connor et al., 2011, Huppert et al., 1988, Quinn et al., 2009c), with the obvious potential to confound results (O'Connor et al., 2011). Several attempts have been made to improve the inter-observer agreement of the mRS: by offering formalised training courses with accreditation on completion (Quinn et al., 2007); the use of semi-structured interviews with pre-specified questions (Wilson et al., 2002, Wilson et al., 2005); and the development of a simplified scoring algorithm (Bruno et al., 2010). The success of these interventions has been mixed. Even after formal training there is considerable inter-observer variability in mRS scores, and also between observers from different countries (O'Connor et al., 2011). A structured interview has been shown to improve the agreement between observers ($k=0.25$ and weighted $k=0.71$ for mRS administered conventionally, versus $k=0.85$ and weighted $k=0.94$ for structured interview; Wilson et al. (2005)), but is time-consuming to administer taking around 15 minutes (Wilson et al., 2002). Use of a short algorithm to assign mRS scores shows comparable reliability to the longer structured interviews, but can be completed in less than two minutes (Figure 2.2.; Bruno et al. (2010)). It was this version that was adopted for use in DARS (Bruno et al., 2010). It is interesting to note that studies validating the use of structured interviews or algorithms to administer the mRS all used small numbers of observers, all of whom had received prior training. In the use of the mRS either delivered "in house" as part of the validation study itself (Wilson et al., 2005) or using existing training packages developed by others (Bruno et al., 2010). It is therefore difficult to unpick the extent to which the observed improvements in inter-observer agreement are attributable to the changes made to the way in which the mRS is administered, or simply to training of the interviewers. Whether similar improvements would be seen if these versions of the mRS

were administered in clinical practice without prior training is not known.

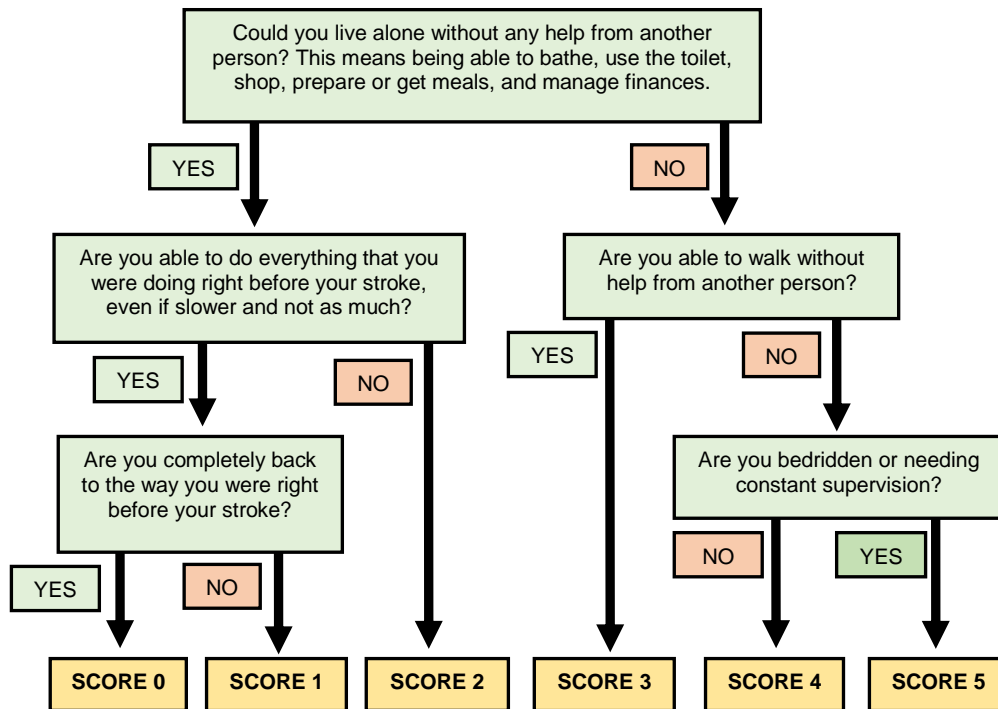


Figure 2.2. Simplified algorithm for assigning mRS score.

Developed by Bruno et al, 2010

A number of cut-off scores for defining “favourable” outcomes on the mRS are reported: ≤ 1 , ≤ 2 or ≤ 3 (Sulter et al., 1999, Steiner et al., 2006). The choice of which to use may influence the results of a trial. For example, post-hoc analysis of a major stroke thrombolysis trial yielded no significant effect for thrombolysis using a cut-off score of ≤ 1 , but a statistically-significant benefit when a cut-off of ≤ 2 was used (Sulter et al., 1999). In trials of interventions to prevent an adverse outcome, a score of 6 (“death”) is sometimes appended to the standard 0-5 metric of the mRS, and outcomes dichotomised as “independent” or “death or dependency”. Several trials in intracerebral haemorrhage have taken this approach, utilising scores of 0-3 to denote a favourable outcome and 4-6 for death or dependency (Steiner et al., 2006, Hallevi et al., 2009). Although adding a “death” term to the metric may appear to be a convenient way to capture the full range of adverse outcomes, it is illogical to regard “death” as lying on the same continuum as “disability”. This illustrates well the perils of arbitrarily altering outcome measures without regard for the underlying construct that the scale is supposed to measure. The mRS will be used here only to describe the distribution of disability in the

sample at T_0 . It will not be entered as a covariate in analyses, and change in mRS scores will not be reported.

2.3.2.3. The GHQ-12

The General Health Questionnaire (GHQ) was originally developed as a global measure of psychological morbidity for use in general practice (Goldberg and Blackwell, 1970, Goldberg and Hillier, 1979). Several variants are available including 12-, 28-, 30, and 60-item versions. The twelve-item version (GHQ-12) was used in DARS. It comprises twelve items: six describing positive psychological states (such as “I feel as though I play a useful part in things”), and six describing negative states (for example, “I have lost sleep over worry”; Hankins (2008a)). The original numerical rating scale scoring system asked patients to rank their experience of these states on a four-point scale: “no more than usual”; “not at all”; “rather more than usual”; and “much more so than usual” (Hankins, 2008b). Each item thus scores between 0 and 3 points, for a total possible score of 36 (Kelly et al., 2008). This was the scoring system adopted in DARS. However, this structure has been criticised for its ambiguity. In particular the response “no more than usual”, when applied to a negative psychological state, may actually be interpreted as denoting the *presence* of this state (Hankins, 2008b). For this reason, updated scoring systems have been devised. The first collapses dichotomises responses as “present” or “absent” (Hankins, 2008b), resulting in a range of scores from 0-12. The second, the “chronic GHQ,” retains the usual 0-3-point scoring system for the “positive” items, but adjusts scoring for the “negative” items such that only the response “not at all” is regarded as healthy (Huppert et al., 1988).

2.3.2.4. The musculoskeletal signs, symptoms and pain manikin (MSK-SSP)

Patients’ self-reported experience of musculoskeletal pain was evaluated in DARS using the MSK-SSP (Keenan et al., 2006). This was developed as a means to facilitate the self-reporting of joint symptoms. Patients are asked to tick boxes on the manikin that correspond to the major joints, to indicate the location of musculoskeletal signs such as joint swelling, or symptoms such as pain or stiffness (Figure 2.3.; Keenan et al. (2006)).

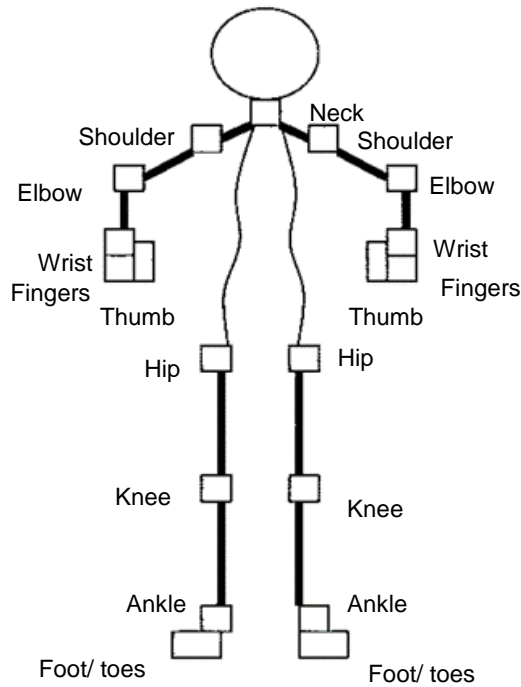


Figure 2.3. The musculoskeletal signs, symptoms, and pain manikin (MSK-SSP)

2.3.2.5. The Fatigue Assessment Scale

A measure of fatigue was included in DARS as a possible effect modifier, although the interaction between post-stroke fatigue (PSF), functional outcome, and other possible effect modifiers is not straightforward. Although there is an association between PSF and dependency in instrumental activities of daily living and poor quality of life, these relationships do not persist once the presence of depression is controlled for (Mead et al., 2007, Wu et al., 2015). There is no apparent association between PSF and cognitive impairment on the Mini Mental-State Examination (MMSE): but the limitations of the MMSE are well known, and there is a suggestion that PSF may be associated with deficits in attention to which the MMSE is relatively insensitive (Wu et al., 2015). Both the absence of a unified definition for PSF, and doubts over whether “PSF” is truly a single construct are problematic when attempting to measure this. Additionally, many of the scales available have been developed for other conditions. A review by Mead et al. (2007) found fifty-five fatigue scales, none of which had been originally developed in stroke populations. Of these, only five (the vitality subscale of the 36-item Short Form

Health Survey, Version 2; the fatigue subscale of the Profile of Mood States; the Fatigue Assessment Scale; the Multidimensional Fatigue Symptom Inventory; and the Brief Fatigue Inventory) had acceptable face validity: other scales included items that could be confounded by the neurological sequelae of stroke (Mead et al., 2007). The Brief Fatigue Inventory was found to contain items that could not be completed easily by patients (Mead et al., 2007). Of the remainder, there were no significant mean differences in scores assigned by different observers: but the Fatigue Assessment Scale (FAS: (Michielsen et al., 2003, Michielsen et al., 2004)) was found to have the narrowest limits of agreement and a high inter-class correlation coefficient of 0.77 (Mead et al., 2007). Unlike the other three scales, which ask about fatigue simply by using a common question stem suffixed by different descriptors (“in the past week I feel run down; worn out; fatigued; sluggish”), the FAS enquires about several different aspects of fatigue such as mental and physical exhaustion and problems concentrating (Mead et al., 2007). Although it has a lower internal consistency than the other scales, it may actually be more useful in clinical practice (Mead et al., 2007). The FAS has ten items, five for physical fatigue and five for mental fatigue. Each is scored as: never (1); sometimes (2); regularly (3); often (4); and always (5). The maximum score is therefore 50. However, assigning a case definition of fatigue based the FAS is difficult, since it depends upon using an arbitrarily-defined cut-off score to distinguish between normal “physiological” fatigue and the more debilitating “pathological” fatigue that may occur in neurological conditions (Duncan et al., 2014). A change in FAS score of 4 points or more has been used to define a clinically-relevant change in fatigue status (Duncan et al., 2014). However, this definition is based upon studies in patients with sarcoidosis (de Kleijn et al., 2011): the validity and clinical relevance of this approach in stroke patients has not been established.

2.3.2.6. The Montreal Cognitive Assessment (MoCA)

When choosing a scale to evaluate cognitive function in DARS participants there were several considerations. The scale had to be quick to administer, and usable with only brief training. It also had to cover cognitive domains that are commonly affected by stroke. As noted above, traditional diagnostic

criteria for “dementia” (The World Health Organisation, 2010, The American Psychiatric Association, 2012) are based upon impairment in memory (O'Brien et al., 2003, Moorhouse and Rockwood, 2008): and yet stroke survivors often display more subtle impairments including in executive function, praxis, and visuospatial abilities (Barker-Collo et al., 2010). The MMSE (Folstein et al., 1975) is amongst the most widely used and cited brief cognitive assessments in dementia screening (Ismail et al., 2010), but has been heavily criticised for its lack of sensitivity to impairment in executive function, visual perception/construction, and abstract reasoning (Nys et al., 2005). Its performance is also dependent on the educational and cultural background of the patients to which it is administered: there is a substantial “ceiling effect” in people of high pre-morbid intelligence with the attendant possibility of false negatives (Pendlebury et al., 2010), whereas older patients, non-English speakers, and those with sensory impairment or a low prior educational attainment may be misidentified as impaired due to difficulty completing some of the items (Ismail et al., 2010).

The MoCA (full original version, in English) was chosen for evaluation of cognitive function in DARS participants. It was developed by Nasreddine et al. (2005) as a screening tool for minor cognitive impairment, and is freely available for non-commercial purposes at www.mocatest.org (©Dr Ziad Nasreddine, MD: used in DARS with permission). Although originally developed for dementia screening, the MoCA has been evaluated for use in stroke populations, and has been shown to detect cognitive dysfunction in patients who were not identified as impaired on the MMSE (Dong et al., 2010, Pendlebury et al., 2010). The assessment takes approximately ten minutes to administer (Nasreddine et al., 2005), and covers several domains including: executive function; phonemic fluency; abstract reasoning; visuospatial functioning; object naming; attention; concentration; working memory; and orientation (Nasreddine et al., 2005).

The maximum total score is 30, with a correction of one point added for anybody who has had 12 years or fewer of formal education (Nasreddine et al., 2005). The validity of calculating summary scores in this way has not been established in stroke patients, although Rasch analysis has demonstrated that

the MoCA is capable of providing interval-level measurement in elderly patients attending a memory clinic (Koski et al., 2009). Summed scores are then typically treated as dichotomous variables, with a cut-off point defined below which patients are deemed to have cognitive impairment. The cut-off chosen is typically a score of ≤ 26 ; but this was derived from a mixed group of patients attending a memory clinic, so may not be appropriate for those with cerebrovascular disease (Pendlebury et al., 2012). Using the accepted cut-off score of ≤ 26 , the MoCA has a sensitivity of 0.97 for minor cognitive impairment after stroke, but a specificity of only 0.19 (Godefroy et al., 2011). By comparison the MMSE has a lower sensitivity (0.86) but higher specificity (0.61) using a cut-off score of ≤ 27 (Godefroy et al., 2011). The cut-off scores for both measures may be adjusted to confer a different balance of sensitivity and specificity: for example, using a cut-off score of ≤ 20 the MoCA has a sensitivity of 0.72 and a specificity of 0.90; using a cut-off score of ≤ 24 , the MMSE has a sensitivity of 0.70, and a specificity of 0.94 (Godefroy et al., 2011).

The MoCA has been used in the a trial of very early mobilisation after stroke (A Very Early Rehabilitation Trial; AVERT) (Bernhardt et al., 2015): interim MoCA data from the first 294 patients randomised were reported in 2011 (Cumming et al., 2011). The distribution of scores in this sample were skewed towards the higher end of the range, with a mean of 21.1 (standard deviation 7.5) and a median of 23 (inter-quartile range 17-27) (Cumming et al., 2011). Nevertheless, the majority of this sample (65%) scored below the defined cut-off score for cognitive impairment (< 26) (Cumming et al., 2011). This is in line with previous estimates of the prevalence of cognitive impairment after stroke, but may also suggest that the accepted cut-off score of < 26 is too high for this sample (Cumming et al., 2011). The poorest performance was seen on tests of executive function (trail making: 50% answered correctly), visuospatial function (cube copying: 50% correct responses), and phonemic fluency (just over 40% generated eleven or more words), and delayed recall of five objects (30% recalled no words, 10% recalled one) (Cumming et al., 2011). Orientation questions were generally answered correctly, with 60% of patients correct on all six (Cumming et al., 2011).

2.3.3. Interpreting the output of rehabilitation outcome measures

The “output” of rehabilitation outcome measures is often quoted as a numerical value. In the case of dichotomous (“yes/ no”) responses such as the RMI (Collen et al., 1991), this is typically a simple sum of all questions endorsed as “yes” by the patient. Scales that utilise polytomous responses (for example the GHQ-12) typically derive a score by weighting responses in some way (Goldberg and Blackwell, 1970, Goldberg and Hillier, 1979). The MoCA derives a summary score by adding together sub-scores that cover several domains of cognitive function (Nasreddine et al., 2005). The summary scores thus derived are assumed to correlate with a patient’s true clinical impairment.

2.3.3.1. Interpreting “summary scores” derived from rehabilitation outcome measures

The legitimacy of deriving such “summary scores” must be questioned. The RMI, for example, defines a range of items ranging from turning over in bed (assumed to be the easiest item on the scale) to running 10m in 4 seconds (assumed to be the hardest item). These mark extremes of the metric, and represent a progression from “lesser” to “greater” mobility. Each item in between marks a greater degree of mobility. The probability that an individual will answer “yes” to a given item is assumed to depend upon the level to which the latent trait under consideration (in this case mobility) exists in them (da Rocha et al., 2013). The items on the scale therefore provide “locations” along the continuum, against which an individual’s performance can be measured (Hobart and Cano, 2009). However, the location of items along the continuum cannot be assumed to be spaced at equal intervals (Hobart and Cano, 2009). In other words, the actual magnitude of any clinical change observed for each one-point change in score may vary considerably depending upon where on the continuum that change occurs. For example, consider two patients whose scores on the RMI change by two points. It is illogical to assume that the functional implications of a change from a score of 5 (standing unsupported for ten seconds) to a score of 7 (ability to walk ten meters or more unaided) are the same as a two point gain from an initial score of 1 (able to turn over in bed) to 3 (can sit on the edge of the bed for ten seconds). It is equally illogical

to state that a patient with a summary score of 10 on the RMI is “twice as mobile” as a patient with a RMI score of 5.

2.3.3.2. Implications for regression modelling

These scales are thus *ordinal* measures, and cannot be assumed to provide interval-level measurement. Scores derived from such scales should not be used in mathematical operations: for example, the calculation of mean scores or change in score between two time points. This poses a dilemma when utilising such scales as outcome variables in a regression analysis. The principles of regression modelling will be described in detail in Part 4: but at present it may be noted that two broad categories of model may be fitted. Linear regression modelling makes efficient use of available data, but relies upon the use of a continuous, as opposed to an ordinal, outcome measure. By contrast, logistic regression modelling allows the prediction of a categorical outcome (either binary or multinomial). Unlike linear regression models, categorical predictors may be fitted to logistic regression models. Such categorical variables may be derived from ordinal scales, by pre-specifying “cut points” and assigning each participant to one of two or more groups (da Rocha et al., 2013). Treating predictor variables in this way dispenses with the need to prove that an ordinal scale can provide interval-level measurement. However, logistic regression models make a less efficient use of the available data than a linear regression models. Furthermore, if the treatment of ordinal scales as categorical variables is to yield meaningful results then the cut-off point(s) must be selected on logical and clinically-relevant grounds. For the RMI, dichotomising scores at <7 or ≥ 7 has a sound theoretical basis, since the ability to walk short distances indoors after a stroke may allow a patient to be sufficiently independent to be discharged home as opposed to a care facility. However, the reasons for choosing a particular cut-off score is, for some scales, not clear. For example, a Barthel Index score of $>95/100$ is commonly cited as signifying almost complete independence in activities of daily living (Quinn et al., 2011), and was thus used in two major thrombolysis trials (Anonymous, 1995, Hacke et al., 1998). However, the rationale for using such a high cut-off score is questionable, particularly since a lower score

would still have signified a reasonable degree of independence (Sulter et al., 1999).

The methods by which models were constructed in this Thesis will be described below. However, the aim of this work was to establish whether a combination of clinical and radiological variables might be helpful in predicting mobility after stroke. Therefore, before discussing how the models were constructed, consideration must first be given to the process for analysing CT scans.

Part 2.4 Analysis of neuroimaging

2.4.1. Analysis of imaging in the DARS trial

2.4.1.1. The need for centralised review of imaging

Plans for the analysis of scans evolved in the early stages of trial setup, when an initial proposal to review imaging at local recruiting centres was found to be impractical. A decision was therefore made to undertake centralised review of scans, but doing so required a substantial amendment to the protocol. This was approved after the trial had opened to recruitment. Imaging could therefore not be obtained for some patients who had been recruited using the initial version of the protocol, since their explicit consent for centralised collation of imaging had not been obtained. For those recruited onto the revised protocol, all available neuroimaging performed in the course of the index admission was despatched to CTRU. All scans had been acquired in the course of the patient's routine care for the index stroke event. The DARS protocol did not stipulate a requirement for additional imaging beyond that which was clinically indicated. Nor was the time of imaging in relation to the index stroke specified.

2.4.1.2. Obtaining scans from recruiting sites

Images were copied to Compact Disc (CD) or Universal Serial Bus (USB) memory stick in Digital Imaging and Communication in Medicine (DICOM) format (National Electrical Manufacturers Association, 2011) for despatch to CTRU. Since both CDs and USB memory sticks were accepted, they will be referred to using the generic term "data storage media". Electronic transfer of files by a secure data link was considered, but was deemed impractical in the time available due to the requirement to obtain individual permission and set-up from every participating recruitment centre. Every effort was made to redact scans of patient-identifiable information (name, NHS number, and treating clinician) prior to despatch. The responsibility for this rested with recruiting centres, and was conducted according to their local protocols. Where complete redaction was not possible, it was agreed with the ethics committee that encryption would be used to protect trial data in transit. Passwords for encrypted media were forwarded to the data management

team at CTRU by email, separately from the images themselves. The possibility that scans would be incompletely anonymised was explained to participants in the patient information sheet and as part of the consent process. Once anonymised, scans and data storage media were identified only by the participant's unique trial identification number, initials, and date of birth. It was the responsibility of the recruiting centres to ensure that the correct scans for the correct patient were sent: once anonymised, CTRU had no direct means of verifying that the images sent were for a particular participant. Data storage media were despatched to CTRU by recorded delivery. Receipt was logged by CTRU, who also confirmed passwords for encrypted media with the recruiting centre.

2.4.1.3. Training JP in CT brain scan interpretation

All scans were co-reported by the Clinical Research Fellow (JP) and one of two experienced consultant neuroradiologists (Dr Jeremy Macmullen-Price, JMP; and Dr Tufail Patankar, TP) using a standardised Case Report Form (CRF) (The Royal College of Radiologists, 2011). Having had no prior experience in interpreting neuroimaging, a specific training programme was devised to allow JP to undertake this role. The initial training consisted of attendance at two organised teaching courses in scan interpretation, The Acute Stroke Training and Assessment in Computerised Axial Tomography course (Emsley et al., 2013) was a two-day theoretical and practical seminar in axial CT interpretation designed to meet the needs of stroke physicians who administer intravenous thrombolysis. The course focused primarily on axial CT imaging, since this is the modality of choice for the evaluation of acute stroke (Emsley et al., 2013). Learning objectives included: recognition of early ischaemic change on CT; recognition of primary intracerebral haemorrhage, subarachnoid haemorrhage, and infarction with haemorrhagic transformation; how to determine the extent of an infarct using the ASPECTS score; and an understanding of advanced imaging techniques (CT perfusion) and imaging appearances of conditions that may mimic stroke (Emsley et al., 2013). The course is accredited by the Royal College of Physicians with ten Continuing Professional Development credits awarded for completion. In addition JP attended the Leeds Third Surgical Neuroradiology course. This was a two-day

course designed to meet the needs of surgical trainees, and covered the principles of neuroimaging techniques and sequences.

2.4.1.4. Procedures for centralised review of scans

Review of the scans was conducted in Leeds Teaching Hospitals NHS Trust (LTHT). Scans were reviewed in small batches, corresponding to the order in which they were received by CTRU, not the order of enrolment of participants in to the trial. An accountability log was established to monitor their despatch from CTRU, receipt by JP, and return to CTRU. Whilst at LTHT data storage media were kept in locked filing cabinets in a secure office, in accordance with the principles of the Data Protection Act (The Stationary Office, 1998). It was not possible to upload scan to LTHT's Picture Archiving and Communication System (Meyer-Ebrecht, 1994, Ratib et al., 1994), since the limited storage capacity of these servers was required to meet clinical demands. Scans were therefore reviewed directly from the storage media using open-access DICOM viewer software. The time to read each scan was not recorded; observers were free to take as long as they deemed necessary to review the images. Scans were read blinded to clinical information (including the Oxford Community Stroke Project classification, the laterality of symptoms, and reports issued by local radiologists) and to treatment allocation. Although this does not reflect the way in which scans would be interpreted in clinical practice, it has been shown that interpreting images with knowledge of the patient's symptoms does not improve detection rates for early ischaemic change (Wardlaw and Mielke, 2005). The published validations of the AISCT have so far utilised blinded interpretation (Wardlaw et al., 2007, Wardlaw et al., 2010). Although observers in validation studies of the AISCT did not appear to "over-call" signs of ischaemic change (Wardlaw et al., 2007), it is not clear whether rates of over-calling would be higher if clinical information were provided.

2.4.1.5. Management of unexpected radiological findings

It was anticipated that, in accordance with standard practice, all scans would have been subjected to routine reporting by local radiologists in order to guide the clinical management of the patient. It was therefore thought unlikely that review of scans by the DARS team would yield findings of which local

clinicians were not already aware. Since the reporting of scans for DARS could be delayed by as much as three years from the point of a patient's enrolment in to the trial, the imaging review provided for the trial could not be relied upon to deliver clinically-significant information in a timely manner. It was also important that the trial team did not seek to influence or guide the clinical management of participants in any way. Feedback on reports was therefore not provided routinely to recruiting centres. However, under limited circumstances, the research team had an ethical obligation to make the treating team aware of certain findings that were unexpected and of clinical significance. A list of findings that would be fed back to the recruiting centres was not specified *a priori*: it was left to the discretion of the consultant neuroradiologist (JM-P or TP) to decide if notification was necessary. It was agreed that information would be communicated to the local PI, with whom responsibility for the clinical management and safety of participants at each centre rested. It was then the responsibility of the PI to ensure that the findings were acted upon appropriately. In practice, this procedure was invoked only twice in the course of the trial: the first for a suspected base of skull meningioma, and the second for a possible aneurysm at the tip of the middle cerebral artery.

2.4.2. Coding scan findings

2.4.2.1. The CT Imaging Interpretation Case Report Form (CRF)

A flow diagram summarising the process of image review is provided in Appendix E. For reporting of ischaemic stroke the AISCT, developed by Wardlaw and Sellar (1994), was utilised. For ICH some additional elements not included in the AISCT were incorporated, due to their prognostic importance. These were configured as an image analysis CRF (Appendix F). One CRF was used for each sequence acquired.

2.4.2.2. Initial coding: image quality and presence of any visible abnormality

Review of each scan began by recording the patient identification number and date of birth, date and time of image acquisition, the modality used (plain CT or CT angiography), and then a subjective judgement of image quality (good,

moderate, or poor) (Wardlaw et al., 2010). The first judgement was whether there were any abnormal findings (both stroke and others). The presence and side of any acute ischaemic lesion was then documented. If more than one ischaemic lesion was present, clinical judgement was used to decide which was the more recent or (in the case of two infarcts of a similar age) the more clinically significant.

2.4.2.3. Coding of acute ischaemic change

The main lesion was coded in detail as below, and secondary lesions were classified as a “second (minor) acute ischaemic lesion” under question 12. Features of early ischaemia were classified as: loss of definition between the cortical grey matter and underlying white matter; loss of the outline of the basal ganglia; and frank hypodensity (Wardlaw et al., 2007, Wardlaw et al., 2010, Wardlaw et al., 2014a). Acute swelling was classified using the AISCT framework and reference diagrams (Wardlaw et al. (2010); Appendix B). MCA lesions were also classified as involving less than or more than a third of this territory (Kaste et al., 1995, Wardlaw et al., 2010). Using the reference diagrams developed for the AISCT, MCA lesions were then further classified as: small cortical; basal ganglia striatocapsular; lateral to ventricle striatocapsular; anterior cortical MCA territory; posterior cortical MCA territory; whole of cortical MCA territory; whole of cortical MCA territory with lateral part of basal ganglia; and whole MCA territory (Wardlaw et al., 2010). Lesions in the ACA and PCA territories were each defined as involving less than 50% of that territory, more than 50%, or complete (Wardlaw et al., 2010). Lacunar lesions were classified as involving: the internal capsule or lentiform; the internal border zone; the centrum semiovale; or the thalamus (Wardlaw et al., 2010). Infarcts involving the anterior and posterior border zones were noted (Wardlaw et al., 2010). Cerebellar lesions were classified as lacunar infarcts, or as involving <50% or >50% of the hemisphere (Wardlaw et al., 2010). Similarly, brain stem lesions were classified as lacunar or as involving less <50% or >50% of the brain stem (Wardlaw et al., 2010). The ASPECTS score (Barber et al., 2000) was recorded for all lesions involving the MCA territory (Wardlaw et al., 2010). The presence of arterial hyperattenuation (suggestive of acute thrombus) (Gacs et al., 1983) in the MCA main stem, the insular

branch of the MCA, the internal carotid artery, the ACA, the PCA, the basilar artery, and the vertebral arteries was recorded (Wardlaw et al., 2010).

2.4.2.4. Coding of ICH

The location of confluent haematomas was classified as: frontal, temporal, parietal, occipital, basal ganglia/ thalamus, internal capsule, brain stem, cerebellum (Bhattathiri et al., 2006). The extent of midline shift (in millimetres), and the presence or absence of intraventricular extension and hydrocephalus were recorded. Haematoma volume was calculated using the formula $\text{Volume} = \text{A} \times \text{B} \times \text{C} / 2$ (Kwak et al., 1983, Kothari et al., 1996). Haemorrhage was classified as: petechial haemorrhage; significant haemorrhagic transformation of an underlying infarct; parenchymal haematoma with no infarct visible; parenchymal haematoma clearly remote from infarct; subdural haematoma; subarachnoid haemorrhage; and extradural haemorrhage (Wardlaw et al., 2014a). For confluent haematomas, the maximum diameter of the lesion was recorded (<3cm, 3-5cm, 5-8cm, and >8cm) (Wardlaw et al., 2014a). If blood was present in more than one location (for example a primary parenchymal haematoma with rupture in to the sub-arachnoid space) then the presence of both was recorded, and clinical judgement utilised in determining which was the most clinically important lesion (Wardlaw et al., 2014a). The presence or absence of changes in the anterior and posterior white matter was noted using the van Swieten scale as grade of 0 (no white matter change), 1 (change restricted to the periventricular region) or 2 (change extending from the lateral ventricle to the cortex) (van Swieten et al., 1990).

2.4.2.5. Coding of other findings

The presence of a second recent infarct was documented (Wardlaw et al., 2014a). Old vascular lesions were classified as: old cortical infarct(s); old striatocapsular infarct(s); old borderzone infarct(s); old lacunar infarct(s); old brainstem/cerebellar infarct(s); and probable old haemorrhage (Wardlaw et al., 2014a). Finally non-stroke lesions were classified as: cerebral tumour; encephalitis; cerebral abscess; and demyelination (Wardlaw et al., 2014a). Brain atrophy was not scored quantitatively, merely recorded as present or absent.

Part 2.5 Statistical analysis procedures

2.5.1. An overview of regression modelling

2.5.1.1. What is regression modelling?

As discussed previously, the purpose of a prognostic model is to combine several predictor variables in order to generate an estimate of the risk of a particular outcome for an individual patient (Steyerberg et al., 2013). These predictions may be useful in informing patients and their families about the likely outcome of their illness, in guiding the treatment of individual patients, and in selecting patients for participation in research trials (Moons et al., 2009b). Two broad types of regression models exist: linear regression, or logistic regression (Stoltzfus, 2011). The choice of which to use will depend upon the outcome variable chosen (in particular, whether it is continuous, binary, or categorical) (Stoltzfus, 2011). In order to understand the rationale for choosing type of model to use, and how its output might be interpreted, it is helpful to summarise briefly the mathematics of regression models.

2.5.1.2. Modelling using continuous predictor and outcome variables: linear regression

Of the two classes of model, linear regression is perhaps the most readily understood, and will thus be discussed here as a prelude to an explanation of logistic regression. A linear regression model makes two assumptions: firstly, that the outcome is a continuous (as opposed to a categorical) variable; secondly, that the relationships between the outcome and predictor variables can be expressed graphically as a straight line (Stoltzfus, 2011). Plotting the relationship between a continuous outcome and a single predictor variable might thus yield a graph similar to Figure 2.4.

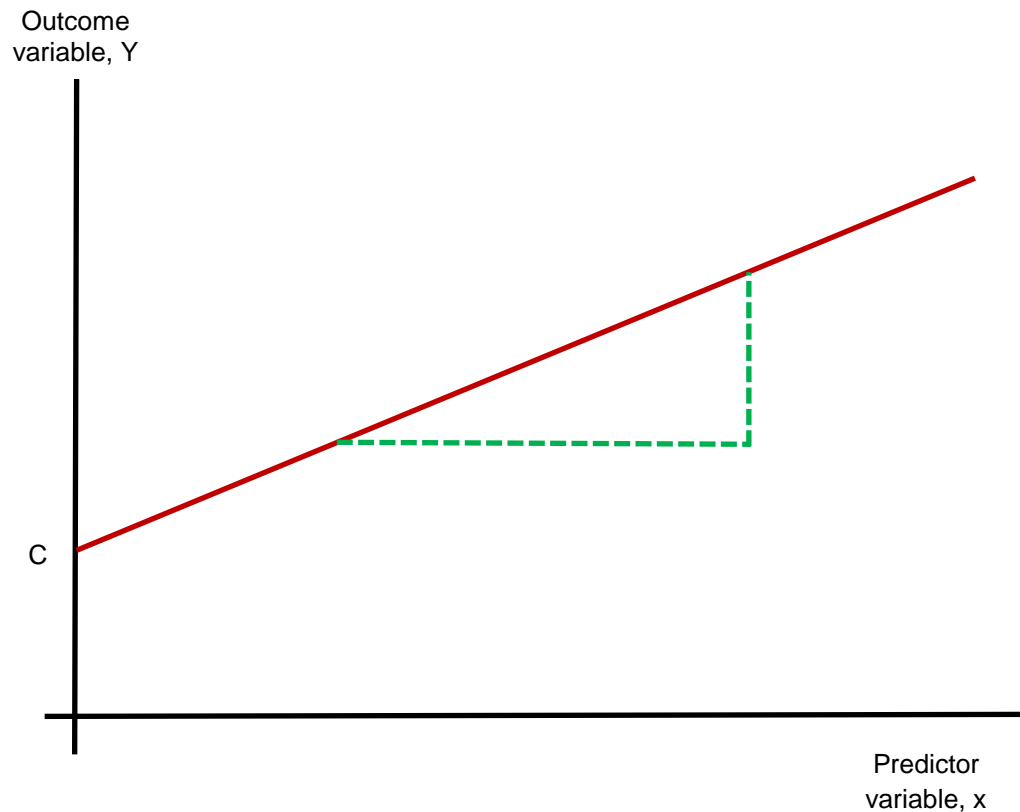


Figure 2.4. Linear relationship between a continuous predictor variable, x , and a continuous outcome variable, Y .

The relationship can be expressed by the equation $Y=mx+c$: in which m is the gradient of the line, and c is its Y -intercept.

The value of the outcome, Y , for any given value of the predictor, x , can be expressed by the equation:

$$Y=mx+c$$

In this case, c is the point at which the line intercepts the y -axis, and m is the gradient of the line (Field, 2013). In linear regression, the analogous term for the Y -axis intercept c is β_0 , also known as the constant for the equation (Field, 2013). This is, in effect, the predicted value of Y before any predictor variables are fitted (Stoltzfus, 2011). Where a single predictor, variable i , is fitted, the gradient of the resulting line may be termed β_i : also known as the coefficient for the variable i (Field, 2013). The values that variable i may assume are termed X_i . Hence, a univariate regression equation using the variable i to predict an outcome variable, \hat{Y} , may be expressed as (Field, 2013):

$$\hat{Y} = \beta_0 + \beta_i X_i$$

If several predictor variables are fitted to a model, then the resulting linear regression equation may be expressed as (Stoltzfus, 2011):

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i$$

In which β_0 is once again the constant (representing the y-axis intercept for a regression line with no predictor variables fitted) and x_i is the value (x) of the i th predictor variable weighted by its coefficient β_i (Stoltzfus, 2011).

Interpreting the output of simple linear regression is comparatively straightforward. In essence, the equation returns a predicted value, \hat{Y} , of the outcome variable, given specified values of each predictor variable X_i (Stoltzfus, 2011). It is possible to assess the contribution that each predictor, i , makes to the model in two ways: firstly, by determining the change in the outcome \hat{Y} for a one-unit change in the value of each variable i (assuming that values of all other predictor variables are held constant); secondly by examining the coefficient, β , for each variable (the greater the value of β_i , the greater the contribution made by the variable i to predicting \hat{Y}) (Stoltzfus, 2011).

2.5.1.3. Modelling using categorical predictor or outcome variables: logistic regression

When a *binary* outcome variable is used, it is obviously impossible to compute an absolute value of \hat{Y} (Stoltzfus, 2011). Instead, the prediction made by the model is of the *probability* (P) of belonging to a specified category (i) of an outcome event \hat{Y} (Stoltzfus, 2011). To facilitate this, the standard linear regression equation quoted above must be expressed on a logarithmic scale as $e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i}$. The probability (P) of belonging to category i of a binary outcome \hat{Y} may thus be expressed as (Stoltzfus, 2011):

$$P(\hat{Y}_i) = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i}}$$

In which e is the base of the natural logarithm; β_0 is a constant (the y-axis intercept for a regression line with no predictor variables fitted); and X_i is the

value of the i th predictor variable (X) weighted by its coefficient β_i (Stoltzfus, 2011).

Note that a binary outcome must have a probability that lies between 0 (the outcome never occurs) and 1 (the outcome always occurs). Since the continuous predictor variables in the above equation may take any value, there is a possibility that the equation may yield values of $P(\hat{Y}_i)$ that are <0 or >1 (Stoltzfus, 2011). This problem may be circumvented by expressing the output of a regression model as an *odds ratio*: the odds for membership of one outcome group (\hat{Y}) divided by the odds of belonging to the other outcome category ($1-\hat{Y}$) (Stoltzfus, 2011). This allows a variant of the standard linear regression equation to be used (Stoltzfus, 2011).

$$\text{Ln}(\hat{Y}/(1-\hat{Y})) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i$$

This output is expressed on a *logarithmic* scale, and is therefore a little more complex to interpret than standard linear regression (Stoltzfus, 2011). The term $\text{Ln}(\hat{Y}/(1-\hat{Y}))$ is essentially the natural log (the “logit”) of an odds ratio for membership of one group versus the other. The influence of each continuous predictor variable, i , on the model is thus expressed as the change in $\text{Ln}(\text{odds})$ of belonging to the specified category of outcome \hat{Y} for each one-unit change in the predictor variable (assuming that values of all other predictor variables are held constant) (Stoltzfus, 2011). This may be converted to a simple odds ratio by raising the base of the natural logarithm, e , to the power of the coefficient β of variable i (Stoltzfus, 2011):

$$\text{OR} = e^{\beta_i}$$

A positive value for the OR suggests that the odds of outcome \hat{Y} increase as the value of variable i increases; conversely, a negative OR implies a negative relationship between the odds of outcome \hat{Y} and variable i (Field, 2013). The statistical significance of the OR for each variable may be determined by examining 95% confidence intervals and p-values (Field, 2013). The odds of the outcome \hat{Y} following a one-unit change in a continuous variable i may be computed by multiplying the “baseline” odds of \hat{Y} by e^{β_i} (Stoltzfus, 2011).

Unlike in linear regression, which relies upon a linear relationship between continuous predictor and outcome variables, dichotomous or categorical predictors may be entered in to a logistic regression model. Where such variables are included, the impact of the variable is still expressed in terms of an odds ratio for a specified category of outcome \hat{Y} : but the interpretation of that odds ratio is complex. For a dichotomous predictor with two possible states (“A” and “B”), one state (for example, A) is nominated as the “basal” state, and the value of e^{β_i} quoted for state B is the *change* in odds ratio that results when a participant moves from state A to state B. Similarly, for a categorical predictor (which includes, for example, states A, B, C, and D) a “basal” state is defined to which the odds ratios for all other states are then referenced. For example, if the “basal” group is state A, then the values of e^{β_i} quoted for states B, C, and D will reflect the change in the odds ratio for the outcome \hat{Y} for participants in those states, relative to state A.

2.5.1.4. Assumptions of logistic regression modelling

Logistic regression makes no assumptions about the normality of the distribution of predictor variables (Bewick et al., 2005). There are, however, several key assumptions which must be tested to ensure the validity of any models derived from logistic regression. The first assumption is that sample group outcomes are uncorrelated, and that there are no duplicated measures amongst the sample (Stoltzfus, 2011). In the case of the DARS sample, this assumption was met since each individual case within the data-set is independent. The second assumption is that there exists a linear relationship between any continuous predictor variables and their natural-log transform (the “linearity of the logit”) (Stoltzfus, 2011). Thirdly, a high degree of correlation between two or more predictor variables (“collinearity”) is undesirable, since this may lead to large standard errors for values of β_i , (Stoltzfus, 2011). Finally, the model must be examined both for adequate fit in general, and also to ensure that there are no outlying cases which are disproportionately influencing the coefficients (Stoltzfus, 2011). The procedure for testing these assumptions will be discussed in detail later.

2.5.2. Modelling walking ability at T1 and T2 in the DARS data-set

2.5.2.1. Summary of the models

Since Levodopa is not effective in promoting recovery of walking ability after a stroke, treatment allocation was disregarded for the purposes of this analysis. The DARS cohort was treated as a large observational data-set. Using binary logistic regression, a series of six models was developed (Table 2.1) to predict ability to walk 10m or more independently at T₁ and at T₂. A full definition of this outcome variable is given below in Section 2.5.2.2.

Table 2.1. Summary of models presented for the “primary infarction, with scan available” (IWS), “primary intracerebral haemorrhage, with scan available” (HWS) and “whole DARS sample” groups.

Model	Analysis population	Outcome measured at	Candidate predictors
1	Primary infarction, with scan available (IWS)	T ₁	Demographic variables; Clinical impairment at T ₀ ; Imaging findings
2	Primary infarction, with scan available (IWS)	T ₂	Demographic variables; Clinical impairment at T ₀ ; Clinical impairment at T ₁ ; Imaging findings
3	Primary intracerebral haemorrhage, with scan available (HWS)	T ₁	Demographic variables; clinical impairment at T ₀ ; Imaging findings
4	Primary intracerebral haemorrhage, with scan available (HWS)	T ₂	Demographic variables; clinical impairment at T ₀ ; clinical impairment at T ₁ ; imaging findings
5	Whole DARS sample	T ₁	Demographic variables; clinical impairment at T ₀
6	Whole DARS sample	T ₂	Demographic variables; clinical impairment at T ₀ ; clinical impairment at T ₁

Analysis was performed using IBM Statistical Package for the Social Sciences (SPSS Statistics), Version 23. Since imaging was not available for a proportion of cases, it was necessary to define two analysis sub-groups. The

analysis for models 1 and 2 considered a sub-group of DARS participants (n=438) who presented with a primary cerebral infarction (as defined by the recruiting centre) and for whom a first CT scan was available for analysis. This group will be referred to as the “infarct with scan” (IWS) group. The analysis for models 3 and 4 considered a sub-group of DARS participants (n=75) who presented with a primary intracerebral haemorrhage (as defined by the recruiting centre) and for whom a first CT scan was available for analysis. This group will be referred to as the “haemorrhage with scan” (HWS) group. Models 5 and 6 considered predictors of walking ability in the DARS sample as a whole (n=593). Since imaging was not available for every patient, only demographic and clinical predictors were considered for inclusion in these models.

2.5.2.2. Definition of primary outcome measure: dichotomised RMI

SR-RMI (Collen et al., 1991) scores were used as the primary outcome measure at T₁ and T₂. When used as an outcome measure, the RMI was dichotomised as “able to walk 10m or more independently (yes/no)”. This was defined as a score of 7 or more, and item 7 answered “yes”, per the following algorithm (Figure 2.5):

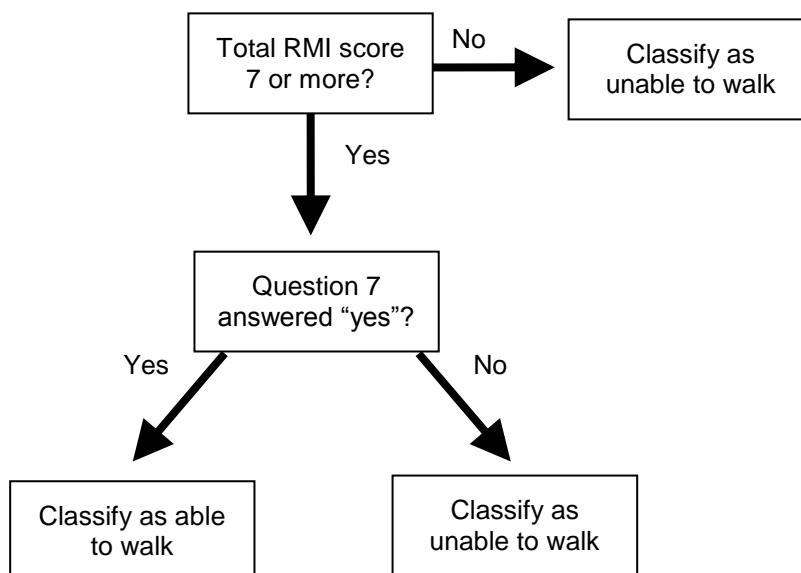


Figure 2.5. Algorithm for dichotomising RMI scores.

2.5.3. Treatment of predictor variables

For simplicity, the treatment of predictor variables will be discussed in terms of: demographic variables at T_0 ; clinical impairment at T_0 and at T_1 ; imaging predictor variables in ischaemic stroke; and imaging predictor variables in ICH.

2.5.3.1. Demographic variables

Age was entered in to the models as a continuous variable. Gender was dichotomised as male/female, and administration of thrombolysis (in the case of infarcts) as yes/no. The OCSF clinical stroke syndrome (Bamford et al., 1991) was used for infarcts only, and was entered as a categorical variable.

2.5.3.2. Clinical impairment at T_0 and T_1

The GHQ-12, FAS, and MoCA were analysed as continuous variables. When entering variables taken at T_0 into models to predict outcomes at T_1 and T_2 , the C-RMI was used as a predictor in preference to the SR-RMI. When variables at T_1 were used as predictors of outcome at T_2 , only the SR-RMI was available. When used as predictors (as opposed to as the outcome measure), both C-RMI and SR-RMI were treated as continuous variables. As discussed above the assumption that the RMI, GHQ-12, FAS, and MoCA provide interval-level measurement (and can thus be treated as continuous as opposed to ordinal scales) is not necessarily legitimate. However, ordinal scales are frequently analysed as interval-level measures: even in high impact-factor stroke and rehabilitation journals (Khan et al., Kozlowski et al., Lu et al., Takahashi et al.). At present, the limitations of treating these scales in this way will merely be acknowledged here. Consideration will be given in the concluding chapter to the principles of psychometrics, including a discussion of methods by which interval-level measurement may be derived from ordinal scales.

The MSK-SSP manikin (Hettiarachchi et al., 2011) was treated as a series of dichotomous variables. “Any MSK pain” was defined as pain in one or more body locus, irrespective of location. Upper-limb pain was defined as pain in one or more upper-limb locus (the shoulder, elbow, wrist, or hand). Lower-

limb pain was defined as pain in one or more lower-limb locus, including hips, knees, ankles, or feet. Recognising the possible confounding effect of central (neuropathic) post-stroke pain, Hettiarachchi et al. (2011) defined this as pain reported in all loci on the side ipsilateral to the clinical stroke syndrome. In DARS, the laterality of stroke symptoms was not recorded, so this distinction could not be reliably made. All reported pain was therefore assumed to be of musculoskeletal origin, whilst acknowledging the limitations of this assumption.

2.5.3.3. Imaging variables in ischaemic stroke

For the present analyses, only the first available plain CT scan performed after stroke was analysed. Wardlaw et al (The IST collaborative group, 2015) used the AISCT template to classify infarcts as small, medium, large, or very large. The same classification was followed in the present analysis. However, only two patients fulfilled the criteria for a “very large” infarct. The categories of “large” and “very large” were therefore combined under the heading of “large infarct”. A separate category of “no visible infarct” (not originally included by Wardlaw et al) was also added. The definition of these categories is summarised in Table 2.2.

Table 2.2. Classification of infarct size

Based on the AISCT template of Wardlaw et al. (The IST collaborative group, 2015). The category “no visible infarct” has also been added.

Classification	Scan findings
No visible infarct	No visible acute ischaemic change
Small infarct	Lacunar infarct; small cortical infarct; small cerebellar infarct; infarct involving less than half of brainstem, ACA territory, or PCA territory
Medium infarct	Striatocapsular infarct; infarct involving anterior or posterior half of peripheral MCA territory; infarct involving more than half of ACA or PCA territory;
Large infarct	Infarct involving: whole of peripheral MCA territory; whole of MCA territory; all of the MCA and ACA territory; all of MCA, ACA, and PCA territories.

Since the basal ganglia and other subcortical structures are thought to play a crucial role in motor learning (Penhune and Steele, 2012, Hikosaka et al., 2002, Doyon et al., 2009), a separate variable was also created which classified ischaemic stroke as: no visible infarct; “cortical” (infarct involving the cortex only); “subcortical” (infarct involving only the basal ganglia, cerebellum, or brain stem); or “both” (infarct affecting both cortical and subcortical structures).

Dichotomised variables were also created for the presence or absence of: any visible abnormality (infarct or other abnormality); a visible acute infarct; a visible acute infarct in the MCA territory; a visible acute infarct in the ACA territory; a visible acute infarct in the PCA territory; a visible acute lacunar infarct; a visible acute borderzone infarct; a visible acute cerebellar infarct; a visible acute brainstem infarct; a visible old vascular lesion (infarct or haemorrhage); any white matter lesions; and any atrophy.

2.5.3.4. Imaging variables in intracerebral haemorrhage

Haematoma volume (in mm³) and midline shift (in mm) were entered as continuous variables. Haematoma location was entered as a categorical variable (frontal, temporal, parietal, occipital, basal ganglia, internal capsule, brain stem, cerebellum). The presence or absence of intraventricular extension or hydrocephalus were treated as dichotomous variables. The presence or absence of a visible old vascular lesion (infarct or haemorrhage), any white matter lesions, and any atrophy were also entered as dichotomous variables.

2.5.3.5. Management of missing data

Data could be “missing” to varying extents, and for several reasons. Complete loss of data occurred when a patient died or withdrew from the trial, or when the research team were unable to contact patients to arrange follow-up. Those who remained in follow-up may have found it difficult to complete the questionnaires due to dominant hand weakness, visuospatial neglect, cognitive dysfunction, or fatigue. Short of complete loss of an entire data-set, the spectrum of missing data therefore ranged from all questionnaires attempted but with some missing items, to non-completion of whole outcome measures.

Up to three missing items on the SR-RMI were imputed, as follows. If items above and below the missing item were the same, then the missing item was assigned the same value as those items. If the first completed item after the missing item had a different value to the items below the missing item, then the missing item was assigned the value of the higher item. Dichotomising the RMI therefore depended upon having no more than three variables missing, and a valid response to question seven. The RMI was therefore classified as “missing” if a completed questionnaire was not received, or if a partially-completed questionnaire was received with more than 3 missing items or question 7 unanswered. If the T₁ SR-RMI was missing, then the patient was classified as “unable to walk” for the purposes intention-to-treat analysis. If the T₂ SR-RMI was missing, but a completed T₁ SR-RMI was available, then the assumption was made that no change in function had occurred in the interim

and the patient was classified on the basis of their response to the T₁ SR-RMI. If at T₂ both T₁ and T₂ SR-RMIs were missing, then the patient was classified as “unable to walk”.

2.5.4. Constructing the models: general considerations

2.5.4.1. Number of predictor variables fitted

When undertaking logistic regression modelling, it is important to understand how many variables it is reasonable to include in the model. Including a large number of variables results in a model that is over-fitted, and not generalisable beyond the data-set in which it was derived (Stoltzfus, 2011). Including variables for which there is an insufficient number of “observed” events may result in inflated values of β_i , with large standard errors (Stoltzfus, 2011). The challenge of model fitting is therefore to construct a parsimonious model, which provides a reasonable explanation of observed data whilst avoiding over-fitting (Stoltzfus, 2011). As a guide, binary outcomes require a minimum of ten outcome events (outcome occurred/did not occur) for each predictor variable included in the model (Stoltzfus, 2011). For example, a model to predict survival derived from a study in which 50 patients lived and 30 died should include at most 3 predictor variables (Stoltzfus, 2011).

2.5.4.2. Entry of predictor variables in to the model

Perhaps the most fundamental decision when fitting a regression model is: how should the variables be entered, and in what order? Three main methods have been proposed for fitting predictor variables (Stoltzfus, 2011). If there is no theoretical basis to support the inclusion of any particular variables, then all may be assumed to have equal importance and are entered in to the model simultaneously. This is the so-called “direct” approach (Stoltzfus, 2011). If previous work has suggested that particular predictors are likely to be important, then these may be entered first and additional variables entered sequentially thereafter. After each variable is entered, the model is re-evaluated to see if the new variable has improved its performance (Stoltzfus, 2011). The final method, stepwise regression, uses pre-defined statistical parameters to determine both the order in which variables are entered in to the model and also whether each variable is subsequently retained or

removed (Stoltzfus, 2011). Two variants of stepwise regression are recognised: the *forward* approach, in which variables are entered one at a time; and the backwards approach, in which all variables are first forced in to the model simultaneously, with subsequent stepwise removal of non-significant variables until only those that make a significant contribution to the model remain (Stoltzfus, 2011). SPSS allows stepwise regression to be automated, with variables included or excluded by the software algorithm on the basis of their statistical significance. However, this approach removes any clinical judgement about what variables are included in the final model (Stoltzfus, 2011). Variables that are of clinical importance but which make a marginal or non-significant contribution to the model may be excluded; or variables that are strongly statistically significant but have little clinically-plausible relationship to the outcome variable may be included. This may result in a model that is over-fitted to the sample in which it was derived (Stoltzfus, 2011). Although some have criticised stepwise approaches to model building on this basis (Field, 2013), others have argued that it is not the stepwise approach per se that is problematic but rather its thoughtless application (Stoltzfus, 2011).

2.5.4.3. Collinearity between predictor variables

Relationships between two or more variables may also profoundly affect the validity of the final model. Two or more predictor variables that are highly correlated are said to be “collinear”. Including collinear variables in a model may bias the coefficients for the variables concerned, or cause them to display a “direction” of effect paradoxical to that which might be expected (O'Brien, 2007). The possibility of collinearity between pairs of predictor variables was therefore explored prior to any formal modelling. A correlation matrix of Spearman's Rho (r) values was constructed, together with associated significance levels. Spearman's r is a non-parametric test of the correlation between two variables, and as such does not make an assumption of normality in the variable pairs it tests (Field, 2013). Interpretations of correlation coefficients vary, but as a guideline values of r of 0.10-0.29 indicate a small correlation, 0.30-0.49 a moderate correlation, and 0.50-1.00 a substantial correlation (Pallant, 2010). A second means of checking for the

presence of collinearity between predictor variables is to calculate tolerance values and Variance Inflation Factors (VIF) for each predictor variable entered (Field, 2013). Tolerance may be defined as “the proportion of variance of the i th independent variable that is not related to the other independent variables in the model” (O'Brien, 2007). The VIF is the reciprocal of the tolerance value (O'Brien, 2007). It has been suggested that tolerance values of <0.1 and VIF values of >10 indicate possible collinearity (Field, 2013), although some question the wisdom of rigorously applying such arbitrary “rules of thumb” (O'Brien, 2007).

Establishing that two variables show collinearity is, perhaps, only the first step: what is more crucial is how this is managed or mitigated. Although it might be tempting to remove one of the pair, there may be equally valid theoretical and statistical grounds to support the inclusion of either (Field, 2013). It is sometimes possible to include a pair of collinear variables in a model, provided that the estimates of β_i are plausible with reasonable standard errors and narrow confidence intervals (O'Brien, 2007). Finally, addressing collinearity does not obviate the need to examine other factors that may influence model stability, such as sample size (O'Brien, 2007).

2.5.5. Procedure for constructing the models

2.5.5.1. Fitting of predictor variables

For each of the six models, binary logistic regression was used to construct a series of univariate models exploring associations between patients' clinical characteristics and imaging variables (where available, in the case of models 1-4), and dichotomised ability to walk 10m or more independently at T_1 or T_2 . Multivariate models were then built using a forward stepwise approach, with manual as opposed to automated selection of variables. In the case of models predicting T_1 outcome, only demographic variables, imaging variables (where available), and T_0 predictors were used. Variables were entered in order of their statistical significance, with the most strongly significant univariate predictor entered first.

2.5.5.2. Assessing the contribution of each variable

After each step, the contribution of each new variable to the model was assessed, as was the contribution of each of the other variables already included. The significance of individual coefficients was calculated using the Wald χ^2 statistic (Bewick et al., 2005), in which:

$$\text{Wald } \chi^2 = (\beta_i / \text{Standard error of } \beta_i)$$

The Wald statistic was then compared to the χ^2 distribution using one degree of freedom, and significance values ascertained (Bewick et al., 2005). A threshold significance of $p < 0.05$ was taken to provide evidence that the variable made a statistically-significant contribution to the model. The magnitude of standard errors for β_i were also considered, as were the 95% confidence intervals for the odds ratio of each predictor variable.

2.5.5.3. Estimating the percentage of variance explained by the model

The overall performance of each iteration of the model was also assessed, by determining the percentage of participants correctly classified by the model and by using Cox&Snell R^2 and Nagelkerke R^2 to provide estimates of the amount of previously unexplained variance in outcome accounted for by the model (Field, 2013). Variables with non-significant coefficients or which did not improve the classification frequency or percentage of explained variance were deleted, unless there was compelling clinical grounds to justify their inclusion.

2.5.5.4. Managing collinearity between predictor variables

Where collinearity was known to exist between two variables, a decision was taken regarding which to omit. In taking this decision, two models were fitted, each identical in all other predictors but containing only one of the collinear variables. The properties of each model were examined, including the percentage of cases correctly classified and Cox&Snell R^2 and Nagelkerke R^2 (as estimates of the variance in outcome explained by each model). Consideration was also given to the magnitude of the coefficients, their standard errors, and significance levels for each of the collinear variables (when fitted to separate models). However, model properties alone were not

the sole factor in determining which of the variables to fit to the final model. For example, stroke services are typically configured to provide the most intensive rehabilitation input in the first few weeks or months after an event. The ability to predict outcomes using variables that would be available at the time of a patient's entry in to a rehabilitation programme might therefore be more clinically useful than a model which utilises predictors recorded some weeks later: even if the latter model has a higher classification rate and explains a greater proportion of the variance in outcome. This was an important consideration where collinearity existed between serial measurements of the same variable.

2.5.5.5. Managing missing data in categorical variables

When entering continuous variables, cases with missing data were included in the model. However, when dichotomised variables were entered cases with missing data were excluded. Recruiting centre was not included as a fixed effect in any of the models, since several centres recruited fewer than five participants. A significance level of 95% ($p=0.05$) was used as the threshold at which a variable was considered for entry in to a multivariate model.

Once a candidate model had been constructed, the overall performance of the model was assessed using the classification rate, estimates of explained variance (Cox&Snell R^2 and Nagelkerke R^2). Each model was then evaluated to ensure that the assumptions of linear regression (Stoltzfus, 2011) had been met, before a final version was settled upon.

2.5.6. Evaluating the models: assumption testing

2.5.6.1. The Box-Tidwell test for linearity of the logit

The assumption of linearity of the logit was assessed using the Box-Tidwell test (Field, 2013). This test depends upon forcing all continuous predictors into a binary logistic model together with their natural-log interaction term (predictor* \ln [Predictor]). Failure of the natural-log interaction terms to make a statistically significant contribution (at $p<0.05$) to this model is taken as evidence that the assumption of linearity had been met (Field, 2013). However, for continuous predictor variables in the DARS dataset (C-RMI, SR-RMI, MoCA, GHQ-12, and FAS), it was possible for participants to achieve a

value/score of 0. Since the natural log of 0 cannot be computed, a visual inspection of raw scores for each continuous predictor variable was performed. Where necessary, a single point was added to the scores of all participants, in order to eliminate values of 0 prior to natural-log transformation of the scores. When the Box-Tidwell test was performed, the interaction terms used for continuous predictor variables were therefore (predictor+1*Ln[predictor+1]). Since all participants were aged ≥ 18 at the time of randomisation, the variable “age” was natural-log transformed without prior adjustment.

2.5.6.2. Hosmer and Lemeshow goodness-of-fit test

Model fit was assessed using the Hosmer and Lemeshow chi-squared goodness of fit test. This test groups observations into deciles based on the predicted probability of the outcome occurring (Bewick et al., 2005). The χ^2 statistic is calculated as follows:

$$\chi^2 = \sum (\text{observed} - \text{expected})^2 / \text{expected}$$

Significance values for χ^2 are calculated using eight degrees of freedom (Bewick et al., 2005). The null hypothesis for this test is that the model under consideration is an adequate fit to the observed data; a p -value of >0.05 was therefore taken as evidence of adequate model fit (Bewick et al., 2005).

2.5.6.3. Testing for the presence of influential cases

As well as the overall fit of the model to the data, it is important to determine if a handful of outlying cases are exerting a disproportionate influence on the model. *Residual values* are a measure of the distance between the “observed” location of a case and its “predicted” location on the regression line (Stoltzfus, 2011). For ease of interpretation they are usually quoted as “standardised” values, with values $>\pm 2.58$ indicating cause for concern (Field, 2013). Case-wise listings of standardised residuals were derived for all six models, and examined for the presence of outlying cases. Where such cases were detected, an examination was made of the source data and where possible a reason for localised poor fit was determined.

Where outlying cases are detected, it is also important to understand how much *influence* they exert over the final model. A small number of outliers may influence the gradient of regression lines, and thereby bias estimates of coefficients. Several methods exist to detect so-called “influential cases”. Cook’s distance (Cook, 1979) is an estimate of the overall influence of a case on the model. It is calculated for each case in the data-set; cases with values >1 are usually deemed to warrant further examination (Field, 2013). The leverage value is an indication of the influence that each observed value of the outcome variable exerts over the predicted values of that variable (Field, 2013). The expected leverage value is calculated as:

$$\text{Leverage} = k+1/N$$

where k is the number of predictors and N is the sample size (Field, 2013). In the absence of any influential cases, all values would be close to this value. Cases with multiples of 2 or 3 times the “expected” value have been suggested as meriting evaluation (Field, 2013). The final method for detecting influential cases is to examine the difference between coefficients when one case is excluded from analysis, versus when all cases are included (Field, 2013). These values, known as DFBeta, are calculated for both the constant and for each predictor variable for each case. Values of >1 are interpreted as giving cause for concern (Field, 2013).

The dilemma is not in detecting outliers or influential cases, but in deciding how they should be managed. When deciding whether to include or exclude such cases, it is important to take in to account the magnitude of the observed change in model parameters as a result of including these cases. This is achieved by fitting a series of models: the first with all cases included, and subsequent iterations with each of the influential cases omitted in turn (Stoltzfus, 2011). If the effect of including a particular case on the overall model parameters is modest, then it is reasonable to retain that case in the analysis (Stoltzfus, 2011). Cases with a more dramatic influence may be omitted, but reasons for doing so must be fully justified.

2.5.7. Summarising the models

For each model, classification tables (Table 2.3.) were derived, and the sensitivity, specificity, and positive and negative predictive values of each model were ascertained:

Table 2.3 Example classification table, illustrating calculation of sensitivity, specificity, and positive and negative predictive values.

		Predicted		% correct
		Not walking	Walking	
Observed	Not walking	A	B	Y
	Walking	C	D	X

The sensitivity of a model refers to the percentage of participants with the target outcome (in this case, walking) who have been correctly classified by the model (Pallant, 2010): value X in the table above. Specificity, or the percentage of patients without the target outcome who are correctly identified, corresponds to value Y in the table above (Pallant, 2010). The positive predictive value is the percentage of patients that the model predicts will have the outcome of interest who do indeed have that outcome (Pallant, 2010). It is calculated as:

$$\text{Positive predictive value} = D/B+D$$

Conversely, the negative predictive value is the percentage of patients that the model predicts will not have the outcome of interest, who actually do not go on to display that outcome (Pallant, 2010). It is calculated as:

$$\text{Negative predictive value} = A/A+C$$

2.5.8. Testing assumptions made for missing outcome data

Where no SR-RMI score was returned at T₁ or T₂, assumptions were made for missing data as outlined above. The impact of these assumptions on model parameters was explored by re-fitting each model with alternative assumptions made (Table 2.4.).

Table 2.4. Summary of alternative assumptions for missing data that were tested.

Condition	Assumption	Applied to
1	Those who did not return a SR-RMI are excluded from analysis	Models 1-6
2	Those who did not return a SR-RMI are able to walk independently.	Models 1-6
3.	Those who did not return a SR-RMI are unable to walk independently.	Models 2, 4, and 6

Assumption 3 was, in effect, the default assumption made for missing data in models derived for T₁ outcome (models 1, 3, and 5). This assumption was therefore only tested for models predicting outcome at T₂ (models 2, 4, and 6). Coefficients, odds ratios, and 95% confidence intervals were calculated for each predictor variable. The percentage of patients correctly classified by the model under each condition was calculated as above. Cox&Snell R² and Nagelkerke R² values were determined, as a measure of the percentage of variance explained by each model. Sensitivity, specificity, and positive and negative predictive values were calculated as above.

Chapter 3. Results

Part 3.1 Summary of recruitment

3.1.1. Summary of recruitment, and sample characteristics

3.1.1.1. Total numbers of participants recruited

The DARS trial opened in July 2011, and recruitment continued until March 2014. During the recruitment phase, a total of 19,509 patients were screened, of whom 1,547 were eligible. Of these, 599 consented to participate, and 593 were randomised: 308 to co-careldopa, and 285 to placebo. The mean interval between stroke onset and randomisation to the trial was 17 days (range 3-59 days, SD 10.06). Recruitment to the trial is summarised in Figure 3.1.

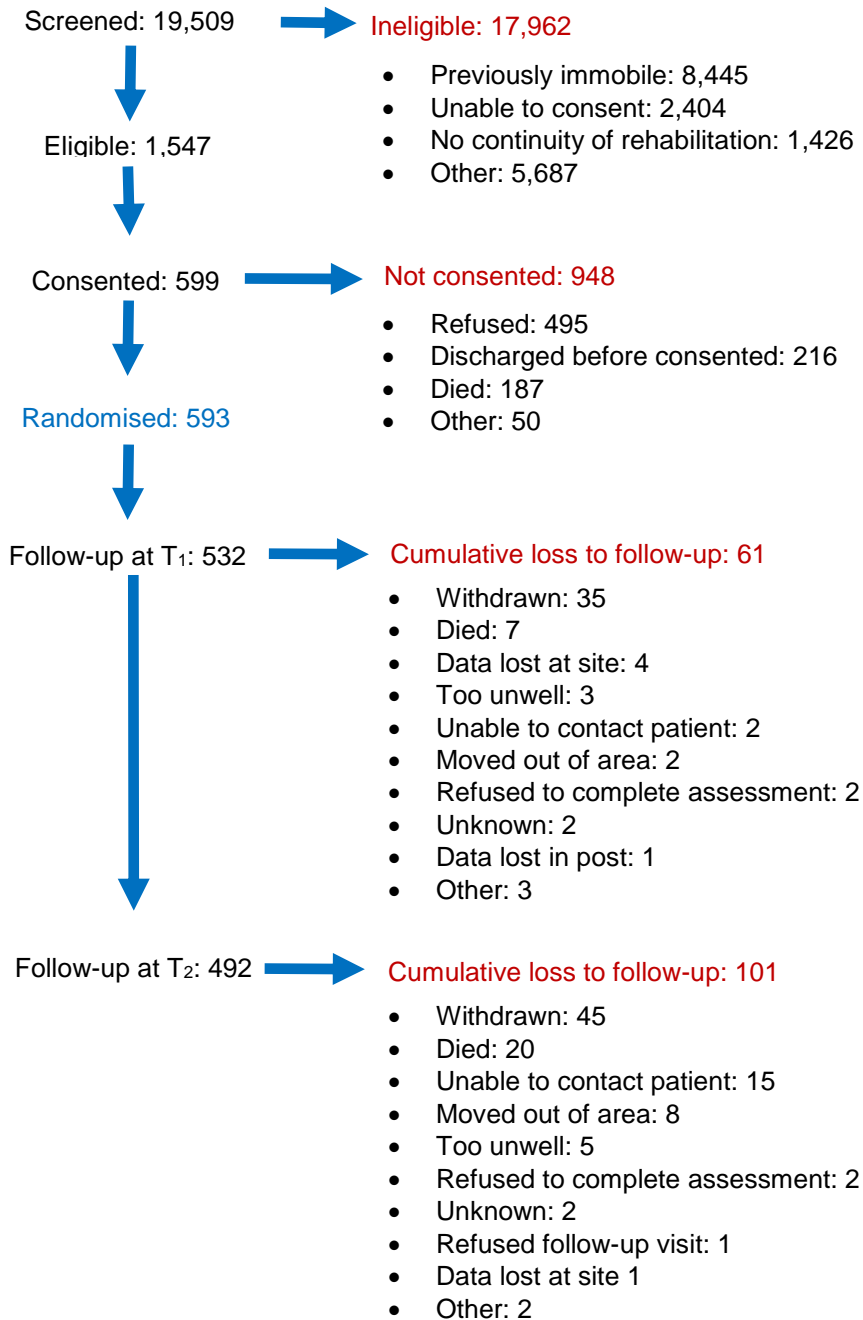


Figure 3.1. Summary of patients screened, eligible, recruited, randomised, and followed up in the DARS trial.

Cumulative loss to follow-up across both arms of the trial was higher than anticipated: 61 patients (10.3%) at T₁, and 101 patients (17.0%) at T₂. This may have resulted in decreased power to detect a difference between the intervention and control arms had the original sample size of 572 patients been retained. In view of this, an increase in the sample size to 593 participants was authorised.

3.1.1.2. Availability of SR-RMI scores at T₁ and T₂

In addition to patients for whom no data were returned, in some cases partially-completed SR-RMI scores were returned for which missing values could not be imputed. When these patients are taken in to account, the numbers of patients for whom SR-RMI scores were not available at T₁ and T₂ are slightly higher than those suggested in Figure 3.1: 69 patients at T₁, and 106 patients at T₂. When the characteristics of those who returned a SR-RMI score at T₁ were compared with those who did not, there was no evidence of a statistically significant difference in the age, the balance of males versus females, or the spectrum of clinical stroke syndromes (classified using the OCSP) between the two groups. There were no statistically significant differences in C-RMI, mRS, MoCA, and GHQ-12 scores at T₀, nor in the presence of musculoskeletal pain at T₀ between those who remained in follow-up at T₁ versus those lost to follow-up. Likewise, there was no statistical evidence of a difference in demographics or clinical impairment at T₀ between those who for whom SR-RMI scores were available at T₂ versus those who did not return a SR-RMI score at this time.

When impairment at T₁ was compared for patients who did and did not return SR-RMI scores at T₂, some important differences were found. Those who remained in follow-up at T₂ had slightly higher MoCA scores at T₁ than those who were lost to follow-up (22.83 for versus 21.06; $p=0.037$), indicating marginally better cognitive performance in this group. The group who remained in follow-up at T₂ also had lower mean GHQ-12 scores (16.32 versus 19.42; $p=0.002$), and lower mean FAS scores (24.49 versus 29.24; $p<0.0005$). This indicates that, in comparison to those who remained in follow-up, those patients who were lost to follow up at T₂ had slightly greater levels of cognitive impairment, depression, and fatigue at the last time-point for which data were available for them. There was no evidence of a statistically significant difference between the group who remained in follow-up at T₂ and those lost to follow up in SR-RMI scores, mRS, or the prevalence of musculoskeletal pain at T₁.

3.1.1.3. Availability of brain imaging

The number of patients for whom imaging was available for analysis is summarised in Figure 3.2.

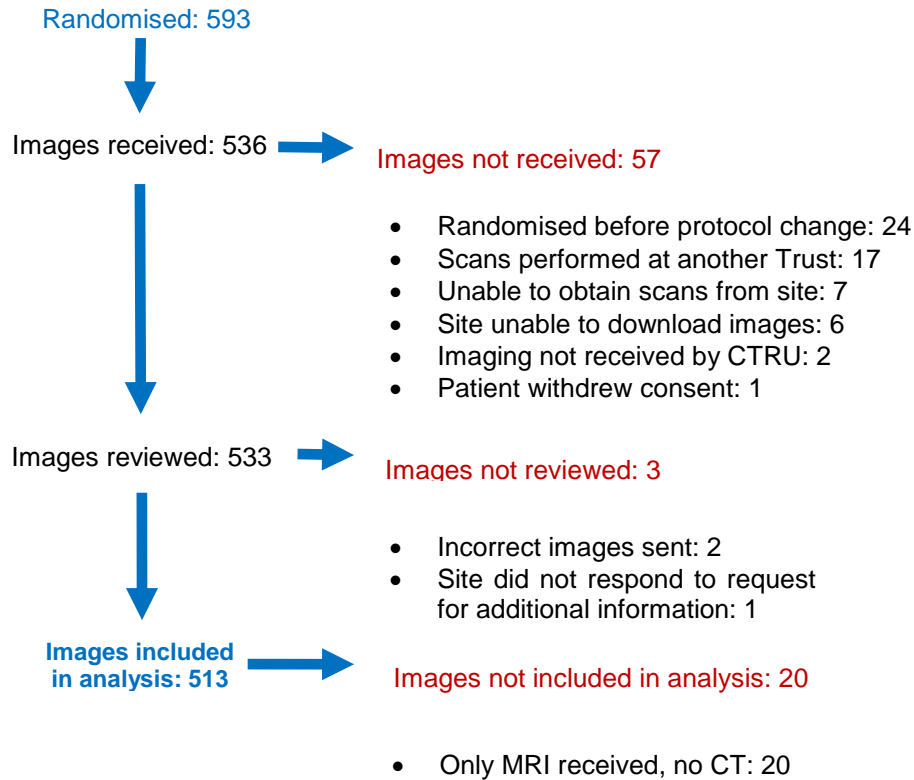


Figure 3.2. Summary of the availability of brain imaging for DARS participants. No imaging could be obtained for 57 of 593 (9.3%) participants. Of note, 24 patients were consented to the initial version of the protocol (which did not explicitly grant permission for despatch of scans to CTRU), and could not subsequently be contacted to obtain consent for review of their scans. Of the 533 patients for whom images were received and reviewed, 20 patients had only MRI images available. These patients were excluded from analysis. A total of 513 patients thus had a first CT scan available for analysis.

Comparing the demographic characteristics of those for whom imaging was available versus those for whom it was not revealed no evidence of a statistically significant difference between the groups in age, gender balance, proportion of infarcts versus haemorrhages, or OCSF clinical stroke

syndrome. Comparison of clinical impairment at T₀, T₁, and T₂ revealed that those for whom scans was obtained had slightly lower mean SR-RMI scores than those for whom imaging was not available at T₁ (mean score 6.71 versus 8.06; p=0.012) and T₂ (6.00 versus 9.43; p=0.016), but there was no statistically significant difference in the proportion of either group who were walking independently at each of these time-points. Nor was there any significant difference in cognitive function, depression, fatigue, or the prevalence of musculoskeletal pain at T₀, T₁, or T₂.

3.1.1.4. Clinical impairment of DARS participants at T₀

The demographics of the DARS sample and the clinical impairment of participants at T₀ are summarised below in Table 3.1.

Table 3.1. Demographic characteristics and clinical impairment in the DARS sample at T₀.

Mean age (range, SD)	68 (20-98, SD 13.23)
Gender	
Female/ Male (%)	229/ 364 (38.6%/ 61.4%)
Type of stroke	
Infarct/ ICH (%)	508/ 85 (85.7%/ 14.3%)
Clinical stroke syndrome (OCSP)*	
N (Missing, %)	508 (1, 0.2%)
TACS	161 (27.2%)
PACS	178 (30.0%)
LACS	116 (19.6%)
POCS	52 (8.8%)
Mobility (C-RMI)	
Mean (range, SD)	2.25 (0-6, 1.791)
Disability (mRS)	
0	4 (0.7%)
1	25 (4.2%)
2	53 (8.9%)
3	214 (36.1%)
4	172 (29.0%)
5	68 (11.5%)
Missing	57 (9.6%)
Cognition (MoCA)	
N (Missing, %)	580 (13, 2.2%)
Mean (range, SD)	20.23 (0-30, 6.308)
Depression (GHQ-12)	
N (Missing, %)	570 (23, 3.5%)
Mean (range, SD)	19.36 (3-36, 6.848)
Musculoskeletal pain	
Missing	10 (1.7%)
Any musculoskeletal pain ¹	236 (39.8%)
Upper limb pain ²	114 (19.2%)
Lower limb pain ³	154 (26.0%)

* Ischaemic stroke only. ¹Any musculoskeletal pain (upper limb, lower limb, spine). ² Any upper limb pain (shoulders, elbow, wrist, small joints of hand). ³ Any lower limb pain (hips, knees, ankles, feet)

3.1.2. The modelling process

3.1.2.1. Caveats for the models presented here

As described in Chapter 2, a series of six models were constructed to predict dichotomised walking ability in the IWS group, the HWS group, and the whole DARS sample at T₁ and T₂. Before describing the construction of these models, some important caveats must be acknowledged. Firstly, the findings presented below represent a post-hoc analysis on an existing data-set. As such, only predictor variables that were collected as part of the DARS trial will be included in these models. Secondly, the numbers of participants available for analysis (particularly in the HWS group) are small. Thirdly, no attempt will be made here to undertake formal validation of the models derived. All of the above limitations will be explored in detail in Chapter 4.

3.1.2.2. Testing for collinearity between predictor variables

Prior to model construction, the presence of collinearity between pairs of continuous predictor variables was assessed using Spearman's r (Field, 2013). A Spearman r of 1 indicate perfect correlation between two variables, whereas a Spearman r of 0 indicates that no correlation exists (Field, 2013). When interpreting Spearman's r , values of 0-0.39 are generally considered to be evidence of a weak association. For example, the correlation between age and FAS score at T₁ has a Spearman r of 0.009). Plotting the two variables on a scatter plot and fitting a line of best fit to this plot allows the absence of any relationship between the two to be visualised clearly (Figure 3.3.):

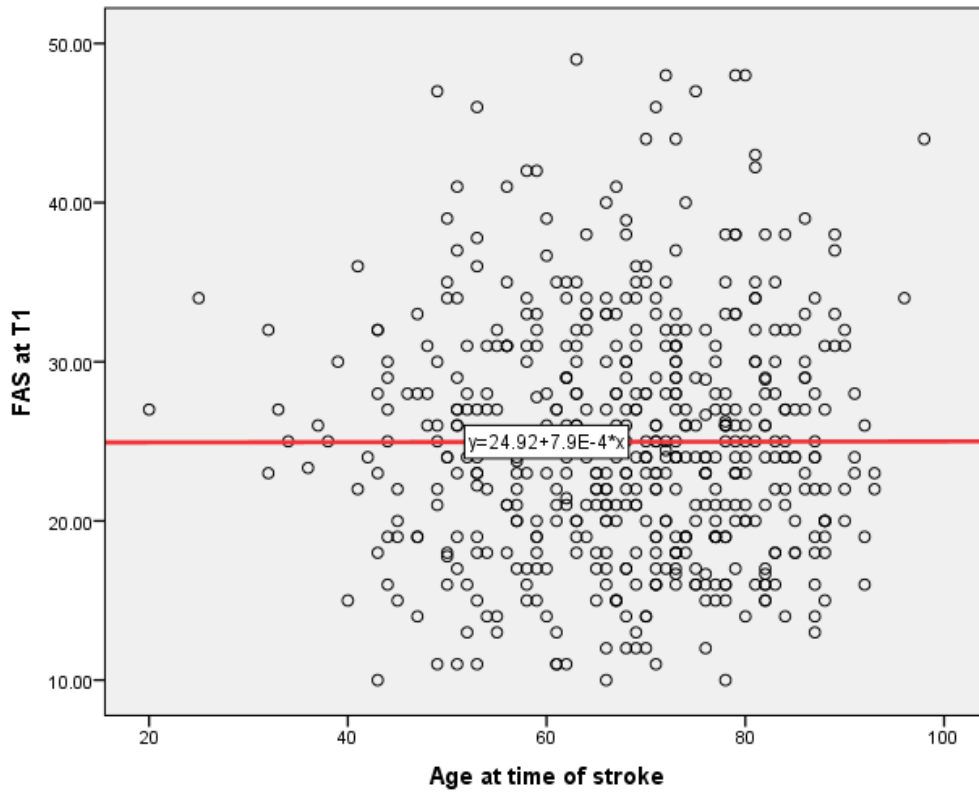


Figure 3.3. Example of a weak correlation between two predictor variables. Age, and FAS score at T₁.

Values of Spearman r of 0.4-0.59 are generally taken to imply an association of moderate strength. An example of such a relationship is that between SR-RMI at T₁ and MoCA at T₁ ($r=0.379$). Once again, this relationship may be visualised using a scatter plot (Figure 3.4.):

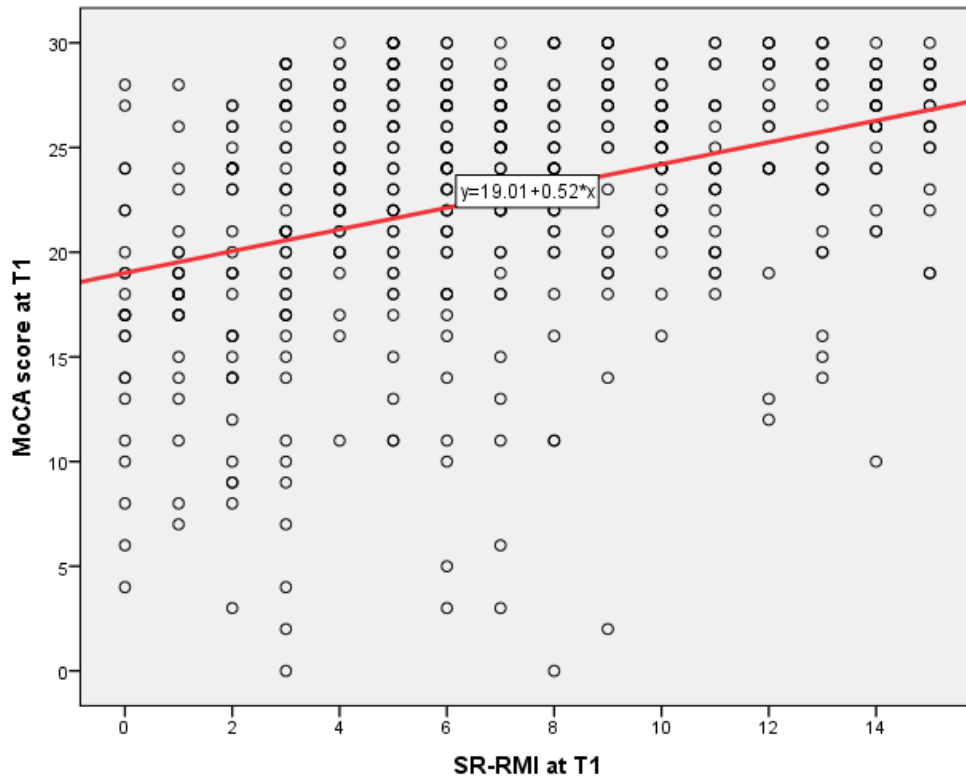


Figure 3.4. Example of a moderate correlation between two predictor variables.

SR-RMI and MoCA at T₁.

Strong associations are generally held to exist for values of r between 0.6-1.0. An example of such a relationship is that between MoCA scores at T₀ and T₁ ($r=0.750$), as demonstrated in Figure 3.5.:

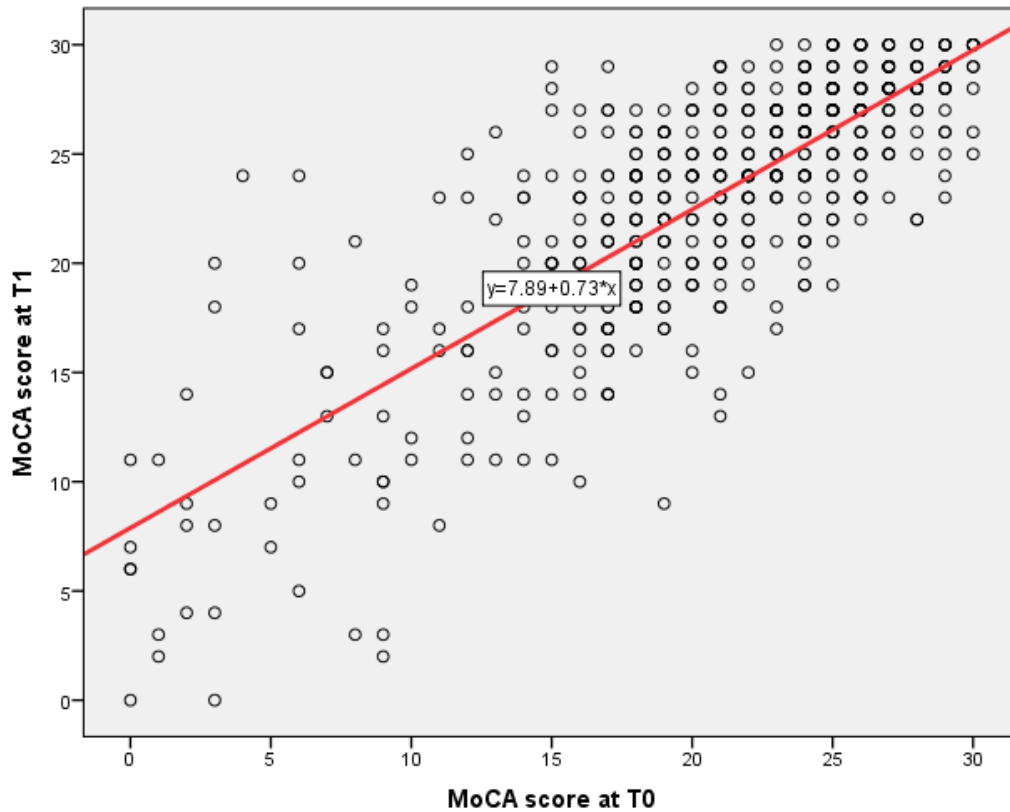


Figure 3.5. Example of a strong correlation between two predictor variables.
MoCA score at T_0 and MoCA score at T_1 .

The sign of Spearman's r provides some indication of the nature of the relationship between the variables: a positive sign implies a positive correlation: i.e. that as the value of one variable increases, so the value of the second variable increases. Conversely, a negative sign implies an inverse relationship: i.e, that as one variable increases, the second decreases. An example of this is the relationship between age and MoCA score at T_1 ($r = -0.351$), which is summarised in Figure 3.6.:

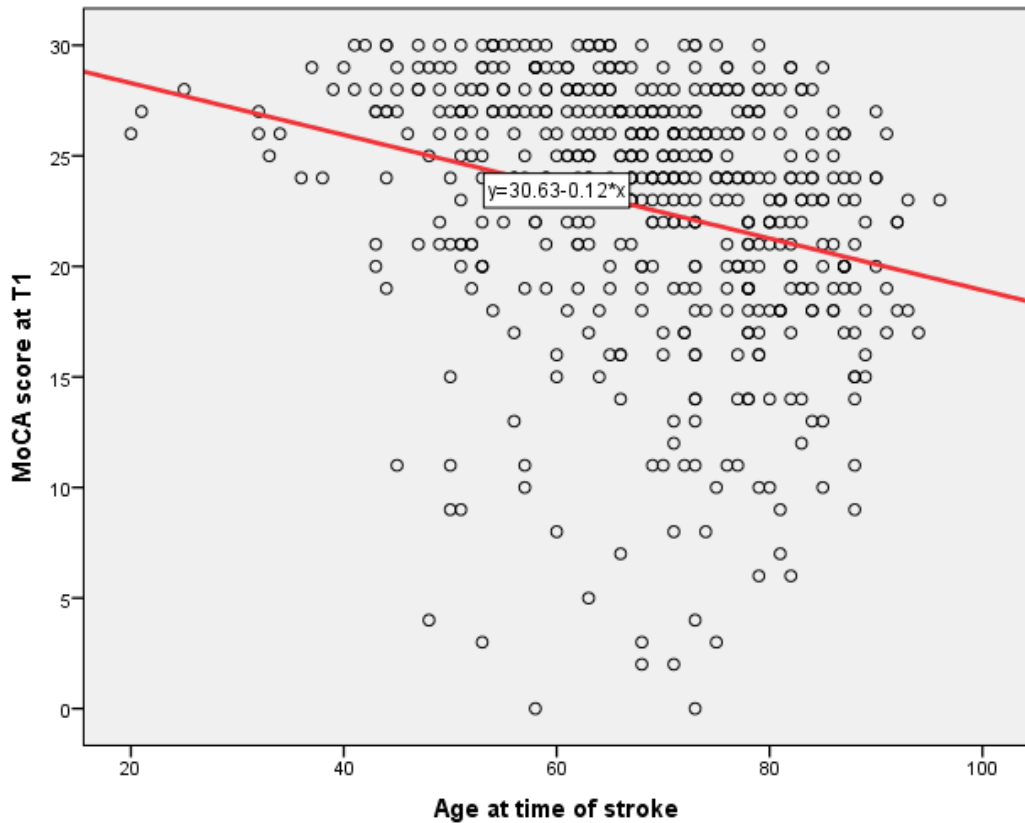


Figure 3.6. Example of a negative correlation between two predictor variables Age and MoCA score at T₁.

Although the Spearman's r provides an estimate of the strength of an association between two variables, it does not prove a *causal* link between them. Furthermore, the p -values cited for each value merely provide an estimate of the probability that the Spearman r value quoted has been obtained by chance: they provide no information about the strength of association between the two variables (Field, 2013). The Spearman's coefficient for each pair of predictor variables used in DARS, together with their associated significance levels, are summarised in Table 3.2

Table 3.2. Correlation matrix for predictor variables

		Age	Variables recorded at T ₀			Variables recorded at T ₁				
			C-RMI	GHQ-12	MoCA	SR-RMI	GHQ-12	MoCA	FAS	
Age	<i>r</i>	1.000	-0.186	-0.052	-0.274	-0.265	0.023	-0.351	-0.009	
	Sig	.	*	0.213	*	*	0.597	*	0.831	
	N	593	593	570	580	524	523	526	523	
Variables recorded at T ₀	C-RMI	<i>r</i>		1.000	-0.187	0.229	0.554	-0.080	0.193	-0.002
		Sig		.	*	*	*	0.066	*	0.958
		N		593	570	580	524	523	526	523
	GHQ-12	<i>r</i>			1.000	-0.104	-0.098	0.343	-0.138	0.251
		Sig			.	0.014	0.027	*	0.002	*
		N			570	562	505	503	508	504
MoCA	<i>r</i>				1.000	0.268	-0.064	0.750	-0.076	
	Sig				.	*	0.146	*	0.087	
	N				580	515	514	520	515	
Variables recorded at T ₁	SR-RMI	<i>r</i>				1.000	-0.293	0.379	-0.202	
		Sig				.	*	*	*	
		N					524	516	509	517
	GHQ-12	<i>r</i>					1.000	-0.165	0.613	
		Sig					.	*	*	
		N						523	509	518
MoCA	<i>r</i>						1.000	-0.143		
	Sig						.	0.001		
	N							526	510	
FAS	<i>r</i>							1.000		
	Sig							.		
	N								523	

* Value of *r* is significant at a level of $p < 0.0005$

The correlation matrix presented above shows strong collinearity for several pairs of predictors, including: C-RMI at T₀ and SR-RMI at T₁ ($r=0.554$); MoCA scores at T₀ and T₁ ($r=0.750$); and GHQ-12 and FAS at T₁ ($r=0.613$). Moderate correlations exist between: age and MoCA at T₁ ($r=-0.351$); GHQ-12 at T₀ and T₁ ($r=0.343$); and SR-RMI and MoCA at T₁ ($r=0.379$). All of these relationships were also strongly statistically significant ($p < 0.0005$). The variable pairs which display strong collinearity will not be fitted to the same model concurrently; the management of more modest degrees of collinearity is discussed in Chapter

2. The fitting of the models themselves will next be discussed, beginning with models 1 and 2 (derived in the IWS group).

Part 3.2 Modelling walking ability in IWS group

3.2.1. Characteristics of the IWS group

3.2.1.1. Comparison of IWS group with the whole DARS sample

Models 1 and 2 were derived in the IWS group: a sub-sample (n=438) of the whole DARS data-set (n=593). Prior to commencing model fitting, the characteristics of the IWS group were compared to those of the whole DARS sample. There was no evidence of a statistically-significant difference in the age profile or gender balance between the two groups. Nor was there evidence of a significant difference in any of the major clinical measures (dichotomous or mean RMI, mRS, MoCA, GHQ-12, FAS, and the prevalence of musculoskeletal pain) at T₀, T₁, and T₂.

3.2.1.2. Determining how many variables might be fitted

The proportion of patients in the IWS group who were able to walk independently at T₁ and T₂ is crucial for estimating the number of variables that might be fitted to the model. At T₁, 176 patients (40.2% of the IWS group) were able to walk independently for 10m or more; 262 (59.8%) were unable to do so. If a ratio of ten patients per observed outcome is applied to the smaller of the two outcome groups, then a maximum of 176/10=17 variables may be fitted to model 1. Allowing for a more stringent ratio of 20 patients per variable would allow up to eight variables to be fitted (Stoltzfus, 2011). Similarly, 221 patients (50.5%) were able to walk independently at T₂, and 217 (49.5%) were unable to do so. This implies that it is reasonable to fit between 10 to 20 variables to Model 2.

3.2.2. Model 1: return to walking at T₁ in IWS group

3.2.2.1. Univariate predictors of outcome at T₁: demographics, and clinical impairment at T₀

Model 1 examined predictors of walking ability at T₁. For this reason, only demographic variables, clinical impairment at T₀, and radiological findings were considered as candidate predictors in this model. The modelling process began by examining univariate associations between possible predictors and

walking ability at T₁. The most statistically significant univariate predictors were C-RMI score and MoCA score at T₀ ($p < 0.0005$) (Table 3.3). Age and OCSP clinical stroke syndrome ($p = 0.005$), thrombolysis status ($p = 0.002$), and gender ($p = 0.033$) were also statistically significant univariate predictors.

Table 3.3. Univariate predictors of independent walking ability at T₁.

Predictors: age, gender, thrombolysis status, OCSF clinical stroke syndrome, and clinical impairment at T₀.

	N	Missing (%)	Mean (Range, SD)	Sig	OR	95% CI for OR		
						Lower	Upper	
Demographic variables	Age	438	0	69.13 (20-98, 12.874)	0.005	0.978	0.964	0.993
	Gender	267	0	-	0.033	0.648	0.435	0.965
	Thrombolysis	109	1 (0.2%)	-	0.002	0.479	0.299	0.769
	OCSF	438	0	-	0.005	-	-	-
	TACS	145	-	-	0.005	-	-	-
	PACS	154	-	-	0.013	1.839	1.138	2.973
	POCS	43	-	-	0.004	2.820	1.403	5.671
	LACS	96	-	-	0.005	2.164	1.264	3.706
Physical impairment at T ₀	C-RMI	438	0	2.25 (0-6, 1.809)	<0.0005	1.719	1.512	1.954
	MoCA	429	9 (2.1%)	20.37 (0-30, 6.268)	<0.0005	1.076	1.039	1.114
	GHQ-12	422	16, (3.7%)	19.13 (3-36, 6.842)	0.475	0.990	0.962	1.018
	Any pain	180	7 (1.6%)	-	0.654	0.914	0.618	1.352
	UL pain	82	7 (1.6%)	-	0.786	1.070	0.656	1.745
	LL pain	120	7 (1.6%)	-	0.045	1.572	1.010	2.449

3.2.2.2. Univariate predictors of outcome at T₁: radiological variables

Univariate associations between walking ability at T₁ and imaging findings are shown in Table 3.4. None of the imaging predictor variables tested attained statistical significance.

Table 3.4. Univariate models of ability to walk 10m or more independently at T₁.

Predictors: vascular territory of infarct; presence of old stroke, white matter lesions, or atrophy; infarct location; infarct size.

	N (%)	Missing (%)	Sig	OR	95% CI for OR	
					Lower	Upper
Any visible abnormality	379 (86.5%)	0	0.840	0.944	0.538	1.655
Any acute infarct	225 (51.4%)	0	0.390	0.846	0.577	1.239
Acute MCA infarct	180 (41.1%)	0	0.510	0.877	0.549	1.295
Acute ACA infarct	8 (1.8%)	0	0.876	0.891	0.210	3.778
Acute PCA infarct	16 (3.7%)	0	0.824	0.889	0.317	2.493
Acute lacunar infarct	13 (3.0%)	2 (0.5%)	0.215	0.438	0.119	1.614
Acute borderzone infarct	5 (1.1%)	2 (0.5%)	0.999	0.000	0.000	-
Acute cerebellar infarct	4 (0.9%)	2 (0.5%)	0.072	3.048	0.903	10.28
Acute brainstem infarct	12 (2.7%)	2 (0.5%)	0.192	4.535	0.468	43.954
Old vascular lesion	119 (27.2%)	4 (0.9%)	0.615	1.116	0.727	1.713
Any white matter lesions	189 (43.2%)	3 (0.7%)	0.926	0.982	0.667	1.445
Atrophy	46 (10.5%)	9 (2.1%)	0.151	1.622	0.838	3.139
Infarct location	437	1 (0.2%)	0.820	-	-	-
Cortical	89 (20.3%)	-	0.514	0.845	0.509	1.402
Subcortical	90 (20.5%)	-	0.716	0.911	0.552	1.505
Both	45 (10.3%)	-	0.408	0.754	0.387	1.471
No visible infarct	213 (48.6%)	-	0.820	-	-	-
Infarct size	429	9 (2.1%)	0.951	-	-	-
Small	58 (13.2%)	-	0.640	0.867	0.477	1.576
Medium	131 (29.9%)	-	0.763	0.934	0.598	1.458
Large	32 (7.3%)	-	0.681	0.851	0.395	1.833
No visible infarct	208 (47.5%)	-	0.951	-	-	-

Of the 225 patients with a visible acute infarct, 180 had a lesion in the MCA territory. The observed frequency of infarcts in other vascular territories was low. In addition, the proportion of cases with missing data (n=2 for lacunar infarcts, borderzone, brainstem, and cerebellum) was high as a proportion of the observed frequency of infarcts in these territories. Estimates for odds ratios for these variables are therefore likely to be significantly biased, as evidenced by their wide confidence intervals.

3.2.2.3. Model 1: construction of an initial model

A multi-variate model was constructed using a forward stepwise approach. Based on the p -values outlined in Tables 3.3 and 3.4, C-RMI at T_0 was entered in to the model first. With the entry of each subsequent variables, the contribution of all variables to the model was re-evaluated and predictors that did not make a significant overall contribution were removed. The first variable to be entered was C-RMI. This made a statistically significant contribution to the model ($p < 0.0005$), explained between 17.5% and 23.7% of the variance in walking ability, and correctly classified 69.2% of cases as walking or not walking independently at T_1 . The addition of MoCA score at T_0 made a statistically significant contribution to the model ($p < 0.017$), but this combination explained only marginally more variance in outcome (19.6%-26.4%) and there was no change in the proportion of cases correctly classified by the model (69.2%). When added sequentially to the model, age, gender, OCSP and thrombolysis status failed to attain statistical significance, nor was there any substantial increase in the proportion of cases correctly classified. The proposed final iteration model therefore comprised C-RMI and MoCA score at T_0 .

3.2.2.4. Model 1: model evaluation and testing assumptions of logistic regression

When the Hosmer and Lemeshow goodness-of-fit test was applied to this model, a value of $p = 0.017$ was obtained, indicating inadequate fit to the observed data. Examination of standardised residuals for evidence of localised misfit revealed five participants with values $> \pm 2.58$. Each of these cases was evaluated in turn. Participant 164 had a total SR-RMI score of 13 at T_1 , which suggests that he was able to walk independently at this time. However, his dichotomised SR-RMI classified him as being unable to walk independently for 10m or more. This probably indicates that a response to question 7 was not provided: and since missing responses to this question were not imputed, he was classified as being unable to walk. Participants 173 and 312 did not return a valid SR-RMI at T_1 , and were therefore assumed to be unable to walk independently. This may account for the discrepancy between their observed (unable to walk) and predicted (able to walk) status.

Participants 238 and 563 had a very low C-RMI and MoCA scores at T₀, and yet were able to walk independently at T₁. It is possible that both simply made a better-than-expected recovery in the face of profound initial impairment. However, their persistently low MoCA scores at T₁ raise the possibility that cognitive impairment might have led them to over-estimate their physical abilities, and thus biased their SR-RMI score at T₁. It is equally possible that communication impairment compromised their ability to complete the MoCA at T₀, and thus led to a score that does not reflect their true cognitive ability. This could therefore have led the model to predict erroneously that they would be unable to walk at T₁.

3.2.2.5. Model 1: revision of model to improve fit

This evidence of localised misfit prompted a re-evaluation of the variables included in the model. On re-evaluating model parameters, the MoCA score at T₀ made a statistically-significant contribution to the model but its overall effect size was small ($\beta=0.044$). Removing this variable resulted in a univariate model which contained only C-RMI at T₀. The Hosmer and Lemeshow goodness-of-fit test was non-significant ($p=0.270$), indicating acceptable overall fit. When assessing local model fit, there were no cases in which standardised residuals exceeded the threshold level of ± 2.58 . Examining the influence of individual cases, there were no cases in which Cook's distance or values of DFBeta exceeded 1. Leverage values ranged from 0.00004 to 0.00962. Sixty-eight patients had leverage values in excess of the predicted value ($1+1/438=0.00457$). The Box-Tidwell test was performed to assess the assumption of a linear relationship between predictor variables and the natural log of the odds ratio of walking independently at T₁. Prior to natural-log transformation, a score of 1 was added to all values of C-RMI, to eliminate scores of 0. C-RMI score was then forced into a model which also contained the interaction terms (C-RMI+1*Ln[C-RMI+1]). The interaction terms did not make a significant contribution to the model, thus indicating that the assumption of linearity of the logit was met.

3.2.2.6. Model 1: final version

The final, univariate, iteration of Model 1 is summarised in Table 3.5.

Table 3.5. Final version of Model 1.

Model contains only C-RMI at T₀.

	B	S.E.	Wald	df	Sig.	OR	95% CI for OR	
							Lower	Upper
C-RMI at T ₀	0.542	0.065	68.608	1	<0.0005	1.719	1.512	1.954
Constant	-1.666	0.189	77.922	1	<0.0005	0.189		

3.2.2.7. Model 1: summary of model characteristics

This model correctly classified 69.2% of cases, and accounted for 17.5%-23.7% of variance in outcome. Each one-point increase in C-RMI at T₀ increased the odds of walking independently at T₁ by 71.9%. The observed versus predicted classification of patients by Model 1 is shown in Table 3.6. From this, the sensitivity, specificity, and positive and negative predictive values can be ascertained.

Table 3.6. Classification table for Model 1.

		Walking independently by T ₁ (predicted)		% correct
		No	Yes	
Walking independently by T ₁ (observed)	No	232	30	88.5
	Yes	105	71	40.3
Overall %				69.2

The sensitivity of Model 1 was low (40.3%), but specificity was higher (88.5%). The positive predictive value of the model is 70.3%. The negative predictive value of model 1 is 68.8%.

3.2.2.8. Model 1: testing assumptions made for missing data

Model 1 was constructed on the basis of an assumption that patients who did not return a SR-RMI score at T₁ were unable to walk independently. This might be termed the “default” assumption. To explore the impact of this assumption on model parameters, Model 1 was re-fitted with two alternative assumptions made. The first (assumption 1) was that patients who did not return a SR-RMI score at T₁ were excluded from analysis. A total of 393 patients were therefore analysed under this assumption. The second (Assumption 2) was that all patients who did not return a SR-RMI score at T₁ were able to walk

independently. The variation in the properties of Model 1 when re-fitted under these alternative assumptions are summarised in Table 3.7.

Table 3.7. Properties of Model 1 when fitted under alternative assumptions for missing SR-RMI scores at T₁.

	Default	Assumption 1	Assumption 2
% correctly classified	69.2%	71.2%	67.6%
Sensitivity	40.3%	62.5%	57.0%
Specificity	88.5%	78.3%	78.3%
Positive predictive value	70.3%	70.1%	72.8%
Negative predictive value	68.8%	72.0%	66.7%
% Variance Cox&Snell R ²	17.5%	21.2%	16.6%
% Variance Nagelkerke R ²	23.7%	28.3%	22.2%
C-RMI at T ₀			
B	0.542	0.623	0.531
OR	1.719	1.865	1.701
95% CI for OR	1.512 – 1.854	1.613 – 2.157	1.493 – 1.939

Model properties changed considerably depending upon what assumptions were made for missing data. The possible range for the percentage of cases correctly classified is between 67.6% (assumption 2) and 71.2% (assumption 1). The percentage increase in the odds of walking independently for each one-point change in C-RMI at T₀ ranged from 70.1% (under assumption 2) to 86.5% (under assumption 1). Values of sensitivity range from 40.3% (under the default assumption) to 62.5% (under assumption 1). Similarly, values of specificity range from 78.3% (assumptions 1 and 2) to 88.5% (default assumption). Positive predictive values ranged from 70.1% (assumption 1) to 72.8% (assumption 2); negative predictive values lay between 66.7% (assumption 1) and 72.0% (assumption 2). The percentage of variance explained by the model ranges from 16.6% (Cox&Snell R², assumption 2) to 28.3% (Nagelkerke R², assumption 1).

3.2.3. Model 2: return to walking at T₂ in IWS group

3.2.3.1. Univariate predictors of outcome at T₂: demographics, and clinical impairment at T₀ and T₁

Model 2 aimed to predict walking ability at T₂, and thus considered demographic variables, clinical impairment at both T₀ and T₁, and radiological findings as possible predictors. Once again, univariate associations between these variables and walking ability at T₂ were first established. The most strongly significant univariate predictors at T₀ were age ($p < 0.0005$), and C-RMI score and MoCA score at T₀ ($p < 0.0005$). SR-RMI and MoCA score at T₁ were also highly significant ($p < 0.0005$) (Table 3.8).

Table 3.8. Univariate predictors of independent walking ability at T₂.

Predictor variables: age, gender, receipt of thrombolysis, OCSP clinical stroke syndrome, impairment at T₀, and impairment at T₂.

	N	Missing (%)	Mean (Range, SD)	p	OR	95% CI for OR Lower Upper		
Demographic variables	Age	438	0	69.13 (20-98, 12.874)	<0.0005	0.962	0.947	0.978
	Gender	267	0	-	0.027	1.544	1.049	2.273
	Thrombolysis	109	1 (0.2%)	-	0.195	0.750	0.485	1.159
	OCSP	438	0	-	0.008	-	-	-
	TACS	145	-	-	0.008	-	-	-
	PACS	154	-	-	0.138	1.413	0.895	2.231
	POCS	43	-	-	0.001	3.557	1.691	7.482
	LACS	96	-	-	0.093	1.561	0.929	2.623
Physical impairment at T ₀	C-RMI	438	0	2.25 (0-6, 1.809)	<0.0005	1.506	1.335	1.698
	MoCA	429	9 (2.1%)	20.37 (0-30, 6.268)	<0.0005	1.064	1.030	1.099
	GHQ-12	422	16, (3.7%)	19.13 (3-36, 6.842)	0.857	0.997	0.970	1.026
	Any pain	180	7 (1.6%)	-	0.238	0.794	0.541	1.165
	UL pain	82	7 (1.6%)	-	0.907	0.972	0.601	1.572
	LL pain	120	7 (1.6%)	-	0.099	0.700	0.458	1.069
Physical impairment at T ₁	SR-RMI	393	45 (10.3%)	6.50 (0-15, 4.070)	<0.0005	1.599	1.457	1.755
	MoCA	392	46 (10.5%)	22.53 (0-30, 6.054)	<0.0005	1.119	1.076	1.163
	GHQ-12	393	45 (10.3%)	16.94 (0-36, 6.993)	0.004	0.958	0.930	0.986
	FAS	393	45 (10.3%)	25.24 (10-49, 7.874)	0.135	0.980	0.955	1.005
	Any pain	298	40 (9.1%)	-	0.185	0.732	0.462	1.160
	UL pain	251	40 (9.1%)	-	0.562	1.128	0.750	1.697
	LL pain	177	40 (9.1%)	-	0.042	0.662	0.444	0.985

3.2.3.2. Univariate predictors of outcome at T₂: radiological variables

Univariate associations between walking ability at T₂ and imaging findings are summarised in Table 3.9. None of the imaging variables was a statistically significant predictor of outcome, and estimates for odds ratios for some variables are likely to be biased by low observed event rates and relatively high proportions of missing data.

Table 3.9. Univariate models of independent walking ability at T₂.

Predictors: vascular territory of infarct; presence of old stroke, white matter lesions, or atrophy; infarct location; infarct size.

	N (%)	Missing (%)	Sig	OR	95% CI for OR	
					Lower	Upper
Any visible abnormality	379 (86.5%)	0	0.949	1.018	0.588	1.763
Any acute infarct	225 (51.4%)	0	0.928	1.107	0.699	1.480
Acute MCA infarct	180 (41.1%)	0	0.873	0.969	0.663	1.419
Acute ACA infarct	8 (1.8%)	0	0.167	0.321	0.064	1.609
Acute PCA infarct	16 (3.7%)	0	0.970	0.981	0.362	2.663
Acute lacunar infarct	13 (3.0%)	2 (0.5%)	0.753	0.837	0.277	2.532
Acute borderzone infarct	5 (1.1%)	2 (0.5%)	0.203	0.240	0.027	2.163
Acute cerebellar infarct	4 (0.9%)	2 (0.5%)	0.103	3.000	0.801	11.235
Acute brainstem infarct	12 (2.7%)	2 (0.5%)	0.347	2.972	0.307	28.800
Old vascular lesion	119 (27.2%)	4 (0.9%)	0.992	0.998	0.654	1.521
Any white matter lesions	189 (43.2%)	3 (0.7%)	0.077	0.709	0.485	1.038
Atrophy	46 (10.5%)	9 (2.1%)	0.005	0.392	0.203	0.758
Infarct location	437	1 (0.2%)	0.954	-	-	-
Cortical	89 (20.3%)	-	0.876	1.040	0.634	1.706
Subcortical	90 (20.5%)	-	0.948	1.016	0.621	1.664
Both	45 (10.3%)	-	0.623	0.851	0.447	1.620
No visible infarct	213 (48.6%)	-	0.954	-	-	-
Infarct size	429	9 (2.1%)	0.877	-	-	-
Small	58 (13.2%)	-	0.948	0.981	0.548	1.756
Medium	131 (29.9%)	-	0.798	1.059	0.684	1.640
Large	32 (7.3%)	-	0.479	0.763	0.361	1.614
No visible infarct	208 (47.5%)	-	0.877	-	-	-

3.2.3.3. Model 2: construction of an initial model

A multi-variate model was constructed using a forward stepwise approach. Demographic variables and clinical impairment at T_0 were entered first, in preference to impairment at T_1 . C-RMI at T_0 made a statistically significant contribution to the model ($p < 0.0005$). This variable alone explained 11.1%-14.8% of variance in walking ability, and correctly classified 63.5% of cases as walking or not walking independently at T_2 . The addition of age explained marginally more variance (14.7%-19.6%), and correctly classified 68.5% of cases. The addition of MoCA score at T_0 did not make a statistically significant contribution to a model containing C-RMI and age ($\chi^2 = 2.687$, $p = 0.101$), and the resulting model explained only fractionally more variance than a model containing C-RMI and age alone (15.9%-21.1%). Nor was there any substantial increase in the proportion of cases correctly classified (68.3%). However, the prevalence of cognitive impairment after stroke, and its theoretical importance in motor learning, justify retention of this variable in the model. Neither gender nor OCSF clinical stroke syndrome made a statistically significant contribution to the model; and nor did their inclusion increase the proportion of variance explained or the percentage of patients classified correctly. Both of these variables were removed from the model.

Having entered all predictor variables measured at T_0 , the model thus contained C-RMI and MoCA scores at T_0 and age. Predictor variables measured at T_1 were next entered. The SR-RMI score at T_1 made a highly significant contribution to the model ($p = < 0.0005$); in the presence of this variable the C-RMI score at T_0 became non-significant, indicating that the ability of mobility at T_0 to predict mobility outcomes at T_2 is wholly mediated by mobility at T_1 . C-RMI at T_0 was thus replaced in the model by SR-RMI at T_1 . The resulting model explained 38.3% to 51.2% of unexplained variance in outcome, and correctly classified 79.7% of cases. The MoCA score at T_1 also made a significant contribution to the model ($p = < 0.003$). Although MoCA scores at T_0 retained statistical significance when fitted alongside scores at T_1 , the strong collinearity between scores collected at the two time points raised the possibility that model coefficients could be biased if both were included. MoCA scores at T_0 were therefore removed from the model. GHQ-

12 scores at T₁ did not make a statistically significant contribution to the model, and were thus excluded.

3.2.3.4. Model 2: modification to improve localised misfit

The working “final” iteration of the model thus comprised age, and SR-RMI and MoCA scores at T₁. This model correctly classified 80.1% of cases, and explained between 38.8% and 51.9% of unexplained variance in outcome. However, when the model was examined for evidence of localised misfit, twelve patients had standardised residuals of $>\pm 2.58$. This amounted to 2.74% of the analysis population. Whilst the percentage of cases affected did not exceed the 5% threshold (Field, 2013), the fact that seven cases had standardised residuals of >3 was nevertheless cause for concern.

The variables included in the model were therefore re-evaluated, and an alternative model was constructed. Replacing SR-RMI at T₁ with C-RMI at T₀ resulted in more acceptable model fit, with seven cases having standardised residuals of $>\pm 2.58$. In this iteration, MoCA scores at T₁ remained highly significant, and this variable was therefore considered for inclusion in the final model. However, in clinical practice the timeframe defined as T₀ in DARS (5-42 days after stroke) typically marks the most intense period of rehabilitation intervention. It is desirable to be able to predict medium-term outcomes using information that would be available to professionals in the early acute rehabilitation period (i.e. within six weeks of a stroke). It could be argued that such a model would be more useful in clinical practice than a model which incorporated predictor variables acquired at a later date, even if the model explains a lower percentage of the overall variance in outcome when compared with a model containing variables measured at T₁.

3.2.3.5. Model 2: model evaluation and testing assumptions of logistic regression

Replacing MoCA scores at T₁ with values measured at T₀ thus resulted in a model containing age, and C-RMI and MoCA scores at T₀. This model showed acceptable fit to the data (Hosmer and Lemeshow goodness-of-fit test $p=0.488$), and only two cases had standardised residuals of $>\pm 2.58$. Neither patient returned a RMI score at T₂, and both were therefore assumed to be

unable to walk independently at this time point. This may account for the discrepancy between their observed (unable to walk) and predicted (able to walk) status. The model was also examined for influential cases using Cook's distance, leverage statistics, and DFBeta values for the constant and each of the predictors (Table 28). No cases had a Cook's distance of >1 . Leverage values ranged from 0.00291 to 0.0445. One hundred and fifty-seven patients had a leverage value greater than the expected value of 0.00913). No cases returned DFBeta values >1 for the constant or for any of the predictor variables. The assumption that a linear relationship exists between each predictor variable and the natural log of the odds ratio of walking independently at T_2 was tested using the Box-Tidwell test. Prior to natural-log transformation, a score of 1 was added to all values of C-RMI and MoCA at T_0 , to eliminate scores of 0. The covariates C-RMI, age, and MoCA score were then forced in to a model which also contained the interaction terms (C-RMI+1*Ln[C-RMI+1]), (age*Ln[age]), and (MoCA+1*Ln[MoCA+1]). The natural-log interaction terms did not make a significant contribution to this model ($p>0.05$), indicating that the assumption of linearity of the logit was met for C-RMI, age and total MoCA score at T_0 .

3.2.3.6. Model 2: final version

The final iteration of Model 2 is summarised in Table 3.10.

Table 3.10. Final version of Model 2

Predictor variables: C-RMI, age, and MoCA at T_0 .

	B	S.E.	Wald	df	Sig	OR	95% CI for OR	
							Lower	Upper
C-RMI at T_0	0.395	0.065	37.107	1	<0.0005	1.485	1.308	1.686
Age	-0.032	0.009	12.977	1	<0.0005	0.969	0.952	0.986
MoCA at T_0	0.029	0.018	2.659	1	0.103	1.029	0.994	1.065
Constant	0.775	0.759	1.041	1	0.307	2.170		

3.2.3.7. Model 2: summary of model characteristics

This model accounted for 15.9%-21.1% of the unexplained variance, and correctly classified 68.3% of cases. Each one-point increase C-RMI at T_0 increased the odds of walking independently at T_2 by 48.5%. Each one-year

increase in age decreased the odds of walking independently at T₂ by 3.2%. The smallest overall effect size was seen for MoCA scores at T₀, with each one-point increase increased the odds of walking independently by only 2.9% at T₂. The observed versus predicted classification of patients by Model 2 is shown in Table 3.11.

Table 3.11. Classification table for Model 2

		Walking independently by T2 (predicted)		% correct
		No	Yes	
Walking independently by T2 (observed)	No	149	63	70.3
	Yes	73	144	66.4
Overall %				68.3

The sensitivity of Model 2 is 66.4%, with a specificity of 70.3%. Its positive predictive value was 69.6%, and its negative predictive value was 67.1%.

3.2.3.8. Model 2: testing assumptions made for missing data

The default assumption made for missing data in Model 2 was that all patients maintained the level of mobility they had reached at T₁. Those who had not returned a SR-RMI score at T₁ were assumed to be unable to walk independently at T₂. To explore the impact of this assumption, Model 2 was re-fitted with three alternative assumptions. Assumption 1 excluded patients who did not return a SR-RMI score at T₂ from analysis: data from only 364 participants were fitted under this assumption. Assumption 2 was that all participants who did not return a SR-RMI at T₂ were able to walk independently; conversely, Assumption 3 was that these patients were unable to walk independently at this time-point. The properties of the model under each of these assumptions are summarised in Table 3.12.

Table 3.12. Properties of Model 2 when fitted under alternative assumptions for missing SR-RMI scores at T₂.

	Default	Assumption 1	Assumption 2	Assumption 3
% correctly classified	68.3%	68.3%	66.0%	66.7%
Sensitivity	66.4%	74.4%	85.3%	61.4%
Specificity	70.3%	60.0%	30.0%	71.6%
Positive predictive value	69.6%	72.0%	69.4%	66.8%
Negative predictive value	67.1%	62.9%	52.3%	66.5%
% Variance Cox&Snell R ²	15.9%	18.3%	11.5%	12.5%
% Variance Nagelkerke R ²	21.1%	24.6%	15.8%	16.7%
C-RMI at T ₀				
B	0.395	0.488	0.378	0.356
OR	1.485	1.628	1.460	1.428
95% CI for OR	1.308 – 1.686	1.396 – 1.899	1.273 – 1.674	1.264 – 1.613
Age				
B	-0.032	-0.036	-0.025	-0.021
OR	0.969	0.965	0.976	0.980
95% CI for OR	0.952 – 0.986	0.946 – 0.985	0.958 – 0.993	0.964 – 0.996
MoCA at T ₀				
B	0.029	0.032	0.017	0.028
OR	1.029	1.033	1.017	1.028
95% CI for OR	0.994 – 1.065	0.994 – 1.072	0.983-1.053	0.994 – 1.064

The percentage of cases correctly classified ranged from 66.0% (assumption 2) to 68.3% (default assumption, and assumption 1). The percentage increase in the odds of walking independently for each one-point change in C-RMI at T₀ ranged from 42.8% (assumption 3) to 62.8% (assumption 1). The percentage change in the odds of walking independently with each one-year increase in age ranged from 2.0% (assumption 3) to 3.6% (assumption 1). The percentage change in odds ratio for independent mobility with each one-point increase in MoCA score was between 1.7% (assumption 2) and 3.3% (assumption 1). The sensitivity of Model 2 ranged from 61.4% (assumption 3) to 62.5% (under assumption 1). Values of specificity range from 78.3% (assumptions 1 and 2) to 85.3% (assumption 2). Positive predictive values lay between 66.8% (assumption 3) and 72.0% (assumption 1); negative predictive values between 52.3% (assumption 2) and 66.5% (assumption 3).

The percentage of variance explained by the model ranges from 11.5% (Cox&Snell R^2 , assumption 2) to 24.6% (Nagelkerke R^2 , assumption 1).

Part 3.3 Modelling walking ability in intracerebral haemorrhage

3.3.1. Characteristics of the HWS group

3.3.1.1. Defining the HWS group

Models 3 and 4 were derived in the HWS group: a sub-group of 75 patients (47 men and 28 women) who had sustained a primary intracerebral haemorrhage. Of these patients, 58 (77.3%) had radiological evidence of a parenchymal haematoma with no infarct visible. However, it must be noted that the criterion for inclusion in the HWS group was that the patient had sustained a primary intracerebral haemorrhage *as defined by the recruiting centre*. For this reason, the group also contains eight patients who were thought by the scan review panel (JP and consultant neuroradiologist) to have a parenchymal haematoma clearly remote from a visible infarct, and two who were thought to have radiological evidence of haemorrhagic transformation of an underlying infarct.

3.3.1.2. Determining how many variables might be fitted

It must be acknowledged that the numbers of patients in the HWS group are small, and do not support anything other than an exploratory analysis. In particular, the number of observed outcome events at each time point limit the number of predictor variables which may be fitted. At T₁, 36 patients (48.0%) were able to walk independently for 10m or more; 39 (52.0%) were unable to do so. If a guideline of ten patients per variable is applied to the smaller of the two outcome groups, then a maximum of $36/10=3$ variables may be fitted to Model 3. By T₂, 32 patients (42.7%) were able to walk independently, with 43 (49.5%) unable to do so. Model 4 therefore supports a maximum of three predictor variables.

3.3.2. Model 3: return to walking at T₁ in HWS group

3.3.2.1. Univariate predictors of outcome at T₁

Model 3 examined predictors of mobility at T₁, using demographic details, clinical impairment at T₀, and radiological predictors. Univariate associations

between these variables and walking ability at T₁ are summarised in Table 3.13.

Table 3.13. Univariate predictors of independent walking ability at T₁.

Predictor variables: age; gender; clinical impairment at T₀; haematoma volume; presence of midline shift; haematoma location; presence of hydrocephalus; presence of intraventricular extension; white matter lesions; old stroke lesion.

	N	Missing (%)	Mean (Range, SD)	Sig	OR	95% CI for OR	
						Lower	Upper
Age	75	0	65.85 (32-92, 1.501)	0.030	0.960	0.925	0.996
Male gender	47	0	-	0.833	0.904	0.354	2.309
C-RMI	75	0	1.96 (0-6, 1.664)	0.001	1.934	1.312	2.849
MoCA at T ₀	72	3 (4%)	20.17 (0-30, 5.891)	0.035	1.102	1.007	1.205
GHQ-12 at T ₀	71	4 (5.3%)	20.84 (8-36, 6.836)	0.442	0.973	0.908	1.043
Any pain at T ₀ (yes)	21	3 (4.0%)	-	0.914	1.058	0.382	2.925
UL pain at T ₀ (yes)	12	3 (4.0%)	-	0.463	0.625	0.178	2.192
LL pain at T ₀ (yes)	14	3 (4.0%)	-	0.908	0.933	0.290	2.999
ICH volume (mm ³)	67	8 (10.7%)	16.58 (1-88, 17.023)	0.091	0.973	0.942	1.004
Midline shift (mm)	66	9 (12%)	1.99 (0-44, 5.700)	0.635	1.025	0.926	1.134
Haematoma location	64	11 (14.7%)	-				
Frontal	10	-	-	0.711			
Temporal	1	-	-	1.000	0.000	0.000	-
Parietal	6	-	-	0.058	11.667	0.922	147.5
Occipital	1	-	-	1.000	>1000	0.000	-
Basal ganglia	38	-	-	0.212	2.593	0.581	11.56
Internal capsule	2	-	-	0.590	2.333	0.107	50.98
Brainstem	0	-	-	-	-	-	-
Cerebellum	6	-	-	0.999	>1000	0.000	-
Hydrocephalus	12	9 (12%)	-	0.272	0.941	0.138	1.748
Intraventricular extension	26	8 (10.7%)	-	0.378	0.640	0.237	1.726
Old vascular lesion	13	2 (2.7%)	-	0.335	0.547	0.160	1.866
White matter lesions	26	0	-	0.090	0.431	0.163	1.142
Atrophy	5	2 (2.7%)	-	0.580	0.593	0.093	3.774

The only statistically significant univariate predictors of mobility at T₁ were age, and C-RMI and MoCA scores at T₀. However, the HWS group contained only a small number of participants and had a relatively high proportion of missing data at T₁. The high odds ratios and wide confidence intervals for some radiological parameters (in particular, occipital and cerebellar haematoma location) indicate that there are insufficient data to support robust conclusions about these predictors. Nevertheless, fitting a model to these data may generate hypotheses worthy of investigation in a larger data-set.

3.3.2.2. Model 3: construction of an initial model

With these caveats in mind, an initial model containing only C-RMI at T₀ was fitted. This accounted for 18.7%-24.9% of the unexplained variance in outcome, and correctly classified 68% of cases. Age did not make a statistically significant contribution to the model, and was therefore removed. MoCA scores at T₀ were of borderline statistical significance ($p=0.079$), and the resulting model accounted for little more variance in outcome (22.7%-30.3%) than did C-RMI alone. The combination of the two variables correctly classified 73.6% of cases. It is therefore possible that this variable might have attained statistical significance in a larger sample size. It was therefore retained in the model. The proposed final iteration of Model 3 thus contained: C-RMI and MoCA scores at T₀.

3.3.2.3. Model 3: model evaluation and testing assumptions of logistic regression

The Hosmer and Lemeshow test indicates acceptable overall model fit ($p=0.536$). No cases had standardised residuals of $>\pm 2.58$: although this implies that there is no evidence of localised misfit, it does raise the possibility that Model 3 is over-fitted to the small sample available. The model was examined for influential cases using Cook's distance, leverage statistics, and DFBeta values for the constant and each of the predictors. No cases had a Cook's distance of >1 . Leverage values ranged from 0.0191 to 0.189. Twenty-six patients had a leverage value greater than the expected value of 0.04. No cases returned DFBeta values >1 for the constant or for either of the predictor variables.

The Box-Tidwell test was performed to test the assumption of a linear relationship between predictor variables and the natural log of the odds ratio of walking independently at T_1 . Prior to natural-log transformation, a score of 1 was added to all values of C-RMI and MoCA, to eliminate scores of 0. The covariates C-RMI and MoCA score at T_0 were then forced in to a model which also contained the interaction terms $(C-RMI+1*\ln[C-RMI+1])$ and $(MoCA+1*\ln[MoCA+1])$. The natural-log interaction terms did not make a significant contribution to this model ($p>0.05$), indicating that the assumption of linearity of the logit was met for all variables.

3.3.2.4. Model 3: final version

The final iteration of Model 3 is summarised in Table 3.14.

Table 3.14. Final iteration of Model 3.

Predictor variables: C-RMI and MoCA scores at T_0 .

	B	S.E.	Wald	df	Sig	OR	95% CI for OR	
							Lower	Upper
C-RMI at T_0	0.643	0.206	9.704	1	0.002	1.901	1.269	2.849
MoCA at T_0	0.082	0.049	2.814	1	0.093	1.085	0.986	1.193
Constant	-2.721	1.076	6.399	1	0.011	0.066		

3.3.2.5. Model 3: summary of model characteristics

This model explained 22.7%-30.3% of variance in outcome at T_1 , and correctly classified 73.6% of patients as walking or not walking at this time. Taken at face value, Model 3 indicates that each one-point increase in C-RMI score at T_0 is associated with a 90.1% increase in the odds of walking independently at T_1 , and each one-point increase in MoCA score at T_0 is associated with an 85% increase in the odds of walking independently at T_1 . However, this apparently-impressive effect size is likely to reflect inflation of the coefficients due to the small sample size from which Model 3 was derived. Although mobility and cognitive function at T_0 do seem to be of prognostic importance in ICH, these apparent associations must be confirmed in a larger sample size.

The observed versus predicted classification of patients with ICH as “walking independently”/“not walking independently” by T₁ is shown in Table 3.15.

Table 3.15. Classification table for Model 3.

		Walking independently at T ₁ (predicted)		% correct
		No	Yes	
Walking independently at T ₁ (observed)	No	25	9	73.5
	Yes	10	28	73.7
Overall %				73.6

The model had a sensitivity of 73.7% and a specificity of 73.5%. Its positive predictive value was 75.6%, and its negative predictive value was 71.4%.

3.3.2.6. Model 3: testing assumptions made for missing data

The default assumption made for missing outcome data in Model 3 was that patients who did not return a SR-RMI at T₁ score were unable to walk independently. To explore the impact of this assumption, Model 3 was re-fitted with two alternative assumptions made. Assumption 1 was that those who did not return a SR-RMI score at T₀ were excluded from analysis: this restricted the available population to only 60 participants. Assumption 2 was that all participants lost to follow-up were able to walk independently at T₁. The properties of Model 3 under these alternative assumptions are outlined in Table 3.16.

Table 3.16. Properties of Model 3 when fitted under alternative assumptions for missing SR-RMI scores at T₁.

	Default	Assumption 1	Assumption 2
% correctly classified	73.6%	66.1%	70.8%
Sensitivity	73.7%	81.6%	86.3%
Specificity	73.5%	38.1%	33.3%
Positive predictive value	75.6%	70.5%	75.9%
Negative predictive value	71.4%	53.3%	50%
% Variance Cox&Snell R ²	22.7%	24.6%	19.8%
% Variance Nagelkerke R ²	30.3%	33.7%	28.3%
C-RMI at T ₀			
B	0.643	0.841	0.904
OR	1.901	2.319	2.469
95% CI for OR	1.269- 2.849	1.331 – 4.041	1.329 – 4.380
MoCA at T ₀			
B	0.082	0.026	-0.021
OR	1.085	1.027	0.979
95% CI for OR	0.986 – 1.193	0.914 – 1.153	0.887 – 1.080

The percentage of cases correctly classified ranged from 66.1% (assumption 1) to 73.5% (under the default assumption). The percentage increase in the odds of walking independently for each one-point change in C-RMI at T₀ ranged from 84.9% (default assumption) to 146.9% (assumption 2). The coefficients for the MoCA score at T₀ were unstable, remaining positive under the default assumption and assumption 1, but becoming negative under assumption 2. Hence, under assumption 1 each one-point change in MoCA score resulted in an increase in the odds of walking independently at T₁ of 2.7%; under assumption 2, each one-point increase in score resulted in a 2.1% *decrease* in the odds of walking independently. This instability most probably reflects the small sample size in which these estimates were derived. The sensitivity of Model 3 ranged from 73.7% (default assumption) to 86.3% (under assumption 2). The range of possible values for specificity was broad: between 33.3% (assumption 2) and 73.5% (default assumption). Positive predictive values lay between 70.5% (assumption 1) and 75.9% (assumption 2); negative predictive values between 50% (assumption 2) and 71.4% (default assumption). The percentage of variance explained by the model

ranges from 19.8% (Cox&Snell R^2 , assumption 2) to 33.7% (Nagelkerke R^2 , assumption 1).

3.3.3. Model 4: return to walking at T2 in HWS group

3.3.3.1. Univariate predictors of outcome at T₂: demographics, and clinical impairment at T₀ and T₁

Model 4 predicted return to independent walking at T₂ in the HWS group. Univariate associations between walking ability at T₂, and demographics and clinical impairment at T₀ and T₁ are shown in Table 3.17.

Table 3.17. Univariate models of independent walking ability at T₂.

Predictor variables: age; gender; clinical impairment at T₀; clinical impairment at T₂.

	N	Missing (%)	Mean (Range, SD)	<i>p</i>	OR	95% CI for OR Lower Upper	
Demographic predictor variables							
Age	75	0	65.85 (32-92, 1.501)	0.002	0.937	0.898	0.977
Male gender	47	0	-	0.611	0.783	0.305	2.012
Predictor variables at T₀							
C-RMI	75	0	1.96 (0-6, 1.664)	0.002	1.896	1.276	2.815
MoCA	72	3 (4%)	20.17 (0-30, 5.891)	0.006	1.148	1.041	1.267
GHQ-12	71	4 (5.3%)	20.84 (8-36, 6.836)	0.643	0.984	0.918	1.054
Any MSK pain	21	3 (4.0%)	-	0.307	0.587	0.211	1.634
Upper limb pain	12	3 (4.0%)	-	0.248	0.476	0.135	1.676
Lower limb pain)	14	3 (4.0%)	-	0.241	0.493	0.151	1.607
Predictor variables at T₁							
SR-RMI	60	15 (20%)	8.12 (0-15, 4.342)	<0.0005	1.628	1.271	2.086
MoCA	64	11 (14.7%)	23.48 (8-30, 4.694)	0.002	1.250	1.084	1.442
GHQ-12	59	16 (21.3%)	15.71 (3-30, 6.734)	0.072	0.925	0.851	1.007
FAS	59	16 (21.3%)	24.23 (11-40, 6.44)	0.023	0.896	0.814	0.985
Any MSKI pain	47	16 (21.3%)	-	0.526	1.524	0.415	5.598
Upper limb pain	33	16 (21.3%)	-	0.512	1.437	0.486	4.251
Lower limb pain	28	16 (21.3%)	-	0.413	1.579	0.529	4.710

3.3.3.2. Univariate predictors of outcome at T₂: radiological variables

Univariate associations between imaging variables and walking ability at T₂ are shown in Table 3.18.

Table 3.18. Univariate predictors of independent walking ability at T₂.

Predictor variables: haematoma volume; presence of midline shift; haematoma location; presence of hydrocephalus; presence of intraventricular extension; white matter lesions; old stroke lesion.

	N	Missing (%)	Mean (Range, SD)	Sig	OR	95% CI for OR	
						Lower	Upper
ICH volume (mm ³)	67	8 (10.7%)	16.58 (1-88, 17.023)	0.422	0.988	0.960	1.017
Midline shift (mm)	66	9 (12%)	1.99 (0-44, 5.700)	0.492	1.052	0.911	1.214
ICH location	64	11 (14.7%)	-	0.996			
Frontal	10	-	-	0.996			
Temporal	1	-	-	1.000	0.000	0.000	-
Parietal	6	-	-	0.790	1.333	0.161	11.075
Occipital	1	-	-	1.000	>1000	0.000	-
Basal ganglia	38	-	-	0.734	1.282	0.306	5.366
Internal capsule	2	-	-	0.999	>1000	0.000	-
Brainstem	0	-	-	-	-	-	-
Cerebellum	6	-	-	0.697	0.667	0.087	5.127
Hydrocephalus	12	9 (12%)	-	0.041	0.250	0.066	0.942
Intraventricular extension	26	8 (10.7%)	-	0.047	0.355	0.127	0.987
Old vascular lesion	13	2 (2.7%)	-	0.363	0.571	0.171	1.910
White matter lesions	26	0	-	0.157	0.498	0.190	1.307
Atrophy	5	2 (2.7%)	-	0.115	0.165	0.017	1.553

Once again, attention is drawn to the small overall sample size, and to the high proportion of missing data for some T₁ predictor variables. As for Model 3, the numbers of participants with certain radiological features (temporal, occipital, capsular, or cerebellar haematomas) were small, and as such confidence intervals for these lesions are either implausibly large or incalculable. It must again be stressed that models derived from these data are exploratory, and should thus be interpreted with caution.

3.3.3.3. Model 4: construction of an initial model

The modelling process began by fitting the strongest predictors measured at T₀. The initial model containing C-RMI at T₀ only explained 16.9%-22.8% of

variance in outcome, and correctly classified 65.3% of patients. The addition of age made a significant ($p=0.010$) contribution to the model; the combination of this and C-RMI accounted for 24.0%-32.3% of unexplained variance, and correctly classified 69.3% of patients. The addition of MoCA scores at T_0 resulted in a model that explained 29.9%-40.2% of variance in outcome, and correctly classified 75.0% of patients.

Variables measuring clinical impairment at T_1 were next fitted. The addition of SR-RMI at T_1 was statistically significant ($p<0.0005$), but the contribution of C-RMI at T_0 was no longer statistically significant ($p=0.955$). This suggests that the effect of C-RMI at T_0 are entirely mediated by that of SR-RMI at T_1 . A model containing SR-RMI at T_1 explained more variance in outcome (49.0%-67.9%) and correctly classified a greater percentage of patients (86.4%) than the previous iteration containing C-RMI at T_0 . However, there is a strong clinical rationale for including variables measured at T_0 in preference to those measured at T_1 , since T_0 marks the time at which patients would typically be engaged in their early rehabilitation. C-RMI at T_0 was retained in the model, and SR-RMI at T_1 was removed. MoCA score and FAS at T_1 and two CT variables (presence of hydrocephalus and presence of intraventricular extension) all failed to make a statistically significant contribution to the model, and were thus removed. The resulting model contained MoCA and C-RMI scores at T_0 , and age. Although this model explained 29.9%-40.2% of variance in outcome, and correctly classified 75.0% of patients, the small sample size from which it was derived makes it desirable to fit the smallest number of predictor variables possible. The variable with the smallest effect size was age (coefficient -0.049; odds ratio 0.952). Removing this variable resulted in a model containing only MoCA and C-RMI scores at T_0 : a combination which still explained 25.1%-33.8% of variance in outcome, and classified 72.2% of patients correctly.

3.3.3.4. Model 4: model evaluation and testing assumptions of logistic regression

This model showed acceptable fit to the data (Hosmer and Lemeshow goodness-of-fit test, $p=0.563$). Only one case showed evidence of localised misfit (standardised residual -2.808). This patient had a comparatively high C-

RMI score of 5 at T₀, and was able to walk independently by T₁. By T₂ his mobility had deteriorated, with a SR-RMI score of 3. It is possible that he sustained a further stroke or an intercurrent illness in the intervening period. The model was examined for influential cases using Cook's distance, leverage statistics, and DFBeta values for the constant and each of the predictors (Table 53). No cases had a Cook's distance of >1. Leverage values ranged from 0.0202 to 0.181. Thirty-three patients had a leverage value greater than the expected value of 0.04. No cases returned DFBeta values >1 for the constant or for either of the predictor variables.

The Box-Tidwell test was performed to assess the assumption of a linear relationship between predictor variables and the natural log of the odds ratio of walking independently at T₂. Prior to natural-log transformation, a score of 1 was added to all values of C-RMI and MoCA at T₀, to eliminate scores of 0. The covariates C-RMI and MoCA score at T₀ were then forced in to a model which also contained the interaction terms (age*Ln[age]), (C-RMI+1*Ln[C-RMI+1]) and (MoCA+1*Ln[MoCA+1]). The natural-log interaction terms did not make a significant contribution to this model (p>0.05), indicating that the assumption of linearity of the logit was met for all continuous variables included.

3.3.3.5. Model 4: final version

The final iteration of Model 4 is summarised in Table 3.19.

Table 3.19. Final version of Model 4.

Predictor variables: MoCA and C-RMI at T₀.

	B	S.E.	Wald	df	Sig	OR	95% CI for OR	
							Lower	Upper
MoCA at To	0.127	0.053	5.778	1	0.016	1.135	1.024	1.259

C-RMI at T ₀	0.616	0.214	8.305	1	0.004	1.852	1.218	2.816
Constant	-3.298	1.152	8.196	1	0.004	0.037		

3.3.3.6. Model 4: Summary of model characteristics

Each one-point increase in C-RMI score at T₀ appears to result in an 85.2% increase in the odds of walking independently at T₂. The effect size for MoCA score at T₀ is smaller, with each one-point increase resulting in a 13.5% increase in the odds of regaining independent mobility at T₂. The apparent effect sizes for these variables are larger than those seen in Models 1 and 2. However, this is likely to reflect the smaller sample size in which Model 4 was derived. Once again, it must be stressed that the analysis presented here is exploratory in nature: the value of these variables in predicting prognosis requires confirmation in a larger sample.

The observed versus predicted classification of patients with ICH as “walking independently”/“not walking independently” by T₂ is shown in Table 3.20.

Table 3.20. Classification table for Model 4.

		Walking independently by T ₂ (predicted)		% correct
		No	Yes	
Walking independently by T ₂ (observed)	No	18	12	60.0
	Yes	8	34	81.0
Overall %				72.2

Model 4 had a sensitivity of 81.0% and a specificity of 60.0%. The positive predictive value of this model was 73.9%, and its negative predictive value was 69.2%.

3.3.3.7. Model 4: testing assumptions made for missing data

The default assumption made for missing outcome data in Model 4 was that the mobility of patients who did not return a SR-RMI score at T₂ was unchanged since the score at T₁, if this was known. If no SR-RMI score had been returned at T₁, patients were assumed to be unable to walk independently at T₂. To explore the impact of this assumption, Model 4 was re-fitted with three alternative assumptions made for missing outcome data.

Assumption 1 excluded patients who did not return a SR-RMI at T₂ from analysis, and was thus tested in only 56 of the 75 patients in the HWS group. Assumption 2 was that all patients who did not return a SR-RMI at T₂ were able to walk independently, whilst assumption 3 was that those with missing data were unable to walk independently at T₂. The properties of Model 3 when re-fitted under these alternative assumptions are summarised in Table 3.21.

Table 3.21. Properties of Model 4 when fitted under alternative assumptions for missing SR-RMI scores at T₂.

	Default	Assumption 1	Assumption 2	Assumption 3
% correctly classified	72.2%	76.4%	76.4%	70.8%
Sensitivity	81.0%	44.4%	98.1%	75.7%
Specificity	60.0%	91.9%	11.1%	65.7%
Positive predictive value	73.9%	77.3%	76.8%	70.0%
Negative predictive value	69.2%	72.7%	66.7%	71.9%
% Variance Cox&Snell R ²	25.1%	23.1%	9.7%	27.9%
% Variance Nagelkerke R ²	33.8%	32.2%	14.4%	37.2%
MoCA at T ₀				
B	0.127	0.112	0.042	0.160
OR	1.135	1.118	1.043	1.174
95% CI for OR	1.024 – 1.259	0.993 – 1.260	0.950 – 1.144	1.045 – 1.319
C-RMI at T ₀				
B	0.616	0.655	0.480	0.593
OR	1.852	1.926	1.616	1.809
95% CI for OR	1.218 – 2.816	1.156 – 3.208	1.034 – 2.525	1.221 – 2.679

The percentage of cases correctly classified ranged from 70.8% (assumptions 3) to 76.4% (assumptions 1 and 2). The percentage increase in the odds of walking independently for each one-point change in MoCA score at T₀ ranged from 11.8% (assumption 1) to 17.4% (assumption 3). Each one-point change in C-RMI score at T₀ resulted in an increase in the odds of walking independently at T₂ of between 61.6% (assumption 2) and 92.6% (assumption 1). The sensitivity of Model 4 ranged from 44.4% (assumption 1) to 98.1% (assumption 2). The range of possible values for specificity was also broad: between 11.1% (assumption 2) and 91.9% (assumption 1). Positive predictive values lay between 70.0% (assumption 3) and 77.3% (assumption 1); negative predictive values between 66.7% (assumption 2) and 72.7%

(assumption 1). The percentage of variance explained by the model ranges from 9.7% (Cox&Snell R^2 , assumption 2) to 37.2% (Nagelkerke R^2 , assumption 3).

Part 3.4 Modelling walking ability in the whole DARS sample

3.4.1. Characteristics of the DARS sample

3.4.1.1. What predictor variables were fitted to Models 5 and 6?

Models 1-4 presented above were constructed using two sub-groups (IWS and HWS) of the DARS sample. The purpose of examining ischaemic stroke and ICH separately, and using only those patients for whom imaging was available, was to assesses the importance of imaging variables as predictors of mobility. However, none of the CT variables examined had a strong univariate association with outcome at T₁ or T₂. Two additional models were therefore constructed, this time utilising all 593 patients in the DARS sample. Model 5 predicted recovery of independent mobility at T₁; Model 6 predicted mobility outcome at T₂. Only demographic and clinical impairment recorded at T₀ were utilised as predictors in these models. Although the inclusion of variables recorded at T₁ may explain more variance in outcome and correctly classify a greater percentage of patients as mobile or immobile at T₂, models that utilise only information available to the rehabilitation team at the time of the patient's entry in to an acute rehabilitation programme are likely to be more useful in a clinical setting.

3.4.1.2. Determining how many variables might be fitted

At T₁, 252 (42.5%) patients were able to walk independently for 10m or more; 341 (57.5%) were unable to do so. If a guideline of ten patients per variable is applied to the smaller of the two outcome groups, then a maximum of $252/10=25$ variables may be fitted to model 5. Allowing for a more generous ratio of 20 patients per variable would allow up to 12 variables to be fitted (Stoltzfus, 2011). By T₂, 311 (52.4%) patients were able to walk independently, with 282 (47.6%) unable to do so. If a ratio of 10 patients per variable is adopted, then 28 variables could be fitted to model 6. Allowing a more conservative 20 patients per variable would permit 14 variables to be fitted (Stoltzfus, 2011).

3.4.2. Model 5: return to walking at T₁ in the whole DARS sample

3.4.2.1. Univariate predictors of outcome at T₁

Model 5 sought to predict walking ability at T₁ using demographic details and clinical impairment at T₀. Univariate associations between predictor variables and this outcome are summarised in Table 3.22.

Table 3.22. Univariate predictors independent walking ability at T₁.

Predictor variables: age; gender; OCSP clinical stroke syndrome; clinical impairment at T₀.

	N	Missing (%)	Mean (Range, SD)	Sig	OR	95% CI for OR		
						Lower	Upper	
Demographic variables	Age	593	0	68.42 (20-98, 13.232)	<0.0005	0.971	0.958	0.983
	Gender (male)	364	0	-	0.015	1.542	1.085	2.140
	Infarct (versus ICH)	508	0	-	0.020	0.578	0.364	0.918
	OCSP	592	1 (0.2%)	-	0.001	-	-	-
	TACS	161	-	-	0.001	-	-	-
	PACS	178	-	-	0.008	1.836	1.172	2.877
	POCS	52	-	-	0.002	2.747	1.446	5.216
	LACS	116	-	-	0.011	1.913	1.163	3.147
ICH	85	-	-	0.000	2.777	1.612	4.784	
Physical impairment at T ₀	C-RMI	593	0	2.25 (0-6, 1.791)	<0.0005	1.714	1.533	1.916
	MoCA	580	13 (2.2%)	20.23 (0-30, 6.308)	<0.0005	1.074	1.043	1.106
	GHQ-12	570	23 (3.5%)	19.36 (3-36, 6.848)	0.180	0.983	0.960	1.908
	Any pain	236	10 (1.7%)	-	0.395	0.864	0.618	1.209
	UL pain	114	10 (1.7%)	-	0.950	0.987	0.652	1.493
	LL pain	154	10 (1.7%)	-	0.020	0.636	0.433	0.932

3.4.2.2. Model 5: construction of an initial model

The first variable to be fitted was age. Although highly statistically significant ($p < 0.0005$), this explained only 3.6%-4.8% of variance in outcome and correctly classified 59.9% of participants. The addition of C-RMI to the model

explained a much greater percentage of the variance in outcome (18.3%-24.6%), and correctly classified 70.5% of cases. The addition of MoCA score at T₀ explained slightly more variance in outcome (19.2%-25.8%), but did not improve the classification rate (69.4%). The MoCA score was nevertheless retained in the model, due to the theoretical importance of this variable in learning processes.

The OCSF classification was considered for inclusion in the model next. This was developed as a clinical classification of ischaemic stroke only (Bamford et al., 1991), and typically includes only four categories: TACS, PACS, POCS, and LACS. The inclusion of both ischaemic stroke and ICH in the DARS sample necessitated the creation of a fifth category, "ICH," to allow this variable to be utilised in a model predicting outcome in a mixed sample of stroke patients. However, to enter OCSF with an additional "ICH" category alongside the dichotomised variable "infarct versus ICH" risked creating collinearity between the two variables. OCSF (with an ICH category) and stroke type dichotomised as "infarct versus ICH" were therefore entered in to separate models which also included age, C-RMI at T₀, and MoCA at T₀.

A model incorporating OCSF with an "ICH" category explained 21.4%-28.8% of variance in outcome, and correctly classified 71.0% of cases. This was little different from a model including "infarct versus ICH" as a dichotomous variable, which accounted for 21.2%-28.5% of variance in outcome, and correctly classified 70.3% of cases. When OCSF (with an ICH category) and dichotomised "infarct versus ICH" were forced in to the model simultaneously, only "infarct versus ICH" retained statistical significance ($p=0.001$). This variable was therefore retained in the model, and OCSF was removed. The addition of both gender and the presence of lower limb pain failed to contribute significantly to the model, and they were therefore removed.

3.4.2.3. Model 5: model evaluation and testing assumptions of logistic regression

The final iteration of the model thus contained age, C-RMI and MoCA scores at T₀, and infarct versus ICH. This model showed acceptable fit to the data (Hosmer and Lemeshow goodness-of-fit, $p=0.529$). Ten cases (1.69% of the

total sample) had standardised residuals of $>\pm 2.58$. These cases were examined to explore possible reasons for misfit. Cases 164, 173, 238, 312, and 563 all previously showed misfit to Model 1, and the reasons for this have been examined above. Participants 28 and 555 both had an initial C-RMI of 0, and were thus predicted to be unable to walk independently by T₁. However, both made a better recovery than might be anticipated, with an SR-RMI scores at T₁ of 10 and 7 respectively. In the case of participant number 28, this gain in motor function was sustained at T₂ (SR-RMI=12). Patient number 555 unfortunately failed to sustain this initial recovery, and by T₂ had a SR-RMI of 2. Participants 68 and 532 had a comparatively high C-RMI of 6 at T₀, and might therefore have been expected to be walking independently at T₁. However, neither returned an SR-RMI at T₁, and both were therefore classified as being unable to walk. Patient number 99 had an initial C-RMI of 6, and MoCA of 27. Despite being predicted to walk independently at follow-up, his scores at T₁ were unchanged. It is not possible to determine a reason for this with the available data. The model was examined for influential cases using Cook's distance, leverage statistics, and DFBeta values for the constant and each of the predictors. No cases had a Cook's distance of >1 . Leverage values ranged from 0.00270 to 0.03903; 201 patients had leverage values greater than the expected value of 0.00843. No cases returned DFBeta values >1 for the constant or for any of the predictor variables.

The Box-Tidwell test was performed to assess the assumption of a linear relationship between predictor variables and the natural log of the odds ratio of walking independently at T₁. Prior to natural-log transformation, a score of 1 was added to all C-RMI and MoCA scores, to eliminate scores of 0. The covariates age, C-RMI and MoCA score at T₀ were then forced in to a model which also contained the interaction terms (age*Ln[age]), (C-RMI+1*Ln[C-RMI+1]) and (MoCA+1*Ln[MoCA+1]). The categorical term "infarct versus ICH" was omitted from the model. The natural-log interaction terms did not make a significant contribution to this model ($p>0.05$), indicating that the assumption of linearity of the logit was met for age, C-RMI and total MoCA score.

3.4.2.4. Model 5: final version

The final iteration of Model 5 is summarised in Table 3.23.

Table 3.23. Final version of Model 5.

Predictor variables: age, MoCA and C-RMI at T₀, and infarct versus ICH.

	B	S.E.	Wald	df	Sig	OR	95% CI for OR	
							Lower	Upper
Age	-0.017	0.007	5.179	1	0.023	0.983	0.969	0.998
C-RMI at T ₀	0.530	0.060	77.358	1	0.000	1.699	1.509	1.911
MoCA at T ₀	0.041	0.016	6.374	1	0.012	1.042	1.009	1.076
Infarct (versus ICH)	-0.805	0.268	9.028	1	0.003	0.447	0.264	0.756
Constant	-0.509	0.691	0.543	1	0.461	0.601		

3.4.2.5. Model 5: summary of model characteristics

This model explained 20.5%-27.5% of variance in outcome, and correctly classified 70.3% of cases. Once again, the largest effect size seen was for C-RMI. Each one-point increase in score at T₀ was associated with an increase of 69.9% in the odds of walking independently at T₁. The effect of cognitive impairment remained more modest, with each one-point increase in MoCA score at T₀ increasing the odds of walking independently by 4.2% at T₁. Each one-year increase in age resulted in a 1.7% decrease in the odds of walking independently at T₁. Finally, at T₁ those with an ischaemic stroke had a 44.7% reduction in the odds of being mobile when compared with those with intracerebral haemorrhage.

The observed versus predicted classification of patients in the DARS sample as “walking independently”/“not walking independently” by T₁ is shown in Table 3.24.

Table 3.24. Classification table for Model 5.

		Walking independently by T ₁ (predicted)		% correct
		No	Yes	
Walking independently by T ₁ (observed)	No	269	62	81.3
	Yes	110	139	55.8
Overall %				70.3

Model 5 had a sensitivity of 55.8% and a specificity of 81.3%. The positive and negative predictive values of this model were, respectively, 69.2% and 71.0%.

3.4.2.6. Model 5: testing assumptions made for missing data

The default assumption made for missing outcome data in Model 5 was that patients who did not return a SR-RMI score at T₁ were unable to walk independently. To explore the impact of this assumption, the model was re-fitted with two alternative assumptions. Assumption 1 excluded patients who did not return a SR-RMI at T₁ from analysis, and was thus tested in only 393 of the original 593 participants. Assumption 2 was that all patients who did not return an SR-RMI at T₁ were able to walk independently. The properties of Model 5 when fitted under these alternative assumptions are summarised in Table 3.25.

Table 3.25. Properties of Model 5 under alternative assumptions for missing SR-RMI scores at T₁.

	Default	Assumption 1	Assumption 2
% correctly classified	70.3%	71.7%	69.1%
Sensitivity	55.8%	68.3%	72.0%
Specificity	81.3%	74.5%	65.8%
Positive predictive value	69.2%	71.7%	71.3%
Negative predictive value	71.6%	71.6%	66.5%
% Variance Cox&Snell R ²	20.5%	25.0%	19.3%
% Variance Nagelkerke R ²	27.5%	33.4%	25.8%
Age			
B	-0.017	-0.017	-0.012
OR	0.983	0.983	0.989
95% CI for OR	0.969 – 0.998	0.967 – 0.999	0.974 – 1.003
C-RMI at T ₀			
B	0.530	0.609	0.518
OR	1.699	1.839	1.678
95% CI for OR	1.509 – 1.911	1.608 – 2.103	1.487 – 1.893
MoCA at T ₀			
B	0.041	0.041	0.031
OR	1.042	1.042	1.031
95% CI for OR	1.009 – 1.076	1.007 – 1.077	1.001 – 1.063
Infarct versus ICH			
B	-0.805	-1.087	-1.063
OR	0.447	0.337	0.345
95% CI for OR	0.264 – 0.756	0.185 – 0.615	0.199 – 0.598

The percentage of cases correctly classified ranged from 70.3% (default assumption) to 71.7% (assumption 1). The percentage decrease in the odds of walking independently at T₁ with each one-year increase in age ranged from 1.1% (assumption 2) to 7.7% (default assumption and assumption 1). Each one-point increase in C-RMI score at T₀ resulted in an increase in the odds of independent mobility at T₁ of between 67.8% (assumption 2) and 83.9% (assumption 1). The odds of walking independently increased by between 3.1% (assumption 2) and 4.3% (default assumption and assumption 1) for each one-point increase in MoCA score at T₀. The odds of walking

independently for an infarct (versus ICH) were between 34.5% (assumption 2) and 44.7% (default assumption).

The sensitivity of Model 5 ranged from 55.8% (default assumption) to 72.0% (assumption 2). The range of values for specificity was between 65.8% (assumption 2) and 81.3% (default assumption). Positive predictive values lay between 69.2% (default assumption) and 71.7% (assumption 1); negative predictive values between 66.5% (assumption 2) and 71.6% (assumption 1). The percentage of variance explained by the model ranges from 19.3% (Cox&Snell R^2 , assumption 1) to 33.4% (Nagelkerke R^2 , assumption 1).

3.4.3. Model 6: return to walking at T₂ in the whole DARS sample

3.4.3.1. Univariate predictors of outcome at T₂

Model 6 predicted recovery of walking ability at T₂. Univariate associations between this outcome and predictor variables (demographics and impairment at T₀) are summarised in Table 3.26.

Table 3.26. Univariate models of independent walking ability at T₁.

Predictor variable: age; gender; infarct (versus haemorrhage); OCSF clinical stroke syndrome; and clinical impairment at T₀.

		N	Missing (%)	Mean (Range, SD)	Sig	OR	95% CI for OR	
							Lower	Upper
Demographic variables	Age	593	0	68.42 (20-98, 13.232)	<0.0005	0.957	0.944	0.970
	Gender (male)	364	0	-	0.011	1.540	1.104	2.147
	Infarct (versus ICH)	508	0	-	0.422	0.827	0.521	1.314
	OCSF	592	1 (0.2%)	-	0.018			
	TACS	161	-	-	0.018			
	PACS	178	-	-	0.110	1.419	0.924	2.177
	POCS	52	-	-	0.001	3.127	1.591	6.146
	LACS	116	-	-	0.163	1.406	0.871	2.270
	ICH	85	-	-	0.066	1.644	0.968	2.793
Physical impairment at T ₀	C-RMI	593	0	2.25 (0-6, 1.791)	<0.0005	1.519	1.367	1.688
	MoCA	580	13 (2.2%)	20.23 (0-30, 6.308)	<0.0005	1.080	1.050	1.111
	GHQ-12	570	23 (3.5%)	19.36 (3-36, 6.848)	0.609	0.994	0.970	1.018
	Any pain	236	10 (1.7%)	-	0.246	0.822	0.590	1.145
	UL pain	114	10 (1.7%)	-	0.973	1.007	0.669	1.518
	LL pain	154	10 (1.7%)	-	0.042	0.681	0.471	0.986

3.4.3.2. Model 6: construction of an initial model

As for model 5, age made a statistically significant contribution ($p < 0.0005$), but explained only a small percentage of the variance in outcome (7.3%-9.8%) and correctly classified 63.9% of cases. The addition of C-RMI made a statistically significant contribution ($p < 0.0005$) to the model. Alongside age, it explained 15.7%-21.0% of the variance in outcome and correctly classified 68.3% of cases. MoCA score at T₀ was next fitted. The resulting model

explained 17.5%-23.4% of the variance in outcome and correctly classified 69.0% of cases. Gender, OCSP clinical stroke syndrome, and lower limb pain all failed to reach statistical significance, and did not contribute any additional explicable variance. Nor did they improve the percentage of patients correctly classified.

3.4.3.3. Model 6: model evaluation and testing assumptions of logistic regression

The model therefore comprised age, and C-RMI and MoCA scores at T₀. This model was an acceptable fit to the data (Hosmer and Lemeshow test $p=0.606$). Examination of residuals for individual cases demonstrated 7 cases (1.18% of the 593 cases included) in which standardised residuals were $>\pm 2.58$ (Table 78). Two patients (68 and 173) had standardised residuals >3 . Five participants (numbers 68, 84, 173, 532, and 547) had relatively high C-RMI scores at T₀, but failed to return a SR-RMI at T₂. All were classified as being unable to walk, despite relatively favourable C-RMI scores at T₀. Participant number 99 had an initial C-RMI score of 6, but failed to achieve any improvement in mobility at either T₁ or T₂. It is not possible to establish a definite cause for this with the data available, although the fact that he developed both upper and lower limb musculoskeletal pain between the initial assessment at T₀ and first follow-up at T₁ may provide some indication of why he failed to progress in rehabilitation.

The model was examined for influential cases using Cook's distance, leverage statistics, and DFBeta values for the constant and each of the predictors (Table 79). No cases had a Cook's distance of >1 . No cases returned DFBeta values >1 for the constant or for any of the predictor variables. Leverage values ranged from 0.00215 to 0.0327: 202 participants had leverage values greater than the expected value of 0.00674.

The Box-Tidwell test was performed to assess the assumption of a linear relationship between predictor variables and the natural log of the odds ratio of walking independently at T₂. Prior to natural-log transformation, one point was added to all C-RMI and MoCA scores to eliminate scores of 0. The covariates age, C-RMI, and MoCA score at T₀ were then forced in to a model

which also contained the interaction terms (age*Ln[age]), (C-RMI+1*Ln[C-RMI+1]) and (MoCA+1*Ln[MoCA+1]). The natural-log interaction terms did not make a significant contribution to this model ($p>0.05$), indicating that the assumption of linearity of the logit was met for age, C-RMI and total MoCA score at T₀.

3.4.3.4. Model 6: final version

The final iteration of Model 6 is summarised in Table 3.27.

Table 3.27. Final version of Model 6

Predictor variables: age, C-RMI, and MoCA at T₀.

	B	S.E.	Wald	df	Sig	OR	95% CI for OR	
							Lower	Upper
Age	-0.035	0.007	21.601	1	<0.0005	0.966	0.952	0.980
C-RMI at T ₀	0.378	0.057	43.506	1	<0.0005	1.459	1.304	1.633
MoCA at T ₀	0.046	0.015	9.219	1	0.002	1.048	1.017	1.079
Constant	0.747	0.650	1.321	1	0.250	2.111		

3.4.3.5. Model 6: Summary of model characteristics

This model explained 17.5%-23.4% of variance in outcome, and correctly classified 69.0% of patients as walking/not walking independently at T₂. Each one-point increase in C-RMI at T₀ resulted in a 45.9% increase in the odds of walking independently at T₂. A one-point increase in the MoCA score at T₀ conferred a more modest 4.8% increase in the odds of walking independently at T₂. Each one-year increase in age was associated with a 3.5% reduction in the odds of walking independently at T₂. The observed versus predicted classification of patients in the DARS sample as “walking independently”/“not walking independently” by T₂ is shown in Table 3.28.

Table 3.28. Classification table for Model 6.

		Walking independently by T2 (predicted)		% correct
		No	Yes	
Walking independently by T2 (observed)	No	185	89	67.5
	Yes	91	215	70.3
Overall %				69.0

Model 6 had a sensitivity of 70.3%, a specificity of 67.5%, a positive predictive value of 70.7%, and a negative predictive value of 67.0%.

3.4.3.6. Model 6: testing assumptions made for missing data

The default assumption under which Model 6 was fitted was that patients who did not return SR-RMI scores at T₂ maintained the same level of mobility that had been recorded at T₁. Those who had also failed to return a SR-RMI at T₁ were assumed to be unable to walk independently at T₂. In order to test the robustness of model parameters, Model 6 was re-fitted using three alternative assumptions. The first excluded those who had not returned a SR-RMI at T₂ from analysis, and was tested in only 487 of the original 593 participants. The second was that all patients who did not return a SR-RMI at T₂ were able to walk independently at this time. The third was that those who did not return a SR-RMI at T₂ were unable to walk independently. The properties of Model 6 when fitted under these alternative assumptions are summarised in Table 3.29.

Table 3.29. Properties of Model 6 when fitted under alternative assumptions for missing SR-RMI scores at T₂.

	Default	Assumption 1	Assumption 2	Assumption 3
% correctly classified	69.0%	60.1%	68.6%	66.4%
Sensitivity	70.3%	79.5%	88.9%	62.8%
Specificity	67.5%	57.6%	27.2%	69.9%
Positive predictive value	70.7%	71.6%	71.3%	67.3%
Negative predictive value	67.0%	65.1%	54.7%	65.6%
% Variance Cox&Snell R ²	17.5%	19.0%	11.7%	13.6%
% Variance Nagelkerke R ²	23.4%	25.7%	16.2%	18.1%
Age				
B	-0.035	-0.036	-0.026	-0.022
OR	0.966	0.965	0.974	0.979
95% CI OR	0.952 – 0.980	0.948 – 0.982	0.959 – 0.989	0.965 – 0.992
C-RMI at T ₀				
B	0.378	0.461	0.356	0.336
OR	1.459	1.586	1.427	1.399
95% CI for OR	1.304 – 1.633	1.385 – 1.815	1.264 – 1.610	1.257 – 1.556
MoCA at T ₀				
B	0.046	0.050	0.030	0.046
OR	1.048	1.051	1.030	1.048
95% CI for OR	1.017 – 1.079	1.017 – 1.087	1.000 – 1.061	1.017 – 1.079

The percentage of cases correctly classified ranged from 60.1% (assumption 1) to 68.6% (assumption 2). The percentage decrease in the odds of walking independently at T₂ with each one-year increase in age ranged from 2.1% (assumption 3) to 3.5% (assumption 1). Each one-point increase in C-RMI score at T₀ was associated with an increase in the odds of independent mobility at T₂ of between 39.9% (assumption 3) and 58.6% (assumption 1). The odds of walking independently increased by between 3.0% (assumption 2) and 5.1% (assumption 1) for each one-point increase in MoCA score at T₀. The sensitivity of Model 6 ranged from 62.8% (assumption 3) to 88.9%

(assumption 2). The range of values for specificity was between 27.2% (assumption 2) and 69.9% (assumption 3). Positive predictive values lay between 67.3% (assumption 3) and 71.6% (assumption 1); negative predictive values between 54.7% (assumption 2) and 65.6% (assumption 3). The percentage of variance explained by the model ranges from 11.7% (Cox&Snell R^2 , assumption 2) to 25.7% (Nagelkerke R^2 , assumption 1).

Chapter 4. Discussion

Part 4.1 Summary of results

4.1.1. The aims of this study

This Thesis sought to develop a series of models to predict return to independent walking after stroke at T₁ and T₂ (roughly eight weeks and six months after stroke). Few models have been developed to predict mobility outcomes specifically. Those that have tend to rely upon an assessment of clinical impairment (van de Port et al., 2006a). Whether or not including an assessment of structural brain impairment using imaging variables offers any additional predictive value over and above an assessment of clinical impairment alone remains uncertain (Stinear and Ward, 2013, Dawes et al., 2008). However the models derived here examined a series of predictor variables including limitation of activity and impairment in brain structure (of which CT imaging was a surrogate marker) in the hope of predicting recovery of the ability to walk 10m or more independently at up to six months after stroke.

4.1.2. Summary of models

4.1.2.1. Summary of key findings

The most striking finding of the work presented here is that none of the brain imaging variables examined here predict walking ability at T₁ and T₂ (up to six months after stroke). Whilst the small sample size of the HWS group is acknowledged, this suggests that models based upon measures of clinical impairment alone might provide a more reliable prediction of rehabilitation potential than those incorporating imaging variables. With this in mind, it is worth examining in more detail the findings of models 5 and 6, which were developed in the whole DARS sample and did not incorporate radiological variables. In both of these models the strongest predictor variable was C-RMI. In model 5, this variable alone increased the percentage of explicable variance from 3.6-4.8% (for age alone) to 18.3-24.6% (for age and C-RMI). The

percentage of cases correctly classified increased from 59.9% (age alone) to 70.5% (age plus C-RMI). Similarly, in model 6 the addition of C-RMI alongside age to the model increased the variance explained from 7.3-9.8% to 15.7-21.0%, although there was a more modest rise in the percentage of cases correctly classified (from 63.9% with age alone to 68.3% of cases with age plus C-RMI). When additional variables were fitted (MoCA at T_0 and infarct versus ICH for model 5, and MoCA at T_0 for model 6) the final iterations of the models explained only marginally more variance than age plus C-RMI (20.5-27.5% for model 5; 17.5-23.4% for model 6), and resulted in little or no improvement in the percentage of patients correctly classified (70.3% for the final iteration of model 5; 69.0% for the final iteration of model 6). This suggests that a clinician-scored RMI performed within eight weeks of stroke (defined here as " T_0 ", and representing the time window during which patients were randomised to DARS) is a useful predictor of independent walking ability at up to six months. Depending upon the assumptions that were made for missing data, model 5 had a sensitivity of between 55.8-72.0%, and a specificity of 65.8-81.3%; model 6 had a sensitivity of 62.8-88.9% and a specificity of 27.2-69.9%.

4.1.2.2. Key predictor variables

The variables that were most consistently predictive of outcome across all models were C-RMI and MoCA scores at T_0 . Of the two, the largest effect size was seen for C-RMI. In the IWS group, each one-point increase in this score increased the odds of walking independently at by 71.9% at T_1 , and 48.5% at T_2 . Estimates in the HWS group are derived from a much smaller sample, and must thus be treated with caution. Taken at face value, a one-point increase in C-RMI at T_0 is associated with a 90.1% increase in the odds of walking independently at T_1 , and an 85.2% increase in the odds of walking independently at T_2 . In the DARS sample as a whole, each one-point increase was associated with a 69.9% increase in the odds of walking independently at T_1 , and a 45.9% increase in the odds of walking independently by T_2 . Cognitive function at T_0 was also a significant predictor of outcome, albeit with a smaller effect size than C-RMI. In the IWS group, each one-point increase MoCA score at T_0 increased the odds of walking independently by 2.9% at T_2 .

Findings in the HWS group must be viewed with caution, but a one-point increase in MoCA scores at T_0 was associated with an 8.5% increase in the odds of regaining independent mobility by T_1 and a 13.5% increase in the odds of walking independently at T_2 in this group. In the whole DARS sample, a one-point increase in MoCA score at T_0 was associated with a 4.2% increase in the odds of walking independently at T_1 , and a 4.8% increase in the odds of regaining independent mobility at T_2 . Age at stroke onset attained significance only in models 2, 5, and 6. In the IWS group, each one-year increase in age was associated with a 3.2% decrease in the odds of walking independently at T_2 . Similarly, in the DARS sample as a whole a one-year increase in age was associated a 1.7% reduction in the odds of walking independently at T_1 , and a 3.5% reduction in the odds of walking independently at T_2 . The type of stroke (infarct versus ICH) was a significant predictor variable only in Model 5 (outcome at T_1 in the whole DARS sample). Here, those with an ischaemic stroke had a 44.7% reduction in the odds of being mobile at T_1 when compared with those with intracerebral haemorrhage.

4.1.2.3. Classification of patients and percentage variance explained

The models presented here correctly classified between 68% and 73% of participants as walking/ not walking independently at the specified end-points: lower than that achieved by other models which successfully predicted RMI score to within ± 2 points in 81% of patients (van de Port et al., 2006a). They also account for only around 20-30% of variance in outcome at best: lower than the 50% accounted for in other models (van de Port et al., 2006a).

The success of the models in classifying patients as walking/not walking independently may be judged by examining how accurately an outcome might be predicted if no model were fitted at all. For example, in the DARS sample as a whole, 42.5% of patients were mobile at T_1 , and 57.5% were immobile. Without any further modelling, one can state that DARS participants have a higher prior probability of being immobile at T_1 than of being mobile. Thus, if an individual were selected from the sample at random and a guess were made about their likely mobility at T_1 , it would be reasonable to predict that the participant would be immobile at that time. If the same assumption were made for every patient in the sample, then without any further modelling

walking ability at T_1 would be correctly predicted in 57.5% of patients. Fitting Model 5 to the data increases the accuracy of prediction, and allows 70.3% of patients to be correctly classified. This sounds impressive: but in reality, only 12.8% more participants are correctly classified by the model than would be the case using prior probability alone.

Furthermore, a patient's rehabilitation potential is currently estimated not by statistical models, but following an expert assessment by an experienced multi-disciplinary team. Little has been published about how such decisions are made, and what factors inform therapists' prognostication. And yet, an apparent "educated guess" by skilled therapists may be as reliable in predicting recovery of walking ability as a regression model (Kwakkel et al., 2000). In order to be useful in clinical practice, a prognostic model would not only need to result in a much higher rate of correct classification than has been achieved here: it must also be demonstrably superior to the judgement of experienced therapists.

4.1.2.4. Sensitivity, specificity, and positive and negative predictive values

If these models are to be used in clinical practice, one must also bear in mind the differing balances of sensitivity and specificity that each displays. Models 1 and 5, both predictive of walking ability at T_1 , have relatively poor sensitivity (40.3% for Model 1, 55.8% for Model 5) but higher specificity (88.5% for Model 1, 81.3% for Model 5). This implies that both are relatively poor predictors of who *will* return to walking by T_1 , but more reliably classify those who will *not* do so. Models derived to predict mobility outcomes at T_2 (Model 2 and Model 6) have moderate sensitivity (68.3% for Model 2; 70.3% for Model 6) and specificity (70.3% for Model 2; 67.5% for Model 6). They thus perform moderately well in predicting both positive ("walking") and negative ("not walking") outcomes. Models 3 and 4 were derived from a small sample of patients with ICH, and must therefore be interpreted cautiously. In Model 3, sensitivity and specificity are balanced (73.7% and 73.5% respectively); Model 4 has a higher sensitivity (81.0%) than specificity (60.0%), indicating a greater reliability in identifying those who will return to walking than those who will not.

4.1.2.5. Stability of models under alternative assumptions for missing data

Estimates for the effect sizes of each of the predictors fitted changed considerably depending on the assumptions made for missing outcome data. The most worrying example of this was seen in Model 3, in which a one-point increase in MoCA score at T₀ was associated with a 2.7% *increase* in the odds of walking independently at T₁ under assumption 1, but a 2.1% *decrease* in the odds of independent mobility under assumption 2. This example is extreme, and reflects the small sample size from which these estimates were derived. Nevertheless, even models fitted to larger samples demonstrated changes in effect size when alternative assumptions were made for missing data. In Model 5, derived from the whole DARS sample, a single-point increase in C-RMI score at T₀ was associated with a 67.8% increase in the odds of walking independently at T₁ under assumption 2, but an increase of 83.9% under assumption 1. In Model 6 (outcome at T₂), effect size estimates for C-RMI ranged from 39.9% (assumption 3) to 58.6% (assumption 1). Furthermore, the sensitivity and specificity of the models also varied considerably when alternative assumptions for missing data were tested. The widest ranges were in models predicting outcome at T₂, since a greater proportion of the sample had been lost to follow-up at this time than at T₁. In Model 2, sensitivity ranged from 61.4% to 85.3%, whilst specificity lay between 30.0% and 71.6%. In Model 6, sensitivity was 62.8%-88.9% and specificity 27.2%-69.9%. Estimates of positive predictive value tended not to vary by more than a few percentage points. The greatest ranges of values observed was in Models 2 (66.8%-72.0%) and 3 (70.0%-77.3%). The greatest range in estimates of negative predictive values was in Model 6 (54.7%-65.9%).

4.1.3. The need for formal validation of the models

The models presented above all suggest that a patient's level of mobility and cognitive function early after stroke might be useful in predicting walking ability. However, they show marked instability in both estimates of the effect sizes of covariates and of other key parameters such as sensitivity and specificity when alternative assumptions are made for missing data. This

raises concerns that the performance of the models derived here may also vary if applied to a different sample.

Such instability might arise when the characteristics of the sample in which a model was developed are different from the population to which it is ultimately applied. For example, a model developed in a high-income country might later be applied in a low- or middle-income country (Vickers and Cronin, 2010), or a model developed in an adult sample might be extrapolated to children (Moons et al., 2015, Moons et al., 2009a).

Over-fitting of a model may also result in the inclusion of predictors that have little value beyond the derivation sample or, conversely, exclusion of predictors that might be important in a different cohort (Moons et al., 2009a).

It is therefore recommended that models undergo validation before introduction into clinical use. Ideally validation should be conducted in a sample different from the original (Altman and Royston, 2000). In practice however, validation is more usually performed by splitting the original data set, with half of the sample being used for model derivation and half for validation (Altman and Royston, 2000, Labarere et al., 2014, Moons et al., 2015). Although convenient, this practice is unsatisfactory. Firstly, the sample size available for derivation of the model is reduced, resulting in loss of statistical power (Altman and Royston, 2000, Labarere et al., 2014, Moons et al., 2015). Secondly, unless caution is exercised the characteristics of the “derivation” and “validation” samples may be virtually identical (Altman and Royston, 2000). For these reasons, “split sample” validation was not performed for the models presented here.

When considering whether it might be appropriate to undertake formal validation of the models presented here, one must consider what is already known about them and their possible clinical utility. The proportion of variance in outcome that they explain is relatively modest, and the classification rates they achieve are little better than prior probability alone and perhaps no better than the clinical judgement of a skilled multidisciplinary team. For these reasons, the models developed here cannot at present be recommended for use in clinical practice in their current form.

In clinical practice knowledge of the variables that might influence recovery, and an estimation of their possible effect sizes, might still provide useful information to inform the judgement of rehabilitation professionals and thus to guide discussions with patients and families. Although the findings presented above must be interpreted with caution, they do suggest that those with the most profound mobility impairment and cognitive dysfunction after stroke are less likely to recover the ability to walk. The development of reliable models to predict mobility will depend in part upon identifying the most crucial predictor variables. Examining the contribution of the variables included in the models presented here might be a starting point for this process.

Part 4.2 Comparing the findings of this study with previous literature

As summarised above, the variables that were included in Models 1-6 were: C-RMI (mobility); MoCA (cognitive function); age; and type of stroke (infarct or ICH). Several variables that might have been expected to contribute to prediction of walking ability did not make a statistically-significant contribution to multivariate models. These included gender; OCSP clinical stroke syndrome; depression; fatigue; and musculoskeletal pain. Finally, none of the imaging predictor variables that were assessed were independent predictors of which patients might walk again.

4.2.1. Variables that were predictors of walking ability in the models presented here

4.2.1.1. Mobility early after stroke

The variable with the largest effect across any of the models was C-RMI at T₀. The finding that a patient's level of mobility in the early stages after stroke is predictive of their later ability to walk again was noted previously by Shum et al. (2014) who used a modified version of the RMI (M-RMI) (Lennon and Johnson, 2000) administered at 3 days to predict walking ability at 28 days after stroke. Unlike the original RMI (Collen et al., 1991), which has 15 questions with binary responses, the M-RMI has eight domains with polytomous responses ranging from 0 (unable to perform) to 5 (independent) (Lennon and Johnson, 2000). An M-RMI score of ≥ 18.5 at day 3 predicted independent walking ability at day 28 with a sensitivity of 85% and specificity of 75% (Shum et al., 2014).

A patient's initial level of mobility appears to be an important predictor of not only later independent mobility, but also of discharge destination after stroke (Sommerfeld and von Arbin, 2001, Brauer et al., 2008). Sommerfeld and von Arbin (2001) found that those achieving RMI scores of ≥ 4 at 10 days after stroke had a 29-fold increase in the risk of being discharged home within three months (relative to those with scores of < 4). A similar pattern was noted by Brauer et al. (2008), who utilised a different measure of mobility, the Motor

Assessment Score (Carr et al., 1985). Each one-point increase on the “gait” sub-scale of this measure increased the odds of stroke survivors returning home 1.66-fold, and each single-point increase in the “balance” sub-scale increased the odds of discharge home by a factor of 1.28.

If initial mobility is a useful indicator of long-term outcome, it is useful to ask what *aspect* of “mobility” is of greatest prognostic importance. The DARS trial, for example, enrolled only participants who had an initial C-RMI score of <7 at T₀: and were thus initially unable to walk independently (Bhakta et al., 2014). Hence, although the effects of C-RMI in models 1-6 were quoted above as the percentage change in the odds of walking independently at follow-up associated with a one-point increase in C-RMI score, it is perhaps more accurate to specify that this prediction holds true only for values of C-RMI *in the range of 0-6*. Initial differences in low-level mobility functions (short of actually walking) may therefore have an important bearing on a patient’s subsequent outcome.

If the RMI is examined more closely, then It could be argued that RMI values within the range of 2 (able to transfer from lying to sitting in bed) to 5 (ability to stand for 10 seconds unaided) (Collen et al., 1991) appear to measure the underlying construct of *balance* rather than mobility. Indeed, re-establishing adequate balance, first in sitting and then in standing, are necessary prerequisites for walking. Interestingly, a model constructed by van de Port et al. (2006a) to predict RMI scores at one year after stroke found sitting balance (measured using the trunk control test) to be a significant predictor of this outcome. Indeed, this accounted for 8% of observed variance in RMI scores at 12 months (van de Port et al., 2006a). Interestingly, the range of the metrics for the “gait” and “balance” sub-scales of the Mobility Assessment Scale (Carr et al., 1985), utilised by Brauer et al. (2008) as a predictor of discharge destination, do align with the lower-end scores of the RMI (Collen et al., 1991). Perhaps future modelling in patients who are initially immobile should include specific measures of balance, as opposed to more global measures of mobility such as the RMI.

4.2.1.2. Cognitive function after stroke

As described in Chapter 1, the phenomenon of cognitive impairment after stroke is complex. It is a common consequence of stroke, both acutely and in the longer term (Gottesman et al., 2010). And given the importance of loop circuits between cortical and sub-cortical structures in motor learning (Doyon et al., 2009, Hikosaka et al., 2002, Penhune and Steele, 2012), and the potential for these pathways to become disrupted by both injury to particular structures as a result of the stroke itself (Schmahmann and Pandya, 2008, Schmahmann et al., 2009) and as a consequence of chronic microvascular disease (Iadecola, 2013, Wardlaw et al., 2013b), one might expect cognitive function to be a strong predictor of recovery of walking ability. However, although MoCA scores did indeed make a statistically significant contribution to the models, the effect observed for each one-point increase was relatively modest in comparison to C-RMI score. This suggests that any relationship between cognition and motor recovery is not straightforward.

Cognitive dysfunction is associated with a decline in mobility at up to three years after stroke (van de Port et al., 2006c), but the precise relationship between motor and cognitive functions remains uncertain (Chen et al., 2013). Whether “cognition” can be regarded as a single underlying construct is debatable: several key cognitive domains are recognised including memory, executive function, praxis, and visuospatial function (Barker-Collo et al., 2010). It is possible that mobility impairment after stroke is due in part to difficulty in allocating limited cognitive resources to a complex task such as walking, rather than to a failure of learning processes *per se*. The phenomenon of dual-task interference, in which performance of a motor task is compromised when the patient is asked to perform a cognitive task simultaneously, is well recognised (Chen et al., 2013). This was initially attributed to an inability to focus attention to two tasks at once, but there is now growing evidence that the underlying impairment is actually in executive function (Chen et al., 2013). Interestingly, impairment in motor function (measured using the Berg balance score, gait velocity, and “timed up-and-go” test) within the first six months after mild stroke also seem to predict the development of cognitive dysfunction at up to two years later (Ben Assayag

et al., 2015). Giving fewer correct answers to questions asked during walking (i.e. under dual-task conditions) also predicted subsequent cognitive decline (Ben Assayag et al., 2015). It is therefore possible that aspects of gait and motor task performance are influenced by subtle impairments in specific cognitive functions, with more florid cognitive dysfunction becoming apparent later on. The link between motor and cognitive dysfunction may therefore be mediated by impairment in specific cognitive domains such as executive function. The MoCA, which was designed as a brief screening test of global cognitive function (Nasreddine et al., 2005), may lack sensitivity for more subtle domain-specific impairments. Although the MoCA does contribute to the overall predictive ability of models 2-5, it is possible that more detailed tests of specific cognitive ability, such as executive function, could account for a greater proportion of observed variance in outcome.

4.2.1.3. Age at stroke onset

Age at stroke onset attained significance only in models 2, 5, and 6. Age has been shown to be a negative correlate of mobility at between one and five years in several models: (van de Port et al., 2006a, Meyer et al., 2015, Sanchez-Blanco et al., 1999) but, in keeping with findings above, its effect is modest. It accounts for only 3% of variance in motor outcome (van de Port et al., 2006a, Bagg et al., 2002), and each one-year increase in age is associated with an increase of only 0.08 points on the leg and trunk function sub-scale of the Rivermead Motor Assessment and a 0.13-point increase in the gross function subscale of this measure (Meyer et al., 2015). It is likely that the effect of age is mediated by other variables such as cognitive impairment or comorbidities (Bagg et al., 2002). For this reason, age must not be the only criterion in deciding whether a patient might benefit from rehabilitation (Bagg et al., 2002).

4.2.1.4. Type of stroke

The association between type of stroke and outcome is uncertain. Although the early mortality from ICH is high (Dennis, 2003), there have been suggestions that those who survive are more likely than those with ischaemic stroke to experience a favourable outcome. This was the case in Model 5, as discussed above. Several other studies also point to better outcomes following

ICH, although the association is by no means proven. A univariate analysis by van de Port et al. (2006a) found that the presence of intracerebral haemorrhage was associated with an increase of 0.176 points on the RMI at one year when compared to ischaemic stroke. Patients with ICH have also been shown to have higher RMI scores at discharge from a rehabilitation programme (mean RMI=6.57, SD \pm 4.33) compared with those with ischaemic stroke (mean RMI=5.42, SD \pm 3.90) (Paolucci et al., 2003), and to show a greater change in the motor subscale of the Functional Independence Measure at discharge from rehabilitation (Δ FIM-motor 22.7 \pm 1.1 for ICH; 20.2 \pm 0.5 for ischaemic stroke) (Kelly et al., 2003). However, although these findings are all *statistically* significant it is debatable whether they meet the threshold for true *clinically significant* change. Also, Meyer et al. (2015) did not detect statistically significant differences between survivors of ischaemic stroke and ICH in Rivermead Motor Assessment (RMA) gross function and arm and trunk subscale scores at any time point. Although there was statistically significant gain of 1.82 points on the RMA arm function subscale for ICH when compared with ischaemic stroke, the clinical relevance of this finding is debatable (Meyer et al., 2015).

A large prospective cohort study also found that patients with ICH were more likely to be dead or dependent at three months and one year (Bhalla et al., 2013). However, data collection in this cohort spanned a 16-year period between 1995 and 2011; hence, not all patients enrolled received what would today be regarded as “gold standard” care (Bhalla et al., 2013). For example, patients with ischaemic stroke were more likely to be managed in the community than those with ICH; a practice that is now regarded as outmoded. In short, the evidence that the type of stroke a patient sustains influences functional prognosis is mixed. Type of stroke therefore cannot at present be used to predict rehabilitation potential.

4.2.2. Clinical predictor variables that did not make a contribution to models 1-6

Several variables that could plausibly have been predictors of mobility ultimately failed to make an independent contribution to the models presented above. These will be examined below.

4.2.2.1. Gender

Gender reached statistical significance as a univariate predictor of outcome in models 1, 2, 5, and 6, with men having more favourable odds of achieving independent mobility at eight weeks and six months than women. However, this variable was not statistically significant on multivariate analysis. These findings are in keeping with those of previous studies, which showed that female gender was associated with a reduction in one-year RMI score compared to male (van de Port et al., 2006a). Once again, this finding did not reach statistical significance in a multivariate analysis (van de Port et al., 2006a). Meyer et al. (2015) failed to demonstrate a statistically significant influence of gender on outcome at any time point up to five years.

4.2.2.2. OCSP clinical stroke syndrome

Clinical stroke syndrome, defined using the OCSP classification (Bamford et al., 1991), was only fitted to models that included patients with ischaemic stroke but did not attain significance in any of these. The OCSP classification has been shown to be of prognostic relevance: but only in the distinction between TACS and other syndromes (PACS, LACS, POCS). At six months after stroke, those with TACS have lower mean RMA scores, a longer length of hospital stay, and greater mortality rates (Pittock et al., 2003). Compared with patients who sustain a LACS, at 3 months after stroke those with TACS are 3.27 (95% CI 2.30-4.66) times more likely to experience limitation of activities (Barthel index of 0-14), and 2.71 (95% CI 1.91-3.85) times more likely to be disabled (Rankin scale of 2-5) (Di Carlo et al., 2006). Although TACS carries clear prognostic implications, the distinction between PACS, LACS, and POCS is uncertain.

4.2.2.3. Depression and fatigue

Depression (GHQ-12) did not attain significance in any of the univariate analyses presented above. The interaction between mood and motor recovery is complex (Chen et al., 2013). Nannetti et al. (2005) found no difference in motor outcome (Fugl-Meyer Assessment Scale) three months after stroke between patients who had been depressed on admission versus those who were not. However, depression has been shown to be associated with long

term deterioration in mobility (defined as a fall of ≥ 2 points on the RMI) at one year (OR 3.44; 95% CI 1.57–7.54) (van de Port et al., 2006b) and at two years (OR 4.2; 95% CI 1.3–13.2) (van Wijk et al., 2006) after stroke. Fatigue reported at T₁ (FAS) was a significant univariate predictor for model 4 ($p=0.023$), but failed to reach significance in the final multivariate model. Other studies, however, have found the presence of fatigue to be predictive of long-term deterioration in mobility: in both a univariate model (OR 3.30; 95% CI 1.09–9.99) and as part of a multi-variate model (OR 3.30; 95% CI 1.09–9.99) alongside depression and cognitive impairment. (van de Port et al., 2006b) It is therefore possible that any effects of depression and fatigue on motor function are exerted on a timescale longer than the six months post-stroke considered here.

4.2.2.4. Musculoskeletal pain

In the DARS sample the proportion of patients reporting musculoskeletal pain in any locus increased from 39.8% at baseline, to 66.9% at eight weeks and 63.2% at six months. The percentage reporting pain in any lower limb locus was 26.0% at baseline, 40.5% at eight weeks, and 43.2% at six months. This is somewhat higher than previous estimates, which indicate that the prevalence of any joint pain within two years of stroke is around 55.4% (Hettiarachchi et al., 2011). The presence of lower limb musculoskeletal pain was a significant univariate predictor in model 1 (assessment at T₀; $p=0.045$), model 2 (assessment at T₁; $p=0.042$); model 5 (assessment at T₀; $p=0.02$), and model 6 (assessment at T₀; $p=0.042$). However, it failed to achieve significance as part of multivariate models. None of the other musculoskeletal pain parameters assessed (presence of any pain, and presence of lower limb pain) made even a univariate contribution to predicting outcome. This is surprising, since previous studies have found that the presence of both stroke-related impairment and hip pain increased the odds of reporting difficulties in standing and mobilising by a far greater magnitude than either of these impairments alone (Hettiarachchi et al., 2011). A possible reason for the discrepancy between the findings presented here and those of Hettiarachchi et al. (2011), may be that the latter study chose to classify pain by laterality and as specific loci; for example “left knee”, “right hip”. The absolute numbers

reporting pain in each locus are not presented, but the wide confidence intervals for some analyses indicate that some combinations of impairments included relatively few participants (Hettiarachchi et al., 2011). The present study, by contrast, chose to classify pain as “any musculoskeletal pain,” “any upper limb pain,” or “any lower limb pain.” This increased the number of patients available for analysis in each group, with commensurately narrower confidence intervals. However, it is possible that grouping pain loci in this way masked clinically significant differences in mobility resulting from pain in particular joints.

4.2.3. Imaging predictor variables

4.2.3.1. What imaging variables were evaluated?

A number of imaging variables were considered as possible predictors of motor recovery. For ischaemic stroke, these included: the presence of any visible abnormality; the presence of visible acute ischaemic change; the location of infarcts (cortical, subcortical, or both); the size of an infarct (small, medium, or large); the presence of old stroke lesions; the presence of white matter lesions; and the presence of atrophy. The ASPECTS score, although associated with “death or dependency” outcome (Barber et al., 2000), makes no assessment of ischaemic change that lies outside of the MCA territory, or of ICH: groups that in total accounted for 69.6% of the DARS sample. For this reason, it was not utilised as a predictor here. For ICH, the variables examined were: haematoma volume and location; the presence of midline shift, intraventricular extension, and hydrocephalus; and the presence of atrophy and old stroke lesions. The protocol for image analysis used here was developed by Wardlaw et al, and has been used in the IST-3 trial (The IST collaborative group, 2015). The predictor variables utilised here for ischaemic stroke are therefore those used by Wardlaw et al in an analysis of imaging findings from the IST-3 data-set (The IST collaborative group, 2015). This presents an opportunity to compare imaging findings between the two data sets, before examining why imaging variables failed to make a significant contribution to predicting outcome in the present study.

4.2.3.2. Comparison of imaging findings in DARS with those of patients in the IST-3 trial

Unlike DARS, the IST-3 trial enrolled only patients with ischaemic stroke, and thus resembles most closely the IWS group from which Models 1 and 2 were derived. It is to this sub-group of the main DARS sample that imaging findings from IST-3 (The IST collaborative group, 2015) will be compared. Firstly, there are striking differences in the sizes of ischaemic lesions reported in the two samples. In the DARS trial, 47.5% of the IWS group had no visible infarct; this was the case in a greater proportion (59%) of the IST-3 sample (The IST collaborative group, 2015). When comparing the size of visible infarcts 13.2% of the IWS group and only 7% of the IST-3 sample had a small infarct; 29.9% of the IWS group and 17% of the IST-3 group had medium infarcts (The IST collaborative group, 2015). In DARS, the frequency of “very large” infarcts (as defined by Wardlaw et al (The IST collaborative group, 2015)) was small, and this category was therefore amalgamated with “large” infarcts to give a total prevalence of “large” infarcts of 7.3%. In IST-3, “large” infarcts accounted for 9% of the total, and “very large” infarcts 8% (The IST collaborative group, 2015). In short, the IWS group had a greater proportion of small and medium-sized infarcts, fewer large lesions, and fewer patients with no visible ischaemia than the IST-3 sample (The IST collaborative group, 2015). Several methodological differences between the two trials may account for these findings. Firstly, the DARS protocol utilised the first available imaging that had been collected for clinical indications, but did not specify when this should be acquired. By contrast, the IST-3 protocol required that imaging be performed within 6 hours of stroke onset (The IST collaborative group, 2015). Since early signs of early ischaemia on CT are often subtle and difficult to detect (Wardlaw and Mielke, 2005) it is possible that, in the IST-3 sample, large or very large lesions were more easily detectable from the outset whilst small and medium-sized infarcts were less obvious and thus more likely to be interpreted as “no visible ischaemia”. Aside from radiological considerations the differences in size of infarcts could also be accounted for by the differing aims, and thus inclusion criteria, of the two studies. DARS randomised patients between five and 42 days after stroke (Bhakta et al., 2014). Only patients who were expected to survive for >2 months after stroke were recruited (Bhakta et al.,

2014). By contrast, IST-3 randomised participants within 6 hours of symptom onset (Whiteley et al., 2006). The exclusion from DARS of those thought likely to die within 2 months, may have biased the DARS sample towards patients with more moderate impairment: and hence those with small-to-medium infarcts on CT.

A further interesting and important difference between the two trials is that the prevalence of atrophy, previous strokes, and WML were much lower in the DARS IWS group than in IST-3. Atrophy was found in 10.5% of DARS participants, versus 77% of those randomised in to IST-3; 27.2% of DARS patients had evidence of a previous stroke, versus 43% of those in IST-3; and 43.2% of DARS patients showed evidence of WML, compared with 52% of IST-3 patients (The IST collaborative group, 2015). In simple terms, DARS participants appeared to have “healthier” brains prior to the index event than those randomised in to IST-3. As noted for “infarct size” above, perhaps patients with advanced atrophy, multiple old strokes, and WML were excluded from the trial: either because they were deemed to have poor rehabilitation potential, or because it was thought unlikely that they would survive more than two months. However, this observation may also reflect in part differences in the age profile of the DARS IWS group, compared with participants enrolled in to IST-3. The DARS trial randomised a greater proportion of younger patients than IST-3 (8% versus 4% in the 18-50 age group, 16% versus 7% in the 51-60 group, 24% versus 12% in the 61-70 group). Peak enrolment for DARS was in the 71-80 age group (31% for DARS, versus 24% for IST-3). IST-3 by contrast specifically sought to expand the indications for thrombolysis beyond accepted parameters, and thus made active attempts to recruit a greater proportion of elderly patients in to the trial (Sandercock et al., 2012). Hence, 46% of those enrolled were in the age group 81-90 (17% for the DARS IWS group), and 7% were aged >90 (2% for DARS IWS group) (The IST collaborative group, 2015). The increasing prevalence of atrophy and WMLs with age may thus account for the different frequency of these findings in the two samples.

4.2.3.3. Other studies using imaging variables to predict outcome

None of the imaging variables considered for inclusion in models 1-6 contributed to their predictive ability. There is a long history in the stroke literature of utilising imaging in outcome prediction: but the most widely-used systems for classifying imaging findings were never intended to predict future walking ability. For example, the ASPECTS score was developed to decide upon eligibility for entry to a thrombolysis trial (Barber et al., 2000), but has since been found to predict good versus poor functional outcome (Dzialowski et al., 2006, Hill et al., 2003) and the risk of haemorrhagic transformation of an infarct following thrombolysis (Hirano et al., 2012). Models to predict outcome from ICH have typically focused upon identifying those who might benefit from aggressive intervention (Godoy et al., 2006, Ruiz-Sandoval et al., 2007, Hemphill et al., 2001, Cheung and Zou, 2003, Cho et al., 2008), reliable prognostication (Tuhim et al., 1991, Cheung and Zou, 2003), or stratification of patients for entry in to clinical trials (Broderick et al., 1993, Godoy et al., 2006, Tuhim et al., 1991, Hemphill et al., 2001, Cheung and Zou, 2003).

Several studies have evaluated the use of imaging to predict outcome in ischaemic stroke. Many were published over ten years ago, and therefore utilised imaging technology that is now obsolete (Johnston et al., 2002, Wardlaw et al., 1998, Saver JL et al., 1999, Candelise L et al., 1991, Johnston et al., 2000). The majority use “death or dependency” as their primary outcome: an outcome measure that is too broad to be of genuine use in planning a rehabilitation programme. Even using such a general measure of outcome, the evidence that neuroimaging variables make a contribution to the model beyond that of clinical variables alone is mixed. Whilst some have found that the presence of any visible infarct on early imaging is associated with an adverse outcome (Wardlaw et al., 1998, Candelise L et al., 1991), more recent reports suggest that this variable confers no extra predictive power beyond that of clinical variables alone (Reid et al., 2010). Similarly, the volume of an infarct on DWI imaging has been shown to be a useful predictor by some (Johnston et al., 2000), but not all (Johnston et al., 2009a, Johnston et al., 2002), studies.

4.2.3.4. Implications for the use of imaging in rehabilitation practice

The role of imaging in predicting specific rehabilitation outcomes therefore remains far from certain, and a detailed evaluation of neuroimaging does not at present form a routine part of rehabilitation prognostication (Stinear and Ward, 2013). Whilst disappointing, the failure of neuroimaging to contribute to predicting the functional prognosis of stroke survivors does have important implications for how limited resources might be used most efficiently in low- and middle-income countries. These nations currently bear the greatest share of the global burden of stroke (Johnston et al., 2009b); and yet access to high-quality acute stroke care is often limited by a combination of cost, lack of trained staff, and geographical centralisation of services in major cities that are inaccessible to the majority of the population (Mendis, 2010). In the absence of acute stroke services, rehabilitation intervention should be the mainstay of stroke management. In countries where access to imaging is restricted (Mendis, 2010), it is useful to note that a thorough clinical assessment of initial impairment, utilising standardised outcome measures, may allow prediction of a patient's prognosis without the need for imaging.

Part 4.3 Limitations of this Thesis

The models described above must be interpreted in the light of several major limitations. Crucial amongst these are: the post-hoc nature of the analysis presented here; the degree to which DARS participants may be regarded as representative of stroke survivors entering rehabilitation programmes in the UK or worldwide; the high rate of loss to follow-up in the DARS sample; and the nature of the outcome measures used and the manner in which they were analysed. All of these limitations will be discussed in the coming section. The analysis of scan findings will be considered separately in Part 4.4. Although all of these factors may ultimately restrict the use of these models in clinical practice, it is perhaps legitimate to regard the findings presented here as an exploratory analysis from which hypotheses worthy of future investigation might be generated.

4.3.1. Analytical considerations

4.3.1.1. Limitations of post-hoc analysis

The first, and possibly the most crucial, limitation of the analysis presented here is that it was conducted post-hoc in a data set originally collected as part of a randomised controlled trial of the effects of administering co-careldopa on motor recovery from stroke. The dangers of such analyses are well recognised. Firstly, the statistical testing of large numbers of predictors increases the potential for type 1 errors to occur: i.e. that a false hypothesis is accepted as correct. This arises when tests of multiple hypotheses yield results that are statistically significant by chance alone (Rothwell, 2005). After all: a p-value of <0.05 merely implies that there is a *95% probability* that the null hypothesis can be rejected. In some cases this can produce results for which there is no rational scientific basis: for example, that patients who were born under the star signs Libra or Gemini do not benefit from aspirin after myocardial infarction ($p=0.02$) (Collins and MacMahon, 2001). In other cases, the consequences for clinical practice can be more profound and even dangerous: for example, the spurious finding that men, but not women, derive benefit from aspirin in ischaemic stroke (Anonymous, 1978) led to the under-treatment of women for at least a decade before new evidence emerged to

challenge this result (Rothwell, 2005). In order to guard against the possibility of type 1 errors a Bonferroni correction may be undertaken, in which the significance level that is accepted is adjusted for the number of hypotheses under test (Dunn, 1961). Such a correction was not undertaken here, and this could have led to variables being included in some models inappropriately. For example, when fitting Model 5, the variable “infarct versus ICH” was significant at the level $p=0.02$ in a univariate analysis. It was thus fitted to the model, where it achieved statistical significance at a level of $p=0.003$. And yet, as discussed above, there is scanty evidence at best to support a hypothesis that this variable is of clinical significance in predicting motor recovery from stroke (Bhalla et al., 2013, Meyer et al., 2015, van de Port et al., 2006a). Models that contain such variables are likely to be over-fitted to the data set in which they were derived, and may thus generalise poorly when applied to other groups.

A second limitation of post-hoc analyses is that any sub-groups that are specified are almost always under-powered in relation to the main sample (Rothwell, 2005): a situation well-illustrated here by models 3 and 4, which were developed in the HWS group. This group consisted of 75 patients, or 12.6% of the whole DARS sample. On univariate analysis, the small number of observed events for some parameters (for example, haemorrhage involving the parietal lobe), resulted in wide confidence intervals and unreliable estimates of their coefficients and odds ratios. Furthermore, instability was also seen in parameters for the final models: most markedly Model 3, in which the MoCA score was by turns both a negative and a positive predictor of walking ability depending upon the assumptions that were made for missing data. Such models may therefore be unreliable, and require formal validation before clinical use.

The derivation of models post hoc from an existing data set also limits the variables that can be fitted to the model to those that have already been collected. These may not be the ones that best predict the outcome in question. This is illustrated well by considering the percentage of variance explained by the models derived above. The largest percentage of variance explained by any of the models above was seen for Model 4 (33.8%,

Nagelkerke R^2); estimates for other models were more typically in the range of 15-25%. The fact that the majority of the variance in mobility outcome remains unexplained by these models indicates that key predictor variables have not been included. The models presented here ultimately included a combination of impairment in bodily functions (mobility, cognition, and type of stroke) and contextual factors (age), although other impairments in brain structure (radiological findings) and in bodily function (pain, fatigue, depression, OCSF), and contextual factors (gender) were also considered for inclusion. It is likely that other impairment-level variables that were not collected here might have been significant predictors of mobility. For example, initial severity of leg weakness (Sanchez-Blanco et al., 1999, Patel et al., 2000, Veerbeek et al., 2011, Wandel et al., 2000), the presence of hemianopia (Sanchez-Blanco et al., 1999, Patel et al., 2000), urinary incontinence (Wade and Hwer, 1987), impairment in sitting balance (Wade and Hwer, 1987, Veerbeek et al., 2011) or standing balance (Kollen et al., 2005), and time between stroke and mobility assessment (Preston et al., 2011, Kwakkel et al., 2006), and activities-level measures such as the BI (Kollen et al., 2006, Paolucci et al., 2008) have all been shown to be associated with limitation of walking ability at six months (Kollen et al., 2005, Patel et al., 2000, Sanchez-Blanco et al., 1999, Veerbeek et al., 2011, Wade and Hwer, 1987, Wandel et al., 2000) or at the time of discharge from inpatient rehabilitation (Paolucci et al., 2008). The National Institute of Health Stroke Scale (NIHSS) (Brott et al., 1989) is a measure of clinical impairment after stroke that is widely used in acute stroke services. It has been shown to be predictive of early mortality after ischaemic stroke (Fonarow et al., 2012) and ICH (Cheung and Zou, 2003), and of functional outcomes (measured using the mRS) (Saver and Altman, 2012). Its relationship to motor outcomes is less certain. A small study of 200 patients (114 of whom were non-ambulant at enrolment) has shown that a combination of age and NIHSS score predict both independent ambulation and recovery of upper limb function at six months (Kwah et al., 2013). The DARS trial might have presented an opportunity to examine the ability of the NIHSS to predict recovery of walking ability in a larger sample; but unfortunately this variable was not collected. Furthermore, in ICF terms recovery from stroke represents a complex

interaction between patients' physical impairments and their environmental/societal contexts (Gutenbrunner et al., 2006, The World Health Organisation, 2011, Dahl, 2002). Although early physical impairment might be important in predicting recovery of mobility, it is likely that factors such as the support from a spouse or partner and family, friends, or carers that a patient is able to access, their home environment and financial situation, and the availability of community resources such as clubs and social groups ultimately have a much greater impact upon a person's quality of life and independence (Meijer et al., 2004).

A final limitation of post-hoc data analysis is that data sets obtained for other purposes may contain biases and can reflect poorly the characteristics of the population to whom a prognostic model might be applied. This issue will be discussed in detail below.

4.3.1.2. Representativeness of the DARS sample

The models developed here were developed in a data-set collected as part of the DARS trial: to date, the largest-ever trial of co-careldopa-augmented stroke rehabilitation. There is, however, a fundamental tension between the requirements of a randomised controlled trial and those of prognostic modelling. When designing a clinical trial, the imperative is to maximise the potential for a true treatment effect to be observed by rigorously controlling for possible confounding variables. For this reason, trials typically set rigid inclusion/exclusion criteria: for example, all DARS participants were expected to survive for more than two months after stroke, and be able to access ongoing rehabilitation on discharge (Bhakta et al., 2014). Practical constraints, such as the requirement for patients to be able to provide informed consent or to be able to swallow medication, further limited the pool of eligible participants. As a result the sample that was ultimately obtained may reflect poorly the whole population of stroke survivors in the UK or worldwide. This poses a problem when interpreting trial results (Rothwell, 2005), since the evidence base for an intervention does not encompass the patients or situations that are commonly encountered in clinical practice.

The lack of generalisability of such data sets also poses a problem for any prognostic models derived from them, since such models are most reliable when derived in samples whose characteristics closely resemble those of the population to which they will ultimately be applied. It is also difficult to control retrospectively for biases that might have been introduced at the time of data collection. In order to understand how useful (or otherwise) the models developed here might be in clinical practice, it is important to understand the DARS sample and how closely it might reflect the population of stroke survivors as a whole.

Around 20-30% of patients who sustain a stroke will die within three months (Department of Health, 2007); and yet, in the DARS trial, all participants were expected to survive for >2 months after recruitment (Bhakta et al., 2014). Just as patients with severe life-limiting strokes were excluded, so those with more minor impairment were excluded by the protocol requirement that participants be unable to walk 10m or more independently at the point of entry in to the trial (Bhakta et al., 2014). The distribution of mRS scores at T_0 indicates that 36.1% of participants had a score of 3, and 29.0% had a score of 4. This indicates a sample with moderate initial impairment. Enrolment of patients with certain other impairments was also impractical. Although around 30% of patients experience cognitive impairment after stroke and 33% have aphasia (The Stroke Association, 2016), those with the most severe degrees of these impairments were excluded from DARS by their inability to give informed consent. Similarly, 45% of stroke survivors experience swallowing problems (The Stroke Association, 2016). In DARS, patients with dysphagia that did not resolve before the end of the recruitment window were excluded, since the trial drug could not be crushed for administration by a nasogastric or gastrostomy tube.

Patients from certain ethnic groups (Afro-Caribbean and South Asian) are more likely to experience strokes at a younger age than those of Caucasian heritage, and are twice as likely to have risk factors for stroke such as diabetes or hypertension (The Stroke Association, 2016). Nevertheless, patients from ethnic minorities are often under-represented in stroke research trials (Cruz-Flores et al., 2011). This is a cause for concern. Firstly, research findings or

prognostic models derived in samples with a non-representative ethnic mix may not be generalisable to patients from minority communities. Secondly, failure to collect data from those from certain communities/backgrounds may limit the ability of clinicians firstly to detect and secondly to understand and mitigate inter-racial differences in rehabilitation outcomes. In DARS, data on patients' ethnic background was not collected: it is therefore not possible to ascertain whether patients enrolled fully represent the ethnic diversity of those who survive a stroke in the UK.

In short, the DARS trial recruited a sample with moderate, predominantly motor, impairment. Patients with more severe degrees of aphasia, cognitive impairment, and dysphagia were not recruited. This sample is therefore not representative of the whole population of people who sustain a stroke worldwide. However, the population of patients who sustain a stroke each year is, perhaps, not an appropriate comparator for the DARS sample. In clinical practice, motor rehabilitation is neither appropriate nor necessary for all those who survive a stroke. For those who sustain a catastrophic stroke, palliative care or planning for discharge to a nursing home might be most appropriate. Conversely, those who sustain a minor stroke with little motor deficit and minimal loss of function will still require rigorous investigation and control of risk factors, but might recover with early supported discharge, without the need for a longer period of inpatient rehabilitation. Hence, although the DARS sample does not reflect the full spectrum of post-stroke impairment, it may more accurately reflect those who enter rehabilitation programmes in the UK. This is an important distinction, since the models developed here will ultimately be applied to patients on rehabilitation units: i.e., who are judged to have the potential to benefit from a period of predominantly physical inpatient rehabilitation.

In DARS, the mean age at randomisation to the trial was 68 years (SD 13.23 years). In a survey of four stroke rehabilitation services in Europe, the mean age of patients on admission to the UK centre was 72.0 (SD 9.5 years) (De Wit et al., 2007). In DARS, 85.7% of patients had an ischaemic stroke and 14.3% of patients had an ICH. Survey data from other UK rehabilitation units indicate that ischaemic stroke accounts for 87% of admissions, and ICH 11%

(De Wit et al., 2007). At the time of enrolment, 13.8% of DARS participants had a mRS score of ≤ 2 , indicating no to moderate disability; this is slightly greater than data from other UK centres, in which 7-11% of patients had a mRS score of ≤ 2 (Putman et al.). In DARS, all patients were unable to walk 10m or more at baseline, with a mean RMI score of 2.25 (SD 1.791). The RMA Leg and Trunk subscale score for patients admitted to rehabilitation programmes in other UK centres is 4, and the mean RMA gross function subscale score was 2 (De Wit et al., 2007): these values indicate respectively an ability to transfer from sitting to standing (equivalent to an RMI score of 4-5) and an ability to transfer from lying to sitting on the side of the bed (equivalent to an RMI score of 2). The baseline level of mobility impairment seen in DARS is therefore comparable to that seen at entry in to inpatient rehabilitation programmes in other UK centres.

Although the baseline characteristics of the DARS sample do reflect those of samples drawn from other rehabilitation units, the proportion of DARS patients who regained independent mobility at T_1 and T_2 is somewhat lower than might be expected. In DARS, 42.5% of patients were walking independently at T_1 and 52.4% at T_2 ; estimates from other studies indicate that 60%-80% of patients are independently mobile at six months post stroke (T_2 in the DARS trial) (Kwakkel and Kollen, 2013). However, such estimates may be confounded by inclusion of patients who were ambulant at the time of enrolment (Preston et al., 2011). The DARS trial enrolled only patients who were unable to walk 10m or more at baseline (Bhakta et al., 2014). Systematic reviews which consider only patients who were non-ambulant at baseline have estimated the probability of walking independently as 0.39-0.60 at three months, and 0.65-0.69 at six months, depending upon whether the patient was managed in a rehabilitation unit or an acute stroke (Preston et al., 2011). The distinction between “acute” and “rehabilitation” units this Australian study makes does not reflect UK practice, in which patients typically receive both acute care and early rehabilitation on an acute stroke unit. Nevertheless, these estimates, derived in patients who were non-ambulant initially, indicate that the proportion of DARS patients who were able to walk independently at T_2 is lower than one might expect. One possible reason for this discrepancy is the comparatively high proportion of DARS patients lost to follow-up at T_2 .

4.3.1.3. Loss to follow-up in the DARS sample

DARS was the first trial of pharmacologically-augmented rehabilitation in which extended follow-up (up to one year) was attempted. However, the proportion of patients lost to follow-up at all end-points was higher than anticipated. By T₁, 61 patients (10.3%) were lost to follow-up and a further eight did not return a SR-RMI score. Hence, the primary outcome measure was unavailable for 69 patients (11.6% of the total sample). By T₂, the cumulative loss to follow-up was 101 patients (17.0%), and a further five did not return usable six-month SR-RMI scores. The total number for whom a SR-RMI was not available at T₂ was therefore 106 patients (17.9% of the total sample enrolled). The primary reasons cited for loss to follow-up at T₂ were: withdrawal from the trial (n=45); death of a patient (n=20); inability of the trial team to contact the patient (n=15); and the patient moving to an area no longer covered by the trial team (n=8).

The percentage of patients lost to follow-up in DARS is higher than in comparable rehabilitation trials. The FLAME trial (Fluoxetine for motor recovery after acute ischaemic stroke) (Chollet et al., 2011) examined the impact of fluoxetine upon motor recovery from stroke. In total 118 patients were randomised across both arms, with 113 remaining in follow-up by 90 days. Only five patients (4.2%) were lost to follow-up (Chollet et al., 2011). A second large multi-centre rehabilitation trial of very early mobilisation after stroke (AVERT) randomised 2104 participants across both arms, to either early mobilisation or usual care (Bernhardt et al., 2015). At three months, 181 (8.6%) were lost to follow-up (Bernhardt et al., 2015). The mortality rates for DARS participants (7 patients (1.1%) at T₁ and 20 patients (3.3%) at T₂) are however lower than that seen in AVERT (160 patients, 7.6%, at 3 months) (Bernhardt et al., 2015).

The DARS intervention was complex (Craig et al., 2008), in that for each participant its successful delivery depended upon an interaction between a number of individuals, spanning several professional disciplines and sometimes two or more services. It is therefore perhaps remarkable that the intervention was delivered successfully in the majority of cases. However, the impact of such high loss rates on the models presented here cannot be

ignored. In particular, the instability in model parameters that was seen when they were fitted under alternative assumptions for missing data has been highlighted above.

4.3.1.4. Outcome measures used

As has been alluded to above, the predictor variables utilised in these models were analysed under an assumption that they provide interval-level measurement: i.e. that a one-point change in their score equates to precisely the same magnitude of clinical change no matter where a patient is located on the metric. However, as illustrated in Chapter 2 using the example of the RMI, this assumption is illogical. This poses a particular problem for regression analysis. When the variables used here were fitted to the models, the output quoted was the change in odds of walking (versus not walking) for each one-point change in a covariate (Stoltzfus, 2011). However, if one cannot be sure that a one-point change in score measured in different patients represents the same magnitude of *clinical* change, then the face validity of such findings is called into question. Nor is it legitimate to perform mathematical operations on ordinal scales, such as the calculation of mean values and standard deviations.

The fundamental difference between ordinal and interval-level measurement is perhaps not as widely appreciated as it should be. And yet, the consequences of making an unfounded assumption that an outcome measure provides interval-level measurement are profound for clinical practice, commissioning of rehabilitation services, and research. This is well illustrated in another field of rehabilitation practice: the management of chronic pain (Kersten et al., 2014). The Pain Visual Analogue Scale (VAS) is a widely-used method for documenting a patient's experience of pain (Kersten et al., 2014). It consists of a line ten centimetres long, with each end of the scale typically anchored with statements such as "no pain" or "worst pain imaginable" (Kersten et al., 2014). Patients are invited to indicate the severity of their pain by placing a mark along the line. Pain severity is typically quoted as a score out of 100, which represents the distance along the scale in millimetres that the mark is positioned (Kersten et al., 2014). Since scores are quoted as a distance in millimetres, they are often assumed to have interval-level

properties and are therefore interpreted as such: for example in the calculation of “change” scores (the difference in VAS scores taken at two time points) (Kersten et al., 2014). The assumption that the pain VAS provides interval-level measurement has recently been challenged. In particular, patients report difficulty in conceptualising their pain as a point on a linear continuum, and thus responses tend to be “clustered” rather than spread evenly along the length of the metric (Kersten et al., 2014). An analysis by Kersten et al. (2014) has recently confirmed that the pain VAS actually behaves as an *ordinal* scale. This fundamentally undermines the assumption that parametric statistics can be applied to this measure.

This is more than just an arcane statistical argument. When results from two randomised controlled trials that had utilised the VAS as an outcome measure were re-analysed, the change in VAS scores quoted in the original papers were found to greatly over-estimate the true change observed when the VAS was first converted to an interval-level measure (Kersten et al., 2014). The impact of these findings cannot be over-stated. Firstly, a change in the VAS score of equal magnitude, but at different locations on the scale, may have fundamentally different clinical implications (Kersten et al., 2012). Secondly, change scores calculated from the VAS may under- or over-estimate true clinical change (Kersten et al., 2012). Thirdly, if change in VAS scores do not reliably measure true clinical change, then the use of this measure to inform decisions about the commissioning of services is questionable (Kersten et al., 2012). Fourthly, if estimates of clinical change derived from the VAS are unreliable, then utilising this measure in sample size calculations may result in under- or over-powered trials from which inappropriate conclusions may be drawn (Kersten et al., 2012).

Caution must therefore be exercised when interpreting the output of Models 1-6. Although the direction and magnitude of effects for each of the covariates fitted to the model may be indicative, they are likely to under- or over-estimate the magnitude of true clinical change. If models are to be derived that reliably predict true clinical change after stroke, then there is a pressing need for robust interval-level scales to measure both key rehabilitation outcomes and predictor variables. The means by which this might be achieved will be

discussed below in Section 5. First however, the limitations of the analysis of brain scans that is presented here will be considered.

4.3.2. Imaging considerations

This study sought to examine whether the characteristics of the stroke, as seen on imaging, contributed to the ability of models to predict motor recovery. Although centres of excellence in imaging research exist (notably the Brain Research Imaging Centre (BRIC) in Edinburgh: www.sbirc.ed.ac.uk), at the time the DARS protocol was under development there was surprisingly little in the published literature to guide non-radiologists who wished to include a secondary analysis of brain imaging in a trial. The Acute Stroke Imaging Research Roadmap (Wintermark et al., 2008), published in 2013, provides a helpful guide to some of the central methodological considerations of robust radiological research and is strongly commended to interested readers. However, it is clear in hindsight that a failure to address some basic considerations when writing the protocol for imaging analysis subsequently led to unanticipated difficulties. The following paragraphs will reflect upon some of the difficulties encountered in this work, and highlight its methodological flaws and their impact upon the models presented here. Although it may also serve as a useful guide to non-experts, those who wish to include an analysis of imaging in a trial are strongly advised to consult experienced academic neuroradiologists from the earliest stages of protocol design.

4.3.2.1. Proportion of scans not available for analysis

A delay in obtaining ethical approval for the centralised review of scans resulted in a failure to collect imaging from the first 24 participants randomised in to the trial. A further 36 scans could not be obtained for other reasons, and 20 were excluded since only MRI imaging and not CT was sent. Hence, 80 participants (13.5% of the DARS sample) did not have imaging available for analysis. Formal testing for differences in the clinical characteristics of those for whom imaging was and was not obtained revealed slightly higher mean SR-RMI scores at T₁ and T₂ in those for whom imaging was not available; however, there was no evidence of a statistically-significant difference in the

proportions of the two groups who were walking independently at eight weeks and six months. Nor were there any other statistically-significant differences in the characteristics of the two groups.

4.3.2.2. The use of routinely-collected as opposed to protocol-specified imaging

Perhaps the most fundamental decision to make when planning an analysis of imaging is whether to utilise scans acquired routinely for clinical purposes, or to stipulate the acquisition of additional imaging beyond that which would normally be considered to be clinically indicated. This will of course depend upon the research question (Wintermark et al., 2008). For example, MRI sequences that are known to have a greater sensitivity and specificity for white matter lesions and lacunes than CT (O'Brien et al., 2003). A study examining the prevalence of these lesions, or changes in them over time, would thus specify that MRI imaging be used. However, requesting additional imaging exclusively for trial purposes inevitably incurs additional costs, which must be included in any grant proposal. As a guide, in 2015-16 NHS England set a funding tariff of £77 for a non-contrast CT of one body area in an adult patient, with an additional £20 for reporting; the tariff for a non-contrast MRI was higher, at £123 plus £22 for reporting (NHS England, 2015). The costs of acquiring and reporting more complex sequences (such as angiography) are commensurately greater. The cost of radiographers' time and patient transport or travelling expenses to and from the scanning facility must also be accounted for.

In DARS, analysis was restricted to the first available CT scan acquired for routine clinical purposes. This pragmatic approach has several advantages: the costs of acquiring imaging specifically for trial purposes are mitigated, and no additional radiation exposure to patients is required. Furthermore, if the clinical utility of a model depends in part upon the predictor variables it utilises being readily collectable in practice (Altman and Royston, 2000), then it is advantageous to construct a model using scans that are already collected routinely. However there are also some clear disadvantages to this approach, including an inability to guarantee that the scanners on which images are

taken are appropriately calibrated, and difficulty in standardising the time at which imaging is acquired in relation to the onset of stroke symptoms.

4.3.2.3. The timing of imaging acquisition

Failing to standardise the time of acquisition of scans has crucial implications for the reliability of reporting, since any scan is only ever a static representation of what is a highly dynamic process. For example, early ischaemic changes on CT are subtle (Grotta et al., 1999, Wardlaw and Mielke, 2005) and may easily be overlooked (Wardlaw et al., 2007). Re-scanning the same patient some weeks later is likely to reveal a mature area of hypodensity typical of an established infarct. Scans acquired within a few hours of stroke onset may thus under-estimate the true structural impairment.

Just as it is difficult to predict the final extent of brain injury from an initial scan, so it may also be difficult to determine the nature of an initial lesion when presented with a scan taken some time after an event. For example, ICH initially returns a hyperdense appearance on CT, but later become hypodense and may closely resemble an old infarct (Macellari et al., 2014, Balami and Buchan, 2012). Similarly, small subcortical infarcts or haemorrhages may give rise to a range of final appearances including: no visible abnormality; a lacune; or white matter lesions (Wardlaw et al., 2013b). In the absence of previous imaging (for example, showing an intracerebral haemorrhage) or a corroborating clinical history (for example, suggesting a lacunar stroke syndrome), clinical judgement must be used to come to a view on the likely clinical significance of scan findings. In DARS, the fact that the timing of imaging in relation to stroke onset was not standardised raises the possibility that patients with very similar patterns of clinical impairment could have displayed markedly different scan findings, depending upon the timing of the imaging. Standardising the timing of imaging may have reduced some of the heterogeneity in this analysis: for example, the true extent of an infarct may have been more easily appreciated had delayed imaging been specified.

4.3.2.4. Expert review of scans

A key requirement for any trial data is that they be collected in a robust and reproducible manner (Wintermark et al., 2008). If scans are to be used to

determine eligibility for participation in a trial, then there may be insufficient time to obtain an expert centralised review by a consultant neuroradiologist. In such cases it may be acceptable to rely upon the interpretation of local stroke physicians, provided that adequate training is provided (Barber et al., 2000). The use of a standardised template may help to improve the reliability of non-expert reports, by drawing the reviewer's attention to each relevant area in turn (Wardlaw et al., 2007, Wardlaw et al., 2010).

If imaging is to be used as a variable in a subsequent analysis, or if advanced imaging techniques are to be used, then reporting by an experienced consultant neuroradiologist is advisable. Centralised review of imaging is recommended, with radiologists blinded to clinical information and treatment allocation (Wintermark et al., 2008). In DARS, the initial grant proposal was for a Research Fellow to code scan findings using written reports from local radiologists. This procedure was later modified to specify centralised review of imaging, since it was recognised that written reports alone were unlikely to be sufficiently detailed to allow accurate coding of findings. However, the additional costs of radiology time had not been included in the original grant application. It was therefore necessary to divert funds from elsewhere in the grant to support this work. As a result, there was sufficient funding to allow scans to be reviewed by JP plus only one consultant neuroradiologist. Consensus reporting by a panel of experts, with formal arbitration mechanisms to settle disagreements between them, might have been more robust. The requirement for centralised reporting of scans also necessitated the establishment of secure procedures for the despatch of scans to CTRU.

4.3.2.5. Data management and quality control

Any systems put in place for receiving images at a centralised reporting facility must be secure, and subject to robust quality control mechanisms (Wintermark et al., 2008). The option of establishing a secure data link between recruiting centres and LTHT was considered, but was deemed impractical given the time taken to approve and establish such links with such a large number of sites. However, the despatch of scans using electronic storage media proved difficult to standardise, and presented several difficulties.

Firstly, a number of sites would only approve transfer of data in encrypted form despite this not being a protocol requirement. Obtaining passwords from sites therefore presented an additional administrative burden. However, the more crucial consideration was that the trial team were unable to implement robust verification procedures for images sent in anonymised form, and were therefore wholly reliant upon quality control procedures at local site to ensure that the correct imaging for the correct patients was downloaded. This had a number of important consequences. Once images had been redacted of patient-identifiable information, they were identifiable only by the participant's trial identification number. There was no way for CTRU to verify that the images had been correctly identified at source. Indeed, a number of cases were noted in which obviously incorrect imaging had been sent. In one case, the same sequence of images for the same patient was copied to two different CDs, each labelled with the trial identification number of a different participant: the error was noticed only because a distinctive necklace the patient was wearing appeared in the scout image. In a second case a CT of the thorax, performed on an earlier occasion for an indication unrelated to the patient's presenting stroke, was sent.

In patients who had had previous neuroimaging studies, the onus was upon the recruiting centre to ensure that only imaging relating to the index stroke (the event for which they were randomised in to the DARS trial) was sent. However, when the dates of image acquisition (as recorded on the images themselves) were compared to the dates of the index stroke, it was noted that in 32 participants the date of imaging *pre-dated* the index stroke. In 24 cases, the recorded scan date preceded the stroke by more than 7 days, with the longest recorded interval being 395 days pre-stroke. At the other extreme, 19 patients had imaging that was apparently conducted >14 days after the index stroke, with the longest time interval being 163 days. The reasons for these discrepancies were impossible to determine. It is possible that, in some cases, the "date of stroke" entered on to the CRF is incorrectly stated, and in fact represents the date on which the patient was transferred from another hospital. It is equally possible that the "first available" imaging may not correspond with the "first imaging performed": in cases where patients were transferred from a regional "hyperacute" stroke unit and subsequently

recruited on a local “acute” stroke unit, imaging performed at the hyperacute centre may not have been made available at the time of transfer. Finally, it is possible that recruiting centres interpreted the instructions to anonymise the images as requiring that all clinical information be redacted from the scans, and that a false date of image acquisition was substituted. Ultimately, this situation illustrates the need for robust quality control procedures in despatching and receiving imaging.

4.3.2.6. Storage and archiving of images

All data collected as part of a clinical trial must be archived securely. The archiving of scan images presented a particular challenge. Although LTHT has a secure Picture Archiving and Communications System, storage capacity for this server was limited and the cost of administrator time to upload images to it could not be met. The trial team were thus obliged to archive images on CDs: a format which may degrade over time. It was therefore agreed that completed radiology CRFs would also be archived as source data, to ensure that a contemporaneous paper-based record of scan findings was retained. However, if archiving of original images is required then the costs of doing so and of preserving the data in a usable format must be included in the initial grant application.

Part 4.4 Directions for future work.

4.4.1. The need for prognostic modelling in stroke

The need for robust models to predict specific rehabilitation outcomes after stroke has, arguably, never been greater. In recent years, many major advances have been made in acute stroke care. In some cases, such as the introduction of thrombolysis for acute ischaemic stroke, the development of simple and reliable prognostic models has been crucial in differentiating those patients who might benefit from the intervention from those at risk of sustaining harm as a result of treatment (Barber et al., 2000).

Although the benefits of acute interventions in stroke care are evaluated largely in terms of preventing death or dependency (Stroke Unit Trialists Collaboration, 2013, Early Supported Discharge Trialists, 2009, Wardlaw et

al., 2014b), such broad outcomes provide very little useful information about specific functional outcomes such as recovery of walking (Preston et al., 2011, Kwakkel and Kollen, 2013). And yet these are precisely the outcomes that are of greatest use when planning a rehabilitation intervention.

4.4.1.1. The possible uses of prognostic models in clinical practice

The development of models to predict specific rehabilitation outcomes may be useful in several respects. Firstly, a model that is able to predict reliably who might walk again could allow more accurate prognostic information to be provided to patients and their families (Craig et al., 2011). This could in turn facilitate goal-setting and discharge planning (Kwakkel and Kollen, 2013). However, in order to be truly useful in this setting any such model would need to offer predictive accuracy superior to that of the clinical judgement of a skilled multi-disciplinary team (Kwakkel et al., 2000), include predictor variables that can easily be collected (Altman and Royston, 2000), and have a scoring system that is simple to apply (Moons et al., 2015).

A second important role for prognostic modelling in rehabilitation practice might be in tailoring the intervention that is delivered to the needs of individual patients. This might sound like an unusual, even counter-intuitive, point to assert: after all, rehabilitation by its nature consists of a “personalised” package of interventions aimed at achieving goals that are specific to an individual (Gutenbrunner et al., 2006). And yet, considerable uncertainty remains about what rehabilitation strategies are most appropriate for individual patients. A good example of such a controversy in stroke care is in deciding the intensity of rehabilitation that should be offered. Although more intensive rehabilitation is associated with shorter lengths of stay in hospital (Slade et al., 2002), recent evidence from a major randomised controlled trial has raised concerns that offering very intensive mobilisation within the first 48 hours of stroke actually results in a *reduction* in the odds of a favourable outcome at three months, and no net benefit in terms of mobility outcome (Bernhardt et al., 2015). This finding is surprising, and merits further consideration. Since the trial enrolled participants with a range of initial clinical impairment from mild to severe (Bernhardt et al., 2015), it is plausible that those with mild initial impairment may in fact *benefit* from early intensive

mobilisation, whilst those with more severe impairment may be harmed by such an approach. A model to predict who might recover the ability to walk again could thus allow the intensity of rehabilitation that is offered to be tailored to the patient's prior probability of regaining independent mobility. Thus, those with mild initial impairment and a high prior probability of walking again could be targeted for early intensive rehabilitation; those with an intermediate probability of walking and moderate impairment may still benefit from active rehabilitation, but at a lesser intensity; those with a low probability of walking (and the most profound impairment) may benefit from a focus upon compensatory rehabilitation strategies, with the aim of maximising participation by modification of contextual factors. An ability to predict future recovery reliably may thus allow the delivery of rehabilitation interventions to be tailored to a patient's prior probability of attaining a specified level of functioning (Steyerberg et al., 2013, Dorresteijn et al., 2011, Kwakkel and Kollen, 2013) thereby allowing resources to be used more efficiently (Steyerberg et al., 2013).

4.4.1.2. Possible uses of prognostic modelling in rehabilitation research

Perhaps the greatest potential of prognostic modelling in rehabilitation lies not in clinical practice, but in research. A number of promising novel rehabilitation strategies are now emerging, such as the use of robotic interventions to augment the intensity and frequency of task-specific practice that can be delivered by therapists (Kwakkel et al., 2008, Sivan et al., 2014), the use of electrical stimulation of peripheral nerves to enhance plasticity in the central nervous system (Dimyan and Cohen, 2011, O'Connor et al., 2014), or direct stimulation of the motor cortex either by magnets or by implanted electrodes (Dimyan and Cohen, 2011). The benefits of these strategies remain as yet unproven, and large-scale trials will ultimately be necessary to establish whether or not they have a place in clinical practice. Having models to predict rehabilitation outcomes such as walking ability might allow patients to be selected for these trials based on their prior probability of benefitting from the intervention. Prognostic modelling may also play a role in the design of future stroke research trials.

As was alluded to in Chapter 1, randomised controlled trials enrol a heterogeneous sample of participants (Dorresteijn et al., 2011). Inherent in the interpretation of their results are the assumptions that each participant has an equal probability of benefitting from the intervention, and that the magnitude of any effect size observed is, on average, uniform across the entire sample (Dorresteijn et al., 2011). End-points are often dichotomised into favourable versus unfavourable outcomes, and are applied to the trial population as a whole. Doing so assumes that the population to which the intervention is applied is homogeneous, and that the direction and magnitude of any treatment effect will be roughly the same for the whole group. And yet, sub-groups within a sample may vary considerably: both in their baseline clinical characteristics and in their response to the intervention (Dorresteijn et al., 2011). Presenting results as an “average” effect size may therefore fail to account for clinically important differences in trial participants (Young et al., 2005). For example, those with profound disability may consider even a modest improvement to be a “good” outcome, whereas those with minor impairment may be dissatisfied with anything less than complete recovery (Young et al., 2005). Furthermore, the existence of significant heterogeneity in a trial sample may compromise statistical power to detect the primary outcome (Makin et al., 2013). Sample size calculations are generally based upon an assumption about the baseline probability of patients developing the outcome of interest: if a sample contains a significant proportion of patients at one extreme of the prognostic spectrum (very good or very poor), then the statistical power to detect a treatment effect will be attenuated (Makin et al., 2013). Perhaps a first step towards improving the design of future rehabilitation trials would be to account for the heterogeneity of trial participants (Dorresteijn et al., 2011). Prognostic models to predict key outcomes of interest (Steyerberg et al., 2013) such as mobility may be used to recruit into trials only those patients who stand the best chance of benefitting from an intervention (Makin et al., 2013). Doing so may reduce the sample size required, but at the expense of decreased recruitment and prolongation of the study (Makin et al., 2013). An alternative approach would be to pre-specify a range of possible trial outcomes, and assign patients differential end-points dependent upon their prior probability of achieving a

specified level of function (Makin et al., 2013). This approach, though not presently used widely in stroke rehabilitation, has the potential to reduce the sample sizes required whilst not impacting upon recruitment (Makin et al., 2013). It is, of course, reliant upon the development of reliable prognostic models.

Models to predict outcomes such as walking may thus have a considerable impact upon both stroke rehabilitation in clinical practice, and rehabilitation research. However, this potential will only be realised if the models developed are able to reliably predict the outcome of interest with a high sensitivity and specificity. A striking, and perhaps important, finding of this study is that a combination of simple clinical measures of impairment that one might expect to find in use on UK stroke units actually explained comparatively little variance in outcome, and correctly classified only marginally more patients as walking/not walking at T₁ and T₂ than prior probability alone. This raises two important questions. Firstly: what might the implications be for the way in which teams arrive at a rehabilitation prognosis? Secondly: what variables might prove to be better predictors of motor outcome?

4.4.1.3. Is statistical modelling any more reliable than “team opinion” in predicting recovery?

As discussed above, if a model is to be useful in clinical practice then it must offer a more reliable prediction of outcome than the opinion of an experienced multi-disciplinary team. It has been suggested previously that models to predict walking ability after stroke offer no greater predictive value than a therapist's assessment (Kwakkel et al., 2000). And yet, the same study revealed that physiotherapists accurately predicted a patient's future walking ability only 48% of the time (Kwakkel et al., 2000). Around 26% of predictions made at five weeks about a patient's walking ability at six months were over-optimistic, whilst 26% underestimated the final level of mobility achieved (Kwakkel et al., 2000). Recovery of arm function was more reliably predicted, with 63.6% of physiotherapists and 59.1% of occupational therapists accurately predicting six-month prognosis on the basis of an assessment at five weeks (Kwakkel et al., 2000). However, this apparent improvement in reliability relative to prediction of walking ability is most likely due to the well-

recognised tendency for arm function to recover incompletely (Kwakkel et al., 2000). It therefore appears that neither existing prognostic models nor clinical judgement alone are reliable predictors of future prognosis. One possible explanation for this finding is that both prognostic models and a professional's clinical judgement depend, either explicitly or implicitly, upon a combination of measures of physical impairment that themselves explain comparatively little variance in outcome. There is therefore a need to understand better the basis upon which a team's prognostic decisions are made and how reliable those decisions might be in predicting key outcomes.

There is also a need to identify variables that might explain a greater proportion of variance in outcome. It is possible that measures focusing upon very specific impairments such as balance (van de Port et al., 2006a) or of executive function might yield more accurate predictions than scales that measure the broader constructs of "mobility" and "cognition". However, the identification of constructs that might be predictive of key rehabilitation outcomes is merely a first step towards model development. Perhaps a more fundamental and urgent imperative is for robust scales with which to measure these constructs.

4.4.1.4. What is the future role of brain imaging in predicting rehabilitation outcomes?

This study aimed to evaluate whether variables derived from early CT imaging made a contribution to models predicting walking ability beyond that achieved by measures of clinical impairment alone. The choice of CT imaging was a pragmatic one, since this modality is already routinely used to guide a patient's acute management, is cheap and quick to acquire in comparison with MRI modalities, and has fewer contra-indications than MRI imaging. However, CT is an imperfect surrogate for a patient's functional impairment (Kobayashi et al., 2009). Nor does structural imaging provide useful information about how different brain structures *interact* to shape recovery (Dayan and Cohen, 2011). It is therefore perhaps unsurprising that CT findings do not contribute to the prediction of a specific functional outcome (recovery of walking ability).

Functional MRI provides a dynamic picture of brain metabolic activity that is assumed to reflect activation of brain structures (Arthurs and Boniface, 2002). This may offer some insight in to the biological basis of recovery from stroke, by elucidating how network interactions between spatially-distributed brain structures facilitate learning. However, although much work has already been done to understand the processes that underpin motor learning (Doyon et al., 2009, Hikosaka et al., 2002, Penhune and Steele, 2012), theoretical models developed in healthy participants cannot be assumed to translate directly to those recovering from stroke. At present, functional imaging remains largely a research tool, and is not performed routinely in clinical practice.

If this modality is to prove useful in predicting rehabilitation outcomes, then the key questions for research are perhaps: what brain structures interact to facilitate motor learning in healthy individuals; how do these processes differ in those with structural brain injury as a consequence of stroke; do specific patterns of disruption of network interactions between brain structures correlate with limitation of motor recovery; does functional MRI imaging allow a more accurate prediction of prognosis to be made than would be obtained using clinical judgement alone or models based on robust measures of clinical impairment; and what are the costs versus benefits of using this technique more widely in clinical practice.

4.4.2. The need for robust outcome measurement in stroke research

4.4.2.1. Why is outcome measurement important?

The need for robust outcome measurement in stroke is urgent and pressing: both for research, prognostic modelling, and for clinical practice. Rehabilitation interventions are, by their nature, complex (Craig et al., 2008): they depend upon a team of professionals working together to restore physical function and modify a patient's social and environmental context, with the overall aim of mitigating disability (a construct defined in terms of limitation of activities and restriction of participation) (Gutenbrunner et al., 2006). However, the efficacy of particular elements of a rehabilitation programme may be hard to discern. Randomised controlled trials, long regarded as the

gold standard of scientific evidence, are by their nature a reductive process that seek to isolate and examine one component part of a wider package of measures (Gutenbrunner et al., 2006). Any genuine effect of such interventions may be small, and difficult to differentiate from that of the wider rehabilitation programme. The ability to detect reliably small changes in function is therefore crucial to evaluating novel rehabilitation interventions.

The development of new rehabilitation interventions will also depend upon an understanding of how different physical impairments may interact to influence recovery. Utilising modelling techniques to explore the process of recovery may allow the identification of key variables that strongly influence outcome. This might, in turn, provide information that will be of value in planning a rehabilitation programme, and also in the development of new interventions that specifically target key components of recovery. Models that are developed in order to predict prognosis may not only enable more reliable information to be provided to patients and families, but could also allow the stratification of patients in to research trials or the differential selection of end-points for trials based upon a patient's prior probability of attaining a specified level of functioning. However, the interpretation of such models depends upon knowing the magnitude of true clinical change that such models are able to predict.

In clinical practice, the ability to detect true clinical change is important not only in planning a package of rehabilitation interventions for individual patients, but also for the commissioning of services. Rehabilitation is known to be effective in reducing physical dependency after stroke (O'Connor et al., 2011). However, demonstrating the cost-effectiveness of such services depends upon the ability to detect change in a patient's level of functioning. Robust outcome measurement is, once again, critical for this process. For example, some scales such as the Barthel Index are known to be insensitive to true clinical change in patients with the highest and lowest scores (Quinn et al., 2011). Such "floor" and "ceiling" effects mean that, in the most- and least-disabled patients, it is possible for clinically-significant changes in performance to occur despite minimal or no change on the BI score (Quinn et al., 2011). The ability to evaluate the success of a rehabilitation intervention,

and its cost-effectiveness, may therefore be compromised if outcome measures are used that do not appropriately target the population in question.

The ability to detect small changes in function, both in rehabilitation trials and in clinical practice, depends upon the use of sensitive, psychometrically-robust outcome measures that are appropriately applied, analysed, and reported. Here, unfortunately, much work remains to be done.

4.4.2.2. The current status of outcome measurement in stroke research

A bewildering array of outcome measures exists, measuring a wide variety of constructs. A recent systematic review found 47 different outcome measures used in 126 stroke trials (Quinn et al., 2009b). The numbers of measures in use for specific impairments (upper limb function, aphasia, and cognition) is larger still (Ali et al., 2013). Despite several calls for standardisation of outcome measurement (Langhorne et al., 2009, Lees et al., 2012), no consensus has been reached on which measures are the most appropriate and standards for a minimum data-set for rehabilitation trials have not been agreed (Lees et al., 2012, Ali et al., 2013). In 2012 the European Stroke Outcomes Working Group (Lees et al., 2012) suggested that the mRS be adopted as the primary outcome measure in stroke trials. Unfortunately this measure is far too broad in scope to be of use in rehabilitation trials. The wide variety of outcome measures used in stroke trials greatly impedes meta-analysis and comparison of trial data (Ali et al., 2013). Worse still, the scales themselves are often arbitrarily changed with scant regard for the impact of those changes on their properties.

A good example of the confusion that surrounds outcome measurement in stroke is the Barthel Index (BI). This is a measure of global physical functioning, covering a number of activities. Since its first publication in 1965 (Mahoney and Barthel, 1965), it has become one of the most widely-used outcome measures in clinical practice (Quinn et al., 2011). Although designed neither for clinical trials nor specifically for stroke, its adaption in to stroke research has been enthusiastic with only the mRS proving more ubiquitous (Quinn et al., 2009b). The original version described ten items: feeding, dressing, grooming, bathing, bladder and bowel continence, toileting,

transfers, mobility, and stair use (Quinn et al., 2011). Each item is scored according to a weighted scale: total scores are out of 100, with higher scores denoting greater independence (Quinn et al., 2011). The BI has been modified several time, including changes to the weighting of items, or alterations to the definitions of items (Quinn et al., 2011). There are presently at least four versions which return summary scores out of 100, all of which are referred to in the literature as the “Barthel Index” (Quinn et al., 2011). A number of expanded versions have been suggested which incorporate items not included in the original score (such as tracheostomy management or cognitive function), apparently in response to local needs (Quinn et al., 2011). Attempts have also been made to derive “short-form” indices, by removing all but the most discriminating items. Interestingly, the items chosen for inclusion or removal seem to vary (Quinn et al., 2011). The construct validity and clinical utility of these “modified” versions cannot be assumed (Quinn et al., 2011). Versions of the BI have been developed for administration by face-to-face interview, direct observation, by telephone, or by self-report (including by post).(Quinn et al., 2011, Gompertz et al., 1994) There is no consensus about which version is the most appropriate for use in stroke trials (Quinn et al., 2011).

4.4.2.3. What are the key properties of a rehabilitation outcome measure?

As we have discussed above rehabilitation outcome measures seek to define a construct, to develop a list of items that range in order from “easiest” to “hardest” or “less” to “more”, and then to determine the position of a patient along the metric according to the location of the items that they endorse (Hobart and Cano, 2009). If the output of such scales is represented as a summary score, then it cannot be assumed that such scores provide *interval level measurement*.

Furthermore, unless all items contained within a scale measure a common underlying construct, the legitimacy of deriving a summary score from them may be called in to question. This property is termed *unidimensionality*. Attempts to measure rehabilitation outcomes may be further compounded by the inter-dependence of the scale used, and the characteristics of the people

in the sample it seeks to measure (Hobart and Cano, 2009). The apparent performance of a group of patients may change depending upon the scale that is used to measure them. For example, when measuring physical functioning in Multiple Sclerosis, a sample of moderately-disabled patients will return a high score on the BI, a mid-range score on the Multiple Sclerosis Impact Scale physical functioning subscale, and a low score on the Medical Outcomes Study 36-item Short Form Health Survey physical functioning dimension (Hobart and Cano, 2009). The apparent performance of an individual patient on any given scale will also depend upon the characteristics of the other patients in the sample. If measured within a sample containing predominantly severely-disabled patients, a person with moderate disability will score within the higher percentiles for physical functioning; conversely the same patient, measured using the same scales but this time amongst a mildly-disabled sample, will score within the lower percentiles (Hobart and Cano, 2009). Put simply, the apparent performance of an individual depends upon the properties of the scale used to measure them, and on the characteristics of the sample of which they are part (Hobart and Cano, 2009). This poses a problem for outcome measurement in rehabilitation practice: unless a means can be found to isolate genuine change in the performance of individual patients from variation in the characteristics of a scale, one cannot be confident that true clinical change has been measured. Ideally, the performance of an individual should be independent of the scale used to measure them, and the properties of a scale should not vary depending upon the characteristics of the sample it is used to measure. This concept is termed *invariance*. Interval-level measurement, unidimensionality, and invariance are sometimes referred to as the key tenets of measurement (Hobart and Cano, 2009). Fortunately, it is possible to evaluate scales systematically to establish whether they meet these requirements.

4.4.2.4. Methods for the formal evaluation of outcome measures

The science of measuring rehabilitation outcomes, a field termed *psychometrics*, has developed enormously in recent years. Traditionally, the validation of a new outcome measure would focus upon three domains: reliability (the extent to which measurements are influenced by random error);

validity (whether a scale actually measures the property it purports to measure); and responsiveness (whether a scale reliably detects clinical change) (Hobart and Cano, 2009). However, classical methods for evaluating scales offer no means of determining whether an outcome measure provides interval-level measurement or displays invariance (Hobart and Cano, 2009). They also depend upon the scale satisfying several assumptions which in fact cannot be mathematically tested, and are therefore considered to be met for most data-sets (Hobart and Cano, 2009).

More recently, a range of rigorous methods have been developed for the evaluation of outcome measures. Amongst them are a series of models proposed by Georg Rasch, and Item Response Theory (Hobart and Cano, 2009). A detailed discussion of these models is beyond the scope of this thesis: interested readers are referred to a comprehensive monograph by Hobart and Cano (2009). Suffice to say that Rasch proposed a set of mathematically-testable hypotheses which, if satisfied, indicate that a scale can be assumed to fulfil the key tenets of measurement (Hobart and Cano, 2009). Unlike Item Response Theory, which seeks to derive models from observed data, Rasch analysis gives primacy to the model itself (Hobart and Cano, 2009). The data obtained from a rating scale are fitted to the Rasch model, and the observed fit is then compared to what would be expected if the scale fulfilled the criteria for optimum measurement (Hobart and Cano, 2009). Any misfit between observed and expected values prompts not a re-evaluation of the model, but a further examination of the data to determine why an item (or set of items) is not performing as expected (Hobart and Cano, 2009).

Rasch analysis allows for the systematic examination of scales to establish if they fulfil the criteria of unidimensionality and invariance. It also allows the appropriateness of the order in which items are arranged in a scale to be evaluated, and for items that do not fit the model to be identified (Tennant and Conaghan, 2007). Bias that arises as a result of differences in the response to individual items between different patient groups within the sample (*differential item functioning*) can also be evaluated (Tennant and Conaghan, 2007). Crucially, the location of items on an ordinal scale can be mapped to a

logarithmic scale: thereby providing true interval-level measurement (Tennant and Conaghan, 2007).

4.4.2.5. The Virtual International Stroke Trials Archive: an opportunity to evaluate outcome measures used in stroke?

Rasch analysis may be used to build scales de novo, or to evaluate the properties of existing scales (Tennant and Conaghan, 2007). The systematic evaluation of rehabilitation scales that are commonly used in stroke would be of enormous value, since having available a battery of validated outcome measures that are proven to fulfil the key tenets of measurement would provide a solid foundation from which research and modelling could then proceed. Such work would, of course, rely upon the existence of a large bank of data derived in stroke patients and covering a variety of relevant outcome measures. The VISTA archive may be such a resource.

VISTA was set up in 2007 to bring together data from major clinical trials, in the hope that doing so would facilitate exploratory analyses of existing data-sets. (Ali et al., 2007) By 2013, its rehabilitation trials offshoot (VISTA-Rehab) contained data-sets from 38 trials, enrolling a total of 10,244 participants (Ali et al., 2013). A total of 44 different outcome measures are included, encompassing both impairment and activities/ participation levels of the ICF (Ali et al., 2013). Unfortunately, the promise of this resource has yet to be realised. Differences in characteristics of the samples from which these measures were recorded confounds any meta-analysis of these data (Ali et al., 2013). If, however, it could be established that the outcome measures contained within VISTA-Rehab display invariance, then exploratory analyses and statistical modelling using pooled data from the VISTA-Rehab bank could proceed with confidence. Furthermore, if the outcome measures contained within VISTA-Rehab could be proven to provide interval-level measurement, then the result would be a large bank of scales for which the magnitude of change can be measured quantitatively. This would be an enormously powerful resource for the design of future rehabilitation trials, since the appropriate measure could be selected from a battery of scales with known psychometric properties and proven validity. The ability to provide interval-level measurement would also allow more accurate power calculations to be

made, since the number of patients enrolled could be tailored to the magnitude of true change anticipated to result from an intervention. This may ultimately reduce the cost of clinical trials: either by preventing the wastage of resources on under-powered trials that are likely to return inconclusive results, or in some cases by allowing the reduction of sample sizes (thereby minimising the time and costs of recruitment). The systematic application of Rasch methods to scales in the VISTA-Rehab bank and the DARS data-set therefore offers a means to establish whether the most commonly used stroke outcome measures fulfil the key tenets of measurement. This work could be completed using existing data; yet its possible impact is substantial.

Part 4.5 Concluding remarks

4.5.1. Potential future uses of outputs from this Thesis

4.5.1.1. Reflections on a complex trial

At the time of its inception DARS was the largest-ever multi-centre randomised controlled trial of a pharmacological intervention to enhance physical recovery after stroke. Delivery of the DARS intervention (a single dose of co-careldopa 45min to 1hour before the start of each therapy session) seems straightforward when set down as a short paragraph in the trial protocol. As Bipin Bhakta himself, ever the optimist, might have said: “How hard can it be?”.

In reality, ensuring the reliable delivery of this ostensibly-simple intervention turned out to be a major challenge for the trials team. There is no absolute definition of what constitutes a “complex” intervention, but the Medical Research Council have suggested that the characteristics of a such an intervention include: a large number of interacting components within the experimental and control groups; the number and difficulty of behaviours required by those delivering or receiving the intervention; and the degree of tailoring or flexibility of the intervention permitted (Craig et al., 2008). In the case of DARS, ensuring that the medication was delivered in accordance with the trial protocol required an unprecedented degree of liaison and interaction between ward nurses and therapy teams (for in-patients), or between

community therapy teams and patients or carers (for those discharged from hospital before their course of treatment was completed).

This required an unprecedented level of training in trial procedures for hospital and community staff who would be involved in delivering the intervention. This was conducted by me and the DARS trial monitor, Lorna Barnard, at a series of face-to-face site initiation visits. The provision of face-to-face training in trial procedures was felt to be the only way to ensure that staff were trained in trial procedures to the standard required for them to deliver the intervention per protocol requirements. This was, however, a costly and time-consuming exercise that was only compounded by the number of centres that ultimately collaborated with DARS.

It was initially intended that the trial would be conducted across a small number of centres within the Yorkshire area. However, when feasibility assessments were requested from potential recruiting centres in the early stages of trial setup, it became apparent that anticipated per-centre monthly recruitment was lower than expected. This led to an initial expansion in the number of centres to 20. Once the trial opened to recruitment, even the modest estimates of 1-2 patients recruited per centre each month were found to be optimistic. This necessitated a further expansion to a final total of over 50 centres. Ultimately, the initial recruitment target was met and exceeded; but maintaining currency in trial procedures for staff at centres that rarely recruited patients was a significant challenge. This was compounded by the tendency of junior therapists to rotate to different posts every 4-6 months.

DARS was an ambitious trial, which overcame a number of significant challenges to deliver a robust answer to an important clinical question. In this respect, it stands not only as a lasting tribute to Bipin, but also as a benchmark for other complex rehabilitation trials. Although the process of setting up and running the DARS trial has not been discussed in detail previously in this Thesis, it is clear that a reflective paper setting out the challenges that were faced by DARS and how they were overcome has much to offer the design of future complex rehabilitation trials.

4.5.1.2. Incorporating analysis of imaging in to a trial

It is not unusual for randomised controlled trials to include an analysis of imaging. This might be as a direct inclusion/exclusion criterion (as for trials of thrombolysis in stroke), or as an outcome measure in itself (for example measuring tumour regression in cancer chemotherapy trials). In DARS, the incorporation of an analysis of brain imaging in to the protocol stemmed initially from a desire to explore the mechanism by which co-careldopa might influence recovery. Although centres of excellence in brain imaging research do exist, there is surprisingly little published literature to guide non-radiologists looking to include imaging analysis in a trial. When designing the DARS protocol, several important radiological considerations were therefore overlooked. For example, the cost of centrally collating imaging and subsequent expert review by experienced neuroradiologists was not incorporated in to the original grant application. Due to the limited funding available, reporting was performed by a single expert plus JP, not by a panel of experts. Information that might have been useful to the radiologists when interpreting scans (for example, the laterality of stroke symptoms) was not collected, as trial paperwork and procedures had been largely finalised by the time the need to do so was identified. Nor were quality control procedures for ensuring that the correct scans were sent to CTRU as robust as they ought to have been. A paper laying out the basic considerations when including imaging analysis in a grant proposal would be useful in the design of future trials.

4.5.1.3. Rasch analysis of outcome measures from the DARS data-set

This Thesis sought to develop a series of models to predict walking ability at up to six months after stroke. However, any such models must be founded upon the rock of robust outcome measurement, rather than the shifting sands of ordinal scales. An enormous variety of outcome measures are currently used in stroke medicine, few of which have been validated using modern psychometric techniques. The manner in which these measures are then analysed and interpreted, for example in the derivation of mean scores or in quoting changes in scores over time, is questionable. This situation is an

impediment to the design, interpretation, and meta-analysis of high quality rehabilitation trials.

There is a need in stroke research and in clinical practice for robust outcome measures that fulfil the key tenets of measurement: interval-level measurement, unidimensionality, and invariance. The use of scales that are proven to be interval-level would allow more efficient linear regression models to be fitted: and their output could legitimately be expressed as the change in the value of an outcome variable for each unit change in a predictor variable. Interval-level outcome measurement will also allow more reliable estimation of effect sizes: when fitting predictor variables statistical models, and when measuring the impact of rehabilitation interventions in clinical practice.

The DARS trial acquired outcome measures covering a variety of impairments and activity limitations at four time points (baseline, eight weeks, six months, and one year after stroke). The systematic application of rigorous psychometric techniques such as Rasch analysis to such a rich data-set would allow a series of outcome measures to be made available that are proven to fulfil the key tenets of measurement.

4.5.2. The implications of this Thesis for clinical practice and research

4.5.2.1. Implications for clinical practice

Although brain imaging findings may have a role in predicting broad outcomes such as death or dependency following stroke, the models presented above cast doubt upon the ability of CT imaging to predict more specific rehabilitation outcomes such as ability to walk independently; at least as far as ischaemic stroke is concerned (models 1 and 2). The models presented for ICH (models 3 and 4) are derived in a smaller sample, with commensurately wide confidence intervals for some imaging variables. By contrast models 5 and 6, based on clinical impairment assessed within the first few weeks after stroke, correctly classified up to 70.3% at T_1 and 69.0% of patients at T_2 . This implies that, in clinical practice, a reasonably accurate estimation of prognosis may be made based on initial clinical impairment alone. Although plain CT imaging remains crucial in guiding the acute management of stroke patients, it cannot

at present be deemed useful for rehabilitation prognostication. This finding is of potential significance in low-resource settings, in which brain imaging might not be readily available.

Advanced imaging techniques such as fMRI have been used to explore the interactions that take place between brain structures during the learning process. However so-called “functional” imaging actually detects an increase in blood oxygen diffusion in to tissues, a finding that is purported to correlate with an increase in neuronal metabolic activity. The actual “function” of those brain areas and how they actually interact to shape learning remains a matter for inference and speculation. Nor is it necessarily legitimate to assume that learning processes in a healthy volunteer under experimental conditions are analogous to those of a stroke patient participating in a rehabilitation programme. Functional imaging therefore remains at present primarily a research tool with little role in routine rehabilitation practice.

The models presented here would certainly require validation in an independent sample prior to clinical use. However, models that are to be used in clinical practice must also be easy to apply, and deliver an output that is readily interpretable by staff. Although binary logistic regression modelling gives some indication of the relative importance of each predictor variable (as measured by the percentage of variance that each predictor explains and the change in odds of the outcome of interest for each one-point change in a predictor variable), the output that they deliver is neither intuitive for a clinician nor easily applicable to an individual patient. A “decision tree”, in which each node is a binary choice and the output is the odds of walking independently at T_1 and T_2 , might allow outcome predictions to be made in a more readily interpretable manner.

4.5.2.2. Implications for research

Perhaps the greatest potential of the models presented here is in rehabilitation research. A model that is able to correctly classify around 70% of patients as able or unable to walk at T_1 and T_2 with a sensitivity of 55.8-72.0% at T_1 and 62.8-88.9% at T_2 might allow patients recruited in to future rehabilitation research trials to be assigned different end-points at the time of

randomisation, based on their prior probability of walking again. There are several advantages to such an approach. Firstly, the sample size required to detect a treatment effect could be reduced, thereby reducing the cost of setting up and running trials. Secondly, accounting for differences in anticipated prognosis would allow researchers to assign outcome measures that are of practical importance to patients. For example, a patient who is unlikely to return to walking might consider achieving independent sitting balance to be an important goal, whereas for those patients who are expected to be able to walk 10m or more an outcome centred around higher levels of mobility (such as walking outdoors, climbing stairs, or running) might be pre-specified. Ultimately such an approach would allow trials to adopt a range of outcome measures that are of direct relevance to pre-specified sub-groups of patients. This approach may also allow detection of more subtle treatment effects than would be apparent if results are analysed based on a single dichotomous outcome.

A further implication of this Thesis is that it calls in to question how useful brain imaging is in predicting rehabilitation potential. Although models incorporating imaging variables are of value in predicting mortality or broad categories such as “independent”/“dead or dependent”, the findings of routinely-acquired CT imaging appear to add nothing beyond an assessment of clinical impairment when predicting more nuanced rehabilitation outcomes such as walking ability. Is it reasonable to continue to fund studies aiming to predict rehabilitation outcomes using imaging variables, when clinical impairment alone appears to be a more reliable predictor of rehabilitation potential?

References

- Department of Health, 2010. Progress in improving stroke care. *London, The Stationary Office*. Available online at <http://www.nao.org.uk/publications/0910/stroke.aspx>.
- ABEN, I., DENOLLET, J., LOUSBERG, R., VERHEY, F., WOJCIECHOWSKI, F. & HONIG, A. 2002. Personality and vulnerability to depression in stroke patients: a 1-year prospective follow-up study. *Stroke*, 33, 2391-5.
- ACLER, M., FIASCHI, A. & MANGANOTTI, P. 2009a. Long-term levodopa administration in chronic stroke patients. A clinical and neurophysiologic single-blind placebo-controlled cross-over pilot study. *Restorative Neurology and Neuroscience*, 27, 277-283.
- ACLER, M., ROBOL, E., FIASCHI, A. & MANGANOTTI, P. 2009b. A double blind placebo RCT to investigate the effects of serotonergic modulation on brain excitability and motor recovery in stroke patients. *Journal of Neurology*, 256, 1152-1158.
- AGUILAR, M. I., HART, R. & PEARCE, L. A. 2007. Oral anticoagulants versus antiplatelet therapy for preventing stroke in patients with non-valvular atrial fibrillation and no history of stroke or transient ischemic attacks. *Cochrane Database of Systematic Reviews*, CD006186.
- AHO, K., HARMSSEN, P., HATANO, S., MARQUARDSEN, J., SMIRNOV, V. E. & STRASSER, T. 1980. Cerebrovascular disease in the community: results of a WHO collaborative study. *Bulletin of the World Health Organization*, 58, 113-30.
- ALBERT-WEISENBERGER, C., SIREN, A.-L. & KLEINSCHNITZ, C. 2013. Ischemic stroke and traumatic brain injury: the role of the kallikrein-kinin system. *Progress in Neurobiology*, 101-102, 65-82.
- ALEXANDER, G. E., DELONG, M. R. & STRICK, P. L. 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, 9, 357-81.
- ALI, M., BATH, P. M. W., CURRAM, J., DAVIS, S. M., DIENER, H.-C., DONNAN, G. A., FISHER, M., GREGSON, B. A., GROTTA, J., HACKE, W., HENNERICI, M. G., HOMMEL, M., KASTE, M., MARLER, J. R., SACCO, R. L., TEAL, P., WAHLGREN, N.-G., WARACH, S., WEIR, C. J. & LEES, K. R. 2007. The Virtual International Stroke Trials Archive. *Stroke*, 38, 1905-10.
- ALI, M., ENGLISH, C., BERNHARDT, J., SUNNERHAGEN, K. S., BRADY, M. & COLLABORATION, V. I.-R. 2013. More outcomes than trials: a call for consistent data collection across stroke rehabilitation trials. *International Journal of Stroke*, 8, 18-24.
- ALTMAN, D. G. & ROYSTON, P. 2000. What do we mean by validating a prognostic model? *Statistics in Medicine*, 19, 453-73.
- AMARENCO, P., BOGOUSLAVSKY, J., CALLAHAN, A., 3RD, GOLDSTEIN, L. B., HENNERICI, M., RUDOLPH, A. E., SILLESEN, H., SIMUNOVIC, L., SZAREK, M., WELCH, K. M., ZIVIN, J. A. & STROKE PREVENTION BY AGGRESSIVE REDUCTION IN CHOLESTEROL LEVELS, I. 2006. High-dose atorvastatin after stroke or transient ischemic attack. *New England Journal of Medicine*, 355, 549-59.

ANDERSON, C. S., HEELEY, E., HUANG, Y., WANG, J., STAPF, C., DELCOURT, C., LINDLEY, R., ROBINSON, T., LAVADOS, P., NEAL, B., HATA, J., ARIMA, H., PARSONS, M., LI, Y., WANG, J., HERITIER, S., LI, Q., WOODWARD, M., SIMES, R. J., DAVIS, S. M., CHALMERS, J. & INVESTIGATORS, I. 2013. Rapid blood-pressure lowering in patients with acute intracerebral hemorrhage. *New England Journal of Medicine*, 368, 2355-65.

ANDERSON, C. S., HUANG, Y., ARIMA, H., HEELEY, E., SKULINA, C., PARSONS, M. W., PENG, B., LI, Q., SU, S., TAO, Q. L., LI, Y. C., JIANG, J. D., TAI, L. W., ZHANG, J. L., XU, E., CHENG, Y., MORGENSTERN, L. B., CHALMERS, J., WANG, J. G. & INVESTIGATORS, I. 2010. Effects of early intensive blood pressure-lowering treatment on the growth of hematoma and perihematomal edema in acute intracerebral hemorrhage: the Intensive Blood Pressure Reduction in Acute Cerebral Haemorrhage Trial (INTERACT). *Stroke*, 41, 307-12.

ANONYMOUS 1978. A randomized trial of aspirin and sulfinpyrazone in threatened stroke. The Canadian Cooperative Study Group. *New England Journal of Medicine*, 299, 53-9.

ANONYMOUS 1993. Secondary prevention in non-rheumatic atrial fibrillation after transient ischaemic attack or minor stroke. EAFT (European Atrial Fibrillation Trial) Study Group. *Lancet*, 342, 1255-62.

ANONYMOUS 1995. Tissue plasminogen activator for acute ischemic stroke. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. *New England Journal of Medicine*, 333, 1581-7.

ANONYMOUS 1996. Thrombolytic therapy with streptokinase in acute ischemic stroke. The Multicenter Acute Stroke Trial--Europe Study Group. *New England Journal of Medicine*, 335, 145-50.

APPELROS, P. & VIITANEN, M. 2004. Prevalence and predictors of depression at one year in a Swedish population-based cohort with first-ever stroke. *Journal of Stroke & Cerebrovascular Diseases*, 13, 52-7.

ARIESEN, M. J., ALGRA, A., VAN DER WORP, H. B. & RINKEL, G. J. 2005. Applicability and relevance of models that predict short term outcome after intracerebral haemorrhage. *Journal of Neurology, Neurosurgery & Psychiatry*, 76, 839-44.

ARIMA, H., WANG, J. G., HUANG, Y., HEELEY, E., SKULINA, C., PARSONS, M. W., PENG, B., LI, Q., SU, S., TAO, Q. L., LI, Y. C., JIANG, J. D., TAI, L. W., ZHANG, J. L., XU, E., CHENG, Y., MORGENSTERN, L. B., CHALMERS, J., ANDERSON, C. S. & INVESTIGATORS, I. 2009. Significance of perihematomal edema in acute intracerebral hemorrhage: the INTERACT trial. *Neurology*, 73, 1963-8.

ARTHURS, O. J. & BONIFACE, S. 2002. How well do we understand the neural origins of the fMRI BOLD signal? *Trends in Neurosciences*, 25, 27-31.

AYATA, C. & ROPPER, A. H. 2002. Ischaemic brain oedema. *Journal of Clinical Neuroscience*, 9, 113-24.

BAGG, S., POMBO, A. P. & HOPMAN, W. 2002. Effect of age on functional outcomes after stroke rehabilitation. *Stroke*, 33, 179-85.

BALAMI, J. S. & BUCHAN, A. M. 2012. Complications of intracerebral haemorrhage. *Lancet Neurology*, 11, 101-18.

BAMFORD, J., SANDERCOCK, P., DENNIS, M., BURN, J. & WARLOW, C. 1991. Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet*, 337, 1521-6.

BANKS, J. L. & MAROTTA, C. A. 2007. Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis. *Stroke*, 38, 1091-6.

BARBER, P. A., DEMCHUK, A. M., ZHANG, J. & BUCHAN, A. M. 2000. Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. ASPECTS Study Group. Alberta Stroke Programme Early CT Score.[Erratum appears in *Lancet* 2000 Jun 17;355(9221):2170]. *Lancet*, 355, 1670-4.

BARKER-COLLO, S., FEIGIN, V. L., PARAG, V., LAWES, C. M. & SENIOR, H. 2010. Auckland Stroke Outcomes Study. Part 2: Cognition and functional outcomes 5 years poststroke. *Neurology*, 75, 1608-16.

BEJOT, Y., ABOA-EBOULE, C., DURIER, J., ROUAUD, O., JACQUIN, A., PONA VOY, E., RICHARD, D., MOREAU, T. & GIROUD, M. 2011. Prevalence of early dementia after first-ever stroke: a 24-year population-based study. *Stroke*, 42, 607-12.

BEN ASSAYAG, E., SHENHAR-TSARFATY, S., KORCZYN, A. D., KLIPER, E., HALLEVI, H., SHOPIN, L., AURIEL, E., GILADI, N., MIKE, A., HALEVY, A., WEISS, A., MIRELMAN, A., BORNSTEIN, N. M. & HAUSDORFF, J. M. 2015. Gait measures as predictors of poststroke cognitive function: evidence from the TABASCO study. *Stroke*, 46, 1077-83.

BERENDS, H. I., NIJLANT, J. M., MOVIG, K. L., VAN PUTTEN, M. J., JANNINK, M. J. & IJZERMAN, M. J. 2009. The clinical use of drugs influencing neurotransmitters in the brain to promote motor recovery after stroke; a Cochrane systematic review. *European journal of physical & rehabilitation medicine*, 45, 621-30.

BERKHEMER, O. A., FRANSEN, P. S., BEUMER, D., VAN DEN BERG, L. A., LINGSMA, H. F., YOO, A. J., SCHONEWILLE, W. J., VOS, J. A., NEDERKOORN, P. J., WERMER, M. J., VAN WALDERVEEN, M. A., STAALS, J., HOFMEIJER, J., VAN OOSTAYEN, J. A., LYCKLAMA A NIJEHOLT, G. J., BOITEN, J., BROUWER, P. A., EMMER, B. J., DE BRUIJN, S. F., VAN DIJK, L. C., KAPPELLE, L. J., LO, R. H., VAN DIJK, E. J., DE VRIES, J., DE KORT, P. L., VAN ROOIJ, W. J., VAN DEN BERG, J. S., VAN HASSELT, B. A., AERDEN, L. A., DALLINGA, R. J., VISSER, M. C., BOT, J. C., VROOMEN, P. C., ESHGHI, O., SCHREUDER, T. H., HEIJOER, R. J., KEIZER, K., TIELBEEK, A. V., DEN HERTOOG, H. M., GERRITS, D. G., VAN DEN BERG-VOS, R. M., KARAS, G. B., STEYERBERG, E. W., FLACH, H. Z., MARQUERING, H. A., SPRENGERS, M. E., JENNISKENS, S. F., BEENEN, L. F., VAN DEN BERG, R., KOUDSTAAL, P. J., VAN ZWAM, W. H., ROOS, Y. B., VAN DER LUGT, A., VAN OOSTENBRUGGE, R. J., MAJOIE, C. B., DIPPEL, D. W. & INVESTIGATORS, M. C. 2015. A randomized trial of intraarterial treatment for acute ischemic stroke.[Erratum appears in *N Engl J Med*. 2015 Jan 22;372(4):394]. *New England Journal of Medicine*, 372, 11-20.

BERNHARDT, J., LANGHORNE, P., LINDLEY, R. I., THRIFT, A. G., ELLERY, F., COLLIER, J., CHURILOV, L., MOODIE, M., DEWEY, H., DONNAN, G. & ON BEHALF OF THE AVERT TRIAL COLLABORATION GROUP 2015. Efficacy and safety of very early mobilisation within 24 h of stroke onset (AVERT): a randomised controlled trial.[Erratum appears in *Lancet*. 2015 Jul 4;386(9988):30; PMID: 26169865]. *Lancet*, 386, 46-55.

BERWAERTS, J., DIJKHUIZEN, R. S., ROBB, O. J. & WEBSTER, J. 2000. Prediction of functional outcome and in-hospital mortality after admission with oral anticoagulant-related intracerebral hemorrhage. *Stroke*, 31, 2558-62.

BEWICK, V., CHEEK, L. & BALL, J. 2005. Statistics review 14: Logistic regression. *Critical Care (London, England)*, 9, 112-8.

BHAKTA, B. B., HARTLEY, S., HOLLOWAY, I., COUZENS, J. A., FORD, G. A., MEADS, D., SACKLEY, C. M., WALKER, M. F., RUDDOCK, S. P. & FARRIN, A. J. 2014. The DARS (Dopamine Augmented Rehabilitation in Stroke) trial: protocol for a randomised controlled trial of Co-careldopa treatment in addition to routine NHS occupational and physical therapy after stroke. *Trials [Electronic Resource]*, 15, 316.

BHALLA, A., WANG, Y., RUDD, A. & WOLFE, C. D. 2013. Differences in outcome and predictors between ischemic and intracerebral hemorrhage: the South London Stroke Register. *Stroke*, 44, 2174-81.

BHATTATHIRI, P. S., GREGSON, B., PRASAD, K. S., MENDELOW, A. D. & INVESTIGATORS, S. 2006. Intraventricular hemorrhage and hydrocephalus after spontaneous intracerebral hemorrhage: results from the STICH trial. *Acta Neurochirurgica - Supplement*, 96, 65-8.

BHATTATHIRI, P. S., GREGSON, B., PRASAD, K. S., MITCHELL, P., SOH, C., MITRA, D., GHOLKAR, A. & MENDELOW, A. D. 2003. Reliability assessment of computerized tomography scanning measurements in intracerebral hematoma. *Neurosurgical Focus*, 15, E6.

BOLAM, J. P., HANLEY, J. J., BOOTH, P. A. & BEVAN, M. D. 2000. Synaptic organisation of the basal ganglia. *Journal of Anatomy*, 196, 527-42.

BONITA, R. & BEAGLEHOLE, R. 1988. Modification of Rankin scale: recovery of motor function after stroke. *Stroke*, 19, 1497-500.

BRAUER, S. G., BEW, P. G., KUYS, S. S., LYNCH, M. R. & MORRISON, G. 2008. Prediction of discharge destination after stroke using the motor assessment scale on admission: a prospective, multisite study. *Archives of Physical Medicine & Rehabilitation*, 89, 1061-5.

BRODERICK, J. P., BROTT, T. G., DULDNER, J. E., TOMSICK, T. & HUSTER, G. 1993. Volume of intracerebral hemorrhage. A powerful and easy-to-use predictor of 30-day mortality. *Stroke*, 24, 987-93.

BROMBERG-MARTIN, E. S., MATSUMOTO, M. & HIKOSAKA, O. 2010. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*, 68, 815-34.

BROTT, T., ADAMS, H. P., JR., OLINGER, C. P., MARLER, J. R., BARSAN, W. G., BILLER, J., SPILKER, J., HOLLERAN, R., EBERLE, R., HERTZBERG, V. & ET AL. 1989. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke*, 20, 864-70.

BROWN, C., HASSON, H., THYSELIUS, V. & ALMBORG, A. H. 2012. Post-stroke depression and functional independence: a conundrum. *Acta Neurologica Scandinavica*, 126, 45-51.

BRUNO, A., SHAH, N., LIN, C., CLOSE, B., HESS, D. C., DAVIS, K., BAUTE, V., SWITZER, J. A., WALLER, J. L. & NICHOLS, F. T. 2010. Improving modified Rankin Scale assessment with a simplified questionnaire. *Stroke*, 41, 1048-50.

CAMPBELL, B. C., MITCHELL, P. J., KLEINIG, T. J., DEWEY, H. M., CHURILOV, L., YASSI, N., YAN, B., DOWLING, R. J., PARSONS, M. W., OXLEY, T. J., WU, T. Y., BROOKS, M., SIMPSON, M. A., MITEFF, F., LEVI, C. R., KRAUSE, M., HARRINGTON, T. J., FAULDER,

K. C., STEINFORT, B. S., PRIGLINGER, M., ANG, T., SCROOP, R., BARBER, P. A., MCGUINNESS, B., WIJERATNE, T., PHAN, T. G., CHONG, W., CHANDRA, R. V., BLADIN, C. F., BADVE, M., RICE, H., DE VILLIERS, L., MA, H., DESMOND, P. M., DONNAN, G. A., DAVIS, S. M. & INVESTIGATORS, E.-I. 2015. Endovascular therapy for ischemic stroke with perfusion-imaging selection. *New England Journal of Medicine*, 372, 1009-18.

CANDELISE L, G, P. & A, M. 1991. Mortality in acute stroke with atrial fibrillation. *Stroke*, 22, 169-174.

CARR, J. H., SHEPHERD, R. B., NORDHOLM, L. & LYNNE, D. 1985. Investigation of a new motor assessment scale for stroke patients. *Physical Therapy*, 65, 175-80.

CASSIDY, E., O'CONNOR, R. & O'KEANE, V. 2004. Prevalence of post-stroke depression in an Irish sample and its relationship with disability and outcome following inpatient rehabilitation. *Disability & Rehabilitation*, 26, 71-7.

CHEN, C., LEYS, D. & ESQUENAZI, A. 2013. The interaction between neuropsychological and motor deficits in patients after stroke. *Neurology*, 80, S27-34.

CHEUNG, R. T. & ZOU, L. Y. 2003. Use of the original, modified, or new intracerebral hemorrhage score to predict mortality and morbidity after intracerebral hemorrhage. *Stroke*, 34, 1717-22.

CHISHOLM-BURNS, M. A. & SPIVEY, C. A. 2008. Pharmacoadherence: a new term for a significant problem. *American Journal of Health-System Pharmacy*, 65, 661-7.

CHO, D. Y., CHEN, C. C., LEE, W. Y., LEE, H. C. & HO, L. H. 2008. A new Modified Intracerebral Hemorrhage score for treatment decisions in basal ganglia hemorrhage--a randomized trial. *Critical Care Medicine*, 36, 2151-6.

CHOI-KWON, S. & KIM, J. S. 2011. Poststroke fatigue: an emerging, critical issue in stroke medicine. *International Journal of Stroke*, 6, 328-36.

CHOLLET, F., TARDY, J., ALBUCHER, J. F., THALAMAS, C., BERARD, E., LAMY, C., BEJOT, Y., DELTOUR, S., JAILLARD, A., NICLOT, P., GUILLON, B., MOULIN, T., MARQUE, P., PARIENTE, J., ARNAUD, C. & LOUBINOUX, I. 2011. Fluoxetine for motor recovery after acute ischaemic stroke (FLAME): a randomised placebo-controlled trial.[Erratum appears in *Lancet Neurol*. 2011 Mar;10(3):205]. *Lancet Neurology*, 10, 123-30.

CLAASSEN, J., BERNARDINI, G. L., KREITER, K., BATES, J., DU, Y. E., COPELAND, D., CONNOLLY, E. S. & MAYER, S. A. 2001. Effect of cisternal and ventricular blood on risk of delayed cerebral ischemia after subarachnoid hemorrhage: the Fisher scale revisited. *Stroke*, 32, 2012-20.

COLLEN, F. M., WADE, D. T., ROBB, G. F. & BRADSHAW, C. M. 1991. The Rivermead Mobility Index: a further development of the Rivermead Motor Assessment. *International Disability Studies*, 13, 50-4.

COLLINS, R. & MACMAHON, S. 2001. Reliable assessment of the effects of treatment on mortality and major morbidity, I: clinical trials. *Lancet*, 357, 373-80.

COOK, R. D. 1979. Influential observations in linear regression. *Journal of the American Statistical Association*, 74, 169-174.

COUTTS, S. B., DEMCHUK, A. M., BARBER, P. A., HU, W. Y., SIMON, J. E., BUCHAN, A. M., HILL, M. D. & GROUP, V. S. 2004. Interobserver variation of ASPECTS in real time. *Stroke*, 35, e103-5.

CRAIG, L. E., WU, O., BERNHARDT, J. & LANGHORNE, P. 2011. Predictors of poststroke mobility: systematic review. *International Journal of Stroke*, 6, 321-7.

CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZARETH, I., PETTICREW, M. & MEDICAL RESEARCH COUNCIL, G. 2008. Developing and evaluating complex interventions: the new Medical Research Council guidance.[Reprint in *Int J Nurs Stud*. 2013 May;50(5):587-92; PMID: 23159157]. *BMJ*, 337, a1655.

CRUZ-FLORES, S., RABINSTEIN, A., BILLER, J., ELKIND, M. S., GRIFFITH, P., GORELICK, P. B., HOWARD, G., LEIRA, E. C., MORGENSTERN, L. B., OVBIAGELE, B., PETERSON, E., ROSAMOND, W., TRIMBLE, B., VALDERRAMA, A. L., AMERICAN HEART ASSOCIATION STROKE, C., COUNCIL ON CARDIOVASCULAR, N., COUNCIL ON, E., PREVENTION, COUNCIL ON QUALITY OF, C. & OUTCOMES, R. 2011. Racial-ethnic disparities in stroke care: the American experience: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, 42, 2091-116.

CUMMING, T. B., BERNHARDT, J. & LINDEN, T. 2011. The montreal cognitive assessment: short cognitive evaluation in a large stroke trial. *Stroke*, 42, 2642-4.

DA CUNHA, C., WIETZIKOSKI, E. C., DOMBROWSKI, P., BORTOLANZA, M., SANTOS, L. M., BOSCHEN, S. L. & MIYOSHI, E. 2009. Learning processing in the basal ganglia: a mosaic of broken mirrors. *Behavioural Brain Research*, 199, 157-70.

DA ROCHA, N. S., CHACHAMOVICH, E., DE ALMEIDA FLECK, M. P. & TENNANT, A. 2013. An introduction to Rasch analysis for Psychiatric practice and research. *Journal of Psychiatric Research*, 47, 141-8.

DAHL, T. H. 2002. International classification of functioning, disability and health: an introduction and discussion of its potential impact on rehabilitation services and research. *Journal of Rehabilitation Medicine*, 34, 201-4.

DAVIS, S. M., BRODERICK, J., HENNERICI, M., BRUN, N. C., DIRINGER, M. N., MAYER, S. A., BEGRUP, K., STEINER, T. & RECOMBINANT ACTIVATED FACTOR, V. I. I. H. T. I. 2006. Hematoma growth is a determinant of mortality and poor outcome after intracerebral hemorrhage. *Neurology*, 66, 1175-81.

DAWES, H., ENZINGER, C., JOHANSEN-BERG, H., BOGDANOVIC, M., GUY, C., COLLETT, J., IZADI, H., STAGG, C., WADE, D. & MATTHEWS, P. M. 2008. Walking performance and its recovery in chronic stroke in relation to extent of lesion overlap with the descending motor tract. *Experimental Brain Research*, 186, 325-33.

DAYAN, E. & COHEN, L. G. 2011. Neuroplasticity subserving motor skill learning. *Neuron*, 72, 443-54.

DE ALMEIDA, A. N., ALHO, E. J. & TEIXEIRA, M. J. 2014. Models of functional cerebral localization at the dawning of modern neurosurgery. *World Neurosurgery*, 81, 436-40.

DE KLEIJN, W. P., DE VRIES, J., WIJNEN, P. A. & DRENT, M. 2011. Minimal (clinically) important differences for the Fatigue Assessment Scale in sarcoidosis. *Respiratory Medicine*, 105, 1388-95.

DE LAAT, K. F., TULADHAR, A. M., VAN NORDEN, A. G., NORRIS, D. G., ZWIERS, M. P. & DE LEEUW, F. E. 2011. Loss of white matter integrity is associated with gait disorders in cerebral small vessel disease. *Brain*, 134, 73-83.

DE WIT, L., PUTMAN, K., SCHUBACK, B., KOMAREK, A., ANGST, F., BAERT, I., BERMAN, P., BOGAERTS, K., BRINKMANN, N., CONNELL, L., DEJAEGER, E., FEYS, H., JENNI, W., KASKE, C., LESAFFRE, E., LEYS, M., LINCOLN, N., LOUCKX, F., SCHUPP, W., SMITH, B. & DE WEERDT, W. 2007. Motor and functional recovery after stroke: a comparison of 4 European rehabilitation centers. *Stroke*, 38, 2101-7.

DEBETTE, S. & MARKUS, H. S. 2010. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis. *BMJ*, 341, c3666.

DELCOURT, C., HUANG, Y., ARIMA, H., CHALMERS, J., DAVIS, S. M., HEELEY, E. L., WANG, J., PARSONS, M. W., LIU, G., ANDERSON, C. S. & ON BEHALF OF THE INTERACT INVESTIGATORS 2012. Hematoma growth and outcomes in intracerebral hemorrhage: the INTERACT1 study. *Neurology*, 79, 314-9.

DELCOURT, C., HUANG, Y., WANG, J., HEELEY, E., LINDLEY, R., STAPF, C., TZOURIO, C., ARIMA, H., PARSONS, M., SUN, J., NEAL, B., CHALMERS, J., ANDERSON, C. & ON BEHALF OF THE INTERACT INVESTIGATORS 2010. The second (main) phase of an open, randomised, multicentre study to investigate the effectiveness of an intensive blood pressure reduction in acute cerebral haemorrhage trial (INTERACT2). *International Journal of Stroke*, 5, 110-6.

DENNIS, M. S. 2003. Outcome after brain haemorrhage. *Cerebrovascular Diseases*, 16 Suppl 1, 9-13.

DEPARTMENT OF HEALTH 2005. Research Governance Framework for Health and Social Care: Second edition, 2005. Available online at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/139565/dh_4_122427.pdf.

DEPARTMENT OF HEALTH 2007. National Stroke Strategy. London, The Stationary Office. Available online at <http://www.dh.gov.uk/health/search/?searchTerms=national+stroke+strategy>.

DESMOND, D. W., REMIEN, R. H., MORONEY, J. T., STERN, Y., SANO, M. & WILLIAMS, J. B. 2003. Ischemic stroke and depression. *Journal of the International Neuropsychological Society*, 9, 429-39.

DI CARLO, A., LAMASSA, M., BALDERESCHI, M., PRACUCCI, G., CONSOLI, D., WOLFE, C. D., GIROUD, M., RUDD, A., BURGER, I., GHETTI, A., INZITARI, D. & EUROPEAN, B. S. O. S. C. G. 2006. Risk factors and outcome of subtypes of ischemic stroke. Data from a multicenter multinational hospital-based registry. The European Community Stroke Project. *Journal of the Neurological Sciences*, 244, 143-50.

DIMYAN, M. A. & COHEN, L. G. 2011. Neuroplasticity in the context of motor rehabilitation after stroke. *Nature Reviews Neurology*.

DIVANI, A. A., MAJIDI, S., LUO, X., SOUSLIAN, F. G., ZHANG, J., ABOSCH, A. & TUMMALA, R. P. 2011. The ABCs of accurate volumetric measurement of cerebral hematoma. *Stroke*, 42, 1569-74.

DOMBROWSKI, S. U., MACKINTOSH, J. E., SNIEHOTTA, F. F., ARAUJO-SOARES, V., RODGERS, H., THOMSON, R. G., MURTAGH, M. J., FORD, G. A., ECCLES, M. P. & WHITE, M. 2013. The impact of the UK 'Act FAST' stroke awareness campaign: content analysis of patients, witness and primary care clinicians' perceptions. *BMC Public Health*, 13, 915.

DONG, Y., SHARMA, V. K., CHAN, B. P., VENKETASUBRAMANIAN, N., TEOH, H. L., SEET, R. C., TANICALA, S., CHAN, Y. H. & CHEN, C. 2010. The Montreal Cognitive Assessment (MoCA) is superior to the Mini-Mental State Examination (MMSE) for the detection of vascular cognitive impairment after acute stroke. *Journal of the Neurological Sciences*, 299, 15-8.

DORRESTEIJN, J. A., VISSEREN, F. L., RIDKER, P. M., WASSINK, A. M., PAYNTER, N. P., STEYERBERG, E. W., VAN DER GRAAF, Y. & COOK, N. R. 2011. Estimating treatment effects for individual patients based on the results of randomised clinical trials. *BMJ*, 343, d5888.

DOWLATSHAHI, D., DEMCHUK, A. M., FLAHERTY, M. L., ALI, M., LYDEN, P. L., SMITH, E. E. & COLLABORATION, V. 2011. Defining hematoma expansion in intracerebral hemorrhage: relationship with patient outcomes. *Neurology*, 76, 1238-44.

DOWLATSHAHI, D., SMITH, E. E., FLAHERTY, M. L., ALI, M., LYDEN, P., DEMCHUK, A. M. & COLLABORATORS, V. 2010. Small intracerebral haemorrhages are associated with less haematoma expansion and better outcomes. *International Journal of Stroke*, 6, 201-6.

DOYON, J., BELLEC, P., AMSEL, R., PENHUNE, V., MONCHI, O., CARRIER, J., LEHERICY, S. & BENALI, H. 2009. Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behavioural Brain Research*, 199, 61-75.

DUDLEY, N. 2011. England's stroke FAST campaign has been evaluated: Health Minister appears to have inadvertently misled Parliament on thrombolysis effects and ROMI. *British Medical Journal*, 342, d243.

DUNCAN, F., GREIG, C., LEWIS, S., DENNIS, M., MACLULLICH, A., SHARPE, M. & MEAD, G. 2014. Clinically significant fatigue after stroke: a longitudinal cohort study. *Journal of Psychosomatic Research*, 77, 368-73.

DUNN, O. J. 1961. Multiple comparisons among means. *Journal of the American Statistical Association*, 56, 52-64.

DZIALOWSKI, I., HILL, M. D., COUTTS, S. B., DEMCHUK, A. M., KENT, D. M., WUNDERLICH, O. & VON KUMMER, R. 2006. Extent of early ischemic changes on computed tomography (CT) before thrombolysis: prognostic value of the Alberta Stroke Program Early CT Score in ECASS II. *Stroke*, 37, 973-8.

EARLY SUPPORTED DISCHARGE TRIALISTS 2009. Services for reducing duration of hospital care for acute stroke patients. Cochrane Database of Systematic Reviews. Issue 2. Art. No.: CD000443. DOI: 10.1002/14651858.CD000443.pub2.

EMSLEY, H. C., BLACKER, K., DAY, E., PUNEKAR, S., O'DONNELL, M., DAVIES, D. P. & TURNBULL, I. W. 2013. Acute Stroke Training and Assessment in Computerised Axial Tomography (ASTRACAT) for stroke physicians: facilitating telestroke implementation in a UK setting. *Journal of Contemporary Medical Education*, 1, 137-144.

ENGELTER, S. T., FRANK, M., LYRER, P. A. & CONZELMANN, M. 2010. Safety of pharmacological augmentation of stroke rehabilitation. *European Neurology*, 64, 325-30.

ERIKSSON, M., ASPLUND, K., GLADER, E. L., NORRVING, B., STEGMAYR, B., TERENT, A., ASBERG, K. H., WESTER, P. O. & RIKS-STROKE, C. 2004. Self-reported depression and use of antidepressants after stroke: a national survey. *Stroke*, 35, 936-41.

FARRELL, B., GODWIN, J., RICHARDS, S. & WARLOW, C. 1991. The United Kingdom transient ischaemic attack (UK-TIA) aspirin trial: final results. *Journal of Neurology, Neurosurgery & Psychiatry*, 54, 1044-54.

FAURE, H. & HRYNASZKIEWICZ, I. 2011. The ISRCTN Register: achievements and challenges 8 years on. *Journal of Evidence-based Medicine*, 4, 188-92.

FAZEKAS, F., KLEINERT, R., ROOB, G., KLEINERT, G., KAPPELLER, P., SCHMIDT, R. & HARTUNG, H. P. 1999. Histopathologic analysis of foci of signal loss on gradient-echo T2-weighted MR images in patients with spontaneous intracerebral hemorrhage: evidence of microangiopathy-related microbleeds. *Ajnr: American Journal of Neuroradiology*, 20, 637-42.

FEIGIN, V. L., BARKER-COLLO, S., PARAG, V., SENIOR, H., LAWES, C. M. M., RATNASABAPATHY, Y., GLEN, E. & GROUP, A. S. 2010. Auckland Stroke Outcomes Study. Part 1: Gender, stroke types, ethnicity, and functional outcomes 5 years poststroke. *Neurology*, 75, 1597-607.

FEIGIN, V. L., FOROUZANFAR, M. H., KRISHNAMURTHI, R., MENSAH, G. A., CONNOR, M., BENNETT, D. A., MORAN, A. E., SACCO, R. L., ANDERSON, L., TRUELSEN, T., O'DONNELL, M., VENKETASUBRAMANIAN, N., BARKER-COLLO, S., LAWES, C. M. M., WANG, W., SHINOHARA, Y., WITT, E., EZZATI, M., NAGHAVI, M., MURRAY, C., GLOBAL BURDEN OF DISEASES, I., RISK FACTORS, S. & THE, G. B. D. S. E. G. 2014. Global and regional burden of stroke during 1990-2010: findings from the Global Burden of Disease Study 2010. *Lancet*, 383, 245-54.

FEIGIN, V. L., LAWES, C. M., BENNETT, D. A., BARKER-COLLO, S. L. & PARAG, V. 2009. Worldwide stroke incidence and early case fatality reported in 56 population-based studies: a systematic review. *Lancet Neurology*, 8, 355-69.

FIELD, A. 2013. *Discovering statistics using IBM SPSS Statistics*. 4th Edition. London; Sage Publications.

FISHER, C. M. 1965. Lacunes: Small, Deep Cerebral Infarcts. *Neurology*, 15, 774-84.

FISHER, C. M. 1968. The arterial lesions underlying lacunes. *Acta Neuropathologica*, 12, 1-15.

FLOEL, A., HUMMEL, F., BREITENSTEIN, C., KNECHT, S. & COHEN, L. G. 2005. Dopaminergic effects on encoding of a motor memory in chronic stroke. *Neurology*, 65, 472-4.

FOGELHOLM, R., AVIKAINEN, S. & MURROS, K. 1997. Prognostic value and determinants of first-day mean arterial pressure in spontaneous supratentorial intracerebral hemorrhage. *Stroke*, 28, 1396-400.

FOLSTEIN, M. F., FOLSTEIN, S. E. & MCHUGH, P. R. 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-98.

FONAROW, G. C., SAVER, J. L., SMITH, E. E., BRODERICK, J. P., KLEINDORFER, D. O., SACCO, R. L., PAN, W., OLSON, D. M., HERNANDEZ, A. F., PETERSON, E. D. & SCHWAMM, L. H. 2012. Relationship of national institutes of health stroke scale to 30-day mortality in medicare beneficiaries with acute ischemic stroke. *Journal of the American Heart Association*, 1, 42-50.

FRENCH, B., THOMAS, L. H., LEATHLEY, M. J., SUTTON, C. J., MCADAM, J., FORSTER, A., LANGHORNE, P., PRICE, C. I., WALKER, A. & WATKINS, C. L. 2007. Repetitive task training for improving functional ability after stroke. *Cochrane Database of Systematic Reviews*, CD006073.

GACS, G., FOX, A. J., BARNETT, H. J. & VINUELA, F. 1983. CT visualization of intracranial arterial thromboembolism. *Stroke*, 14, 756-62.

GARBUSINSKI, J. M., VAN DER SANDE, M. A., BARTHOLOME, E. J., DRAMAIX, M., GAYE, A., COLEMAN, R., NYAN, O. A., WALKER, R. W., MCADAM, K. P. & WALRAVEN, G. E. 2005. Stroke presentation and outcome in developing countries: a prospective study in the Gambia. *Stroke*, 36, 1388-93.

GEBEL, J. M., SILA, C. A., SLOAN, M. A., GRANGER, C. B., MAHAFFEY, K. W., WEISENBERGER, J., GREEN, C. L., WHITE, H. D., GORE, J. M., WEAVER, W. D., CALIFF, R. M. & TOPOL, E. J. 1998. Thrombolysis-related intracranial hemorrhage: a radiographic analysis of 244 cases from the GUSTO-1 trial with clinical correlation. *Global Utilization of Streptokinase and Tissue Plasminogen Activator for Occluded Coronary Arteries*. *Stroke*, 29, 563-9.

GEYH, S., CIEZA, A., SCHOUTEN, J., DICKSON, H., FROMMELT, P., OMAR, Z., KOSTANJSEK, N., RING, H. & STUCKI, G. 2004. ICF Core Sets for stroke. *Journal of Rehabilitation Medicine*, 135-41.

GILLICK, B. T. & ZIRPEL, L. 2012. Neuroplasticity: an appreciation from synapse to system. *Archives of Physical Medicine & Rehabilitation*, 93, 1846-55.

GODEFROY, O., FICKL, A., ROUSSEL, M., AURIBAUT, C., BUGNICOURT, J. M., LAMY, C., CANAPLE, S. & PETITNICOLAS, G. 2011. Is the montreal cognitive assessment superior to the mini-mental state examination to detect poststroke cognitive impairment?: a study with neuropsychological evaluation. *Stroke*, 42, 1712-6.

GODOY, D. A., PINERO, G. & DI NAPOLI, M. 2006. Predicting mortality in spontaneous intracerebral hemorrhage: can modification to original score improve the prediction? *Stroke*, 37, 1038-44.

GOLDBERG, D. P. & BLACKWELL, B. 1970. Psychiatric illness in general practice. A detailed study using a new method of case identification. *British Medical Journal*, 1, 439-43.

GOLDBERG, D. P. & HILLIER, V. F. 1979. A scaled version of the General Health Questionnaire. *Psychological Medicine*, 9, 139-45.

GOMPERTZ, P., POUND, P. & EBRAHIM, S. 1994. A postal version of the Barthel Index. *Clinical Rehabilitation*, 8, 233-239.

GOTTESMAN, R. F., HILLIS, A. E., GOTTESMAN, R. F. & HILLIS, A. E. 2010. Predictors and assessment of cognitive dysfunction resulting from ischaemic stroke. *Lancet Neurology*, 9, 895-905.

GOUW, A. A., VAN DER FLIER, W. M., PANTONI, L., INZITARI, D., ERKINJUNTTI, T., WAHLUND, L. O., WALDEMAR, G., SCHMIDT, R., FAZEKAS, F., SCHELTENS, P., BARKHOF, F. & GROUP, L. S. 2008. On the etiology of incident brain lacunes: longitudinal observations from the LADIS study. *Stroke*, 39, 3083-5.

GOYAL, M., DEMCHUK, A. M., MENON, B. K., EESA, M., REMPEL, J. L., THORNTON, J., ROY, D., JOVIN, T. G., WILLINSKY, R. A., SAPKOTA, B. L., DOWLATSHAHI, D., FREI, D. F., KAMAL, N. R., MONTANERA, W. J., POPPE, A. Y., RYCKBORST, K. J., SILVER, F. L., SHUAIB, A., TAMPIERI, D., WILLIAMS, D., BANG, O. Y., BAXTER, B. W., BURNS, P. A., CHOE, H., HEO, J. H., HOLMSTEDT, C. A., JANKOWITZ, B., KELLY, M., LINARES, G., MANDZIA, J. L., SHANKAR, J., SOHN, S. I., SWARTZ, R. H., BARBER, P. A., COUTTS, S. B., SMITH, E. E., MORRISH, W. F., WEILL, A., SUBRAMANIAM, S., MITHA, A. P., WONG, J. H., LOWERISON, M. W., SAJOBI, T. T., HILL, M. D. & INVESTIGATORS, E. T. 2015. Randomized assessment of rapid endovascular treatment of ischemic stroke. *New England Journal of Medicine*, 372, 1019-30.

GOYAL, M., MENON, B. K., VAN ZWAM, W. H., DIPPEL, D. W., MITCHELL, P. J., DEMCHUK, A. M., DAVALOS, A., MAJOIE, C. B., VAN DER LUGT, A., DE MIQUEL, M. A., DONNAN, G. A., ROOS, Y. B., BONAFE, A., JAHAN, R., DIENER, H. C., VAN DEN BERG, L. A., LEVY, E. I., BERKHEMER, O. A., PEREIRA, V. M., REMPEL, J., MILLAN, M., DAVIS, S. M., ROY, D., THORNTON, J., ROMAN, L. S., RIBO, M., BEUMER, D., STOUCHE, B., BROWN, S., CAMPBELL, B. C., VAN OOSTENBRUGGE, R. J., SAVER, J. L., HILL, M. D., JOVIN, T. G. & COLLABORATORS, H. 2016. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *Lancet*, 387, 1723-31.

GREENBERG, S. M., VERNOOIJ, M. W., CORDONNIER, C., VISWANATHAN, A., AL-SHAHI SALMAN, R., WARACH, S., LAUNER, L. J., VAN BUCHEM, M. A., BRETELER, M. M. & MICROBLEED STUDY, G. 2009. Cerebral microbleeds: a guide to detection and interpretation. *Lancet Neurology*, 8, 165-74.

GROSS, C. G. 2007. The discovery of motor cortex and its background. *Journal of the History of the Neurosciences*, 16, 320-31.

GROTTA, J. C., CHIU, D., LU, M., PATEL, S., LEVINE, S. R., TILLEY, B. C., BROTT, T. G., HALEY, E. C., JR., LYDEN, P. D., KOTHARI, R., FRANKEL, M., LEWANDOWSKI, C. A., LIBMAN, R., KWIATKOWSKI, T., BRODERICK, J. P., MARLER, J. R., CORRIGAN, J., HUFF, S., MITSIAS, P., TALATI, S. & TANNE, D. 1999. Agreement and variability in the interpretation of early CT changes in stroke patients qualifying for intravenous rtPA therapy. *Stroke*, 30, 1528-33.

GUTENBRUNNER, C., WARD, A. B. & CHAMBERLAIN, M. A. 2006. White Book on Physical and Rehabilitation Medicine in Europe. Produced by the Section of Physical and Rehabilitation Medicine, Union Européenne des Médecins Spécialistes (UEMS), European Board of Physical and Rehabilitation Medicine and L'Académie Européenne de Médecine de Réadaptation, in conjunction with the European Society for Physical and Rehabilitation Medicine.

HACHINSKI, V. 2007. The 2005 Thomas Willis Lecture: stroke and vascular cognitive impairment: a transdisciplinary, translational and transactional approach. *Stroke*, 38, 1396.

HACKE, W., KASTE, M., FIESCHI, C., TONI, D., LESAFFRE, E., VON KUMMER, R., BOYSEN, G., BLUHMKI, E., HOXTER, G., MAHAGNE, M. H. & ET AL. 1995. Intravenous thrombolysis with recombinant tissue plasminogen activator for acute hemispheric stroke. The European Cooperative Acute Stroke Study (ECASS). *JAMA*, 274, 1017-25.

HACKE, W., KASTE, M., FIESCHI, C., VON KUMMER, R., DAVALOS, A., MEIER, D., LARRUE, V., BLUHMKI, E., DAVIS, S., DONNAN, G., SCHNEIDER, D., DIEZ-TEJEDOR, E. & TROUILLAS, P. 1998. Randomised double-blind placebo-controlled trial of thrombolytic therapy with intravenous alteplase in acute ischaemic stroke (ECASS II). Second European-Australasian Acute Stroke Study Investigators. *Lancet*, 352, 1245-51.

HACKETT, M. L., YAPA, C., PARAG, V. & ANDERSON, C. S. 2005. Frequency of depression after stroke: a systematic review of observational studies. *Stroke*, 36, 1330-40.

HALLEVI, H., ALBRIGHT, K. C., ARONOWSKI, J., BARRETO, A. D., MARTIN-SCHILD, S., KHAJA, A. M., GONZALES, N. R., ILLOH, K., NOSER, E. A. & GROTTA, J. C. 2008. Intraventricular hemorrhage: Anatomic relationships and clinical implications. *Neurology*, 70, 848-52.

HALLEVI, H., DAR, N. S., BARRETO, A. D., MORALES, M. M., MARTIN-SCHILD, S., ABRAHAM, A. T., WALKER, K. C., GONZALES, N. R., ILLOH, K., GROTTA, J. C. & SAVITZ, S. I. 2009. The IVH score: a novel tool for estimating intraventricular hemorrhage volume: clinical and research implications. *Critical Care Medicine*, 37, 969-74, e1.

HALLEVY, C., IFERGANE, G., KORDYSH, E. & HERISHANU, Y. 2002. Spontaneous supratentorial intracerebral hemorrhage. Criteria for short-term functional outcome prediction. *Journal of Neurology*, 249, 1704-9.

HANKINS, M. 2008a. The factor structure of the twelve item General Health Questionnaire (GHQ-12): the result of negative phrasing? *Clinical Practice & Epidemiology in Mental Health [Electronic Resource]*: CP & EMH, 4, 10.

HANKINS, M. 2008b. The reliability of the twelve-item general health questionnaire (GHQ-12) under realistic assumptions. *BMC Public Health*, 8, 355.

HEMPHILL, J. C., 3RD, BONOVIK, D. C., BESMERTIS, L., MANLEY, G. T. & JOHNSTON, S. C. 2001. The ICH score: a simple, reliable grading scale for intracerebral hemorrhage. *Stroke*, 32, 891-7.

HETTIARACHCHI, C., CONAGHAN, P., TENNANT, A. & BHAKTA, B. 2011. Prevalence and impact of joint symptoms in people with stroke aged 55 years and over. *Journal of Rehabilitation Medicine*, 43, 197-203.

HIKOSAKA, O., NAKAMURA, K., SAKAI, K. & NAKAHARA, H. 2002. Central mechanisms of motor skill learning. *Current Opinion in Neurobiology*, 12, 217-22.

HILL, M. D., ROWLEY, H. A., ADLER, F., ELIASZIW, M., FURLAN, A., HIGASHIDA, R. T., WECHSLER, L. R., ROBERTS, H. C., DILLON, W. P., FISCHBEIN, N. J., FIRSZT, C. M., SCHULZ, G. A., BUCHAN, A. M. & INVESTIGATORS, P.-I. 2003. Selection of acute ischemic stroke patients for intra-arterial thrombolysis with pro-urokinase by using ASPECTS. *Stroke*, 34, 1925-31.

HIRANO, T., SASAKI, M., TOMURA, N., ITO, Y., KOBAYASHI, S. & JAPAN ALTEPLASE CLINICAL TRIAL, G. 2012. Low Alberta stroke program early computed tomography score within 3 hours of onset predicts subsequent symptomatic intracranial hemorrhage in patients

treated with 0.6 mg/kg Alteplase. *Journal of Stroke & Cerebrovascular Diseases*, 21, 898-902.

HOBART, J. & CANO, S. 2009. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technology Assessment (Winchester, England)*, 13, iii.

HODICS, T., COHEN, L. G. & CRAMER, S. C. 2006. Functional imaging of intervention effects in stroke motor rehabilitation. *Archives of Physical Medicine & Rehabilitation*, 87, S36-42.

HUNTINGTON, G. 1872. On chorea. *Original publication in The Medical and Surgical Reporter: A Weekly Journal (Philadelphia: S.W. Butler)*, vol. 26, no. 15 (April 13, 1872), pp. 317-321. Reprinted in: *Journal of Neuropsychiatry & Clinical Neurosciences*, 2003, 15, 109-12.

HUPPERT, F. A., GORE, M. & ELLIOTT, B. J. 1988. The value of an improved scoring system (CGHQ) for the General Health Questionnaire in a representative community sample. *Psychological Medicine*, 18, 1001-6.

IADECOLA, C. 2013. The pathobiology of vascular dementia. *Neuron*, 80, 844-66.

IADECOLA, C. & ANRATHER, J. 2011. Stroke research at a crossroad: asking the brain for directions. *Nature Neuroscience*, 14, 1363-8.

INTERCOLLEGIATE STROKE WORKING PARTY 2011. *The National Sentinel Stroke Audit 2010: Round 7*. London; The Royal College of Physicians. Available online at <https://www.rcplondon.ac.uk/projects/national-sentinel-stroke-audit>.

INZITARI, D., PRACUCCI, G., POGGESI, A., CARLUCCI, G., BARKHOF, F., CHABRIAT, H., ERKINJUNTTI, T., FAZEKAS, F., FERRO, J. M., HENNERICI, M., LANGHORNE, P., O'BRIEN, J., SCHELTENS, P., VISSER, M. C., WAHLUND, L. O., WALDEMAR, G., WALLIN, A., PANTONI, L. & GROUP, L. S. 2009. Changes in white matter as determinant of global functional decline in older independent outpatients: three year follow-up of LADIS (leukoaraiosis and disability) study cohort. *BMJ*, 339, b2477.

ISMAIL, Z., RAJJI, T. K. & SHULMAN, K. I. 2010. Brief cognitive screening instruments: an update. *International Journal of Geriatric Psychiatry*, 25, 111-20.

JENNETT, B. & BOND, M. 1975. Assessment of outcome after severe brain damage. *Lancet*, 1, 480-4.

JOHNSTON, K. C., BARRETT, K. M., DING, Y. H., WAGNER, D. P. & ACUTE STROKE ACCURATE PREDICTION, I. 2009a. Clinical and imaging data at 5 days as a surrogate for 90-day outcome in ischemic stroke. *Stroke*, 40, 1332-3.

JOHNSTON, K. C., CONNORS, A. F., WAGNER, D. P., KNAUS, W. A., WANG, X.-Q. & HALEY, E. C. 2000. A Predictive Risk Model for Outcomes of Ischemic Stroke. *Stroke*, 31, 448-455.

JOHNSTON, K. C., WAGNER, D. P., HALEY, E. C., JR., CONNORS, A. F., JR. & STROKE, R. I. R. T. O. T. M. I. A. 2002. Combined clinical and imaging information as an early stroke outcome measure. *Stroke*, 33, 466-72.

JOHNSTON, S. C., MENDIS, S. & MATHERS, C. D. 2009b. Global variation in stroke burden and mortality: estimates from monitoring, surveillance, and modelling. *Lancet Neurology*, 8, 345-54.

JOVIN, T. G., CHAMORRO, A., COBO, E., DE MIQUEL, M. A., MOLINA, C. A., ROVIRA, A., SAN ROMAN, L., SERENA, J., ABILLEIRA, S., RIBO, M., MILLAN, M., URRRA, X., CARDONA, P., LOPEZ-CANCIO, E., TOMASELLO, A., CASTANO, C., BLASCO, J., AJA, L., DORADO, L., QUESADA, H., RUBIERA, M., HERNANDEZ-PEREZ, M., GOYAL, M., DEMCHUK, A. M., VON KUMMER, R., GALLOFRE, M., DAVALOS, A. & INVESTIGATORS, R. T. 2015. Thrombectomy within 8 hours after symptom onset in ischemic stroke. *New England Journal of Medicine*, 372, 2296-306.

KAMALAKANNAN, S., GUDLAVALLETI VENKATA, M., PROST, A., NATARAJAN, S., PANT, H., CHITALURRI, N., GOENKA, S. & KUPER, H. 2016. Rehabilitation Needs of Stroke Survivors After Discharge From Hospital in India. *Archives of Physical Medicine & Rehabilitation*, 97, 1526-1532.e9.

KASTE, M., FIESCHI, C., TONI, D., LESAFFRE, E., VON KUMMER, R., BOYSEN, G., BLUHMKI, E., HOXTER, G. & MAHAGNE, M. H. 1995. Intravenous thrombolysis with recombinant tissue plasminogen activator for acute hemispheric stroke. The European Cooperative Acute Stroke Study (ECASS). *JAMA*, 274, 1017-25.

KEENAN, A. M., TENNANT, A., FEAR, J., EMERY, P. & CONAGHAN, P. G. 2006. Impact of multiple joint problems on daily living tasks in people in the community over age fifty-five. *Arthritis & Rheumatism*, 55, 757-64.

KELLY, M. J., DUNSTAN, F. D., LLOYD, K. & FONE, D. L. 2008. Evaluating cutpoints for the MHI-5 and MCS using the GHQ-12: a comparison of five different methods. *BMC Psychiatry*, 8, 10.

KELLY, P. J., FURIE, K. L., SHAFQAT, S., RALLIS, N., CHANG, Y. & STEIN, J. 2003. Functional recovery following rehabilitation after hemorrhagic and ischemic stroke. *Archives of Physical Medicine & Rehabilitation*, 84, 968-72.

KERSTEN, P., KUCUKDEVECI, A. A. & TENNANT, A. 2012. The use of the Visual Analogue Scale (VAS) in rehabilitation outcomes. *Journal of Rehabilitation Medicine*, 44, 609-10.

KERSTEN, P., WHITE, P. J. & TENNANT, A. 2014. Is the pain visual analogue scale linear and responsive to change? An exploration using Rasch analysis. *PLoS ONE [Electronic Resource]*, 9, e99485.

KHAN, F., AMATYA, B., ELMALIK, A., LOWE, M., NG, L., REID, I. & GALEA, M. P. An enriched environmental programme during inpatient neuro-rehabilitation: A randomized controlled trial. *Journal of Rehabilitation Medicine*, 48, 417-25.

KOBAYASHI, A., WARDLAW, J. M., LINDLEY, R. I., LEWIS, S. C., SANDERCOCK, P. A., CZLONKOWSKA, A. & GROUP, I. S. T. C. 2009. Oxfordshire community stroke project clinical stroke syndrome and appearances of tissue and vascular lesions on pretreatment ct in hyperacute ischemic stroke among the first 510 patients in the Third International Stroke Trial (IST-3). *Stroke*, 40, 743-8.

KOLLEN, B., KWAKKEL, G. & LINDEMAN, E. 2006. Longitudinal robustness of variables predicting independent gait following severe middle cerebral artery stroke: a prospective cohort study. *Clinical Rehabilitation*, 20, 262-8.

KOLLEN, B., VAN DE PORT, I., LINDEMAN, E., TWISK, J. & KWAKKEL, G. 2005. Predicting improvement in gait after stroke: a longitudinal prospective study. *Stroke*, 36, 2676-80.

KOSIOR, J. C., IDRIS, S., DOWLATSHAHI, D., ALZAWAHMAH, M., EESA, M., SHARMA, P., TYMCHUK, S., HILL, M. D., AVIV, R. I., FRAYNE, R., DEMCHUK, A. M. & INVESTIGATORS, P. R. S. C. I. S. 2011. Quantomo: validation of a computer-assisted methodology for the volumetric analysis of intracerebral haemorrhage. *International Journal of Stroke*, 6, 302-5.

KOSKI, L., XIE, H. & FINCH, L. 2009. Measuring cognition in a geriatric outpatient clinic: Rasch analysis of the Montreal Cognitive Assessment. *Journal of Geriatric Psychiatry & Neurology*, 22, 151-60.

KOTHARI, R. U., BROTT, T., BRODERICK, J. P., BARSAN, W. G., SAUERBECK, L. R., ZUCCARELLO, M. & KHOURY, J. 1996. The ABCs of measuring intracerebral hemorrhage volumes. *Stroke*, 27, 1304-5.

KOTILA, M., NUMMINEN, H., WALTIMO, O. & KASTE, M. 1998. Depression after stroke: results of the FINNSTROKE Study. *Stroke*, 29, 368-72.

KOZLOWSKI, A. J., CELLA, D., NITSCH, K. P. & HEINEMANN, A. W. Evaluating Individual Change With the Quality of Life in Neurological Disorders (Neuro-QoL) Short Forms. *Archives of Physical Medicine & Rehabilitation*, 97, 650-654.e8.

KUTLUBAEV, M. A., DUNCAN, F. H. & MEAD, G. E. 2012. Biological correlates of post-stroke fatigue: a systematic review. *Acta Neurologica Scandinavica*, 125, 219-27.

KUTLUBAEV, M. A. & HACKETT, M. L. 2014. Part II: predictors of depression after stroke and impact of depression on stroke outcome: an updated systematic review of observational studies. *International Journal of Stroke*, 9, 1026-36.

KWAH, L. K., HARVEY, L. A., DIONG, J. & HERBERT, R. D. 2013. Models containing age and NIHSS predict recovery of ambulation and upper limb function six months after stroke: an observational study.[Erratum appears in J Physiother. 2013 Dec;59(4):218]. *Journal of Physiotherapy*, 59, 189-97.

KWAK, R., KADOYA, S. & SUZUKI, T. 1983. Factors affecting the prognosis in thalamic hemorrhage. *Stroke*, 14, 493-500.

KWAKKEL, G., KOLLEN, B. & TWISK, J. 2006. Impact of time on improvement of outcome after stroke. *Stroke*, 37, 2348-53.

KWAKKEL, G. & KOLLEN, B. J. 2013. Predicting activities after stroke: what is clinically relevant? *International Journal of Stroke*, 8, 25-32.

KWAKKEL, G., KOLLEN, B. J. & KREBS, H. I. 2008. Effects of robot-assisted therapy on upper limb recovery after stroke: a systematic review. *Neurorehabilitation & Neural Repair*, 22, 111-21.

KWAKKEL, G., VAN DIJK, G. M. & WAGENAAR, R. C. 2000. Accuracy of physical and occupational therapists' early predictions of recovery after severe middle cerebral artery stroke. *Clinical Rehabilitation*, 14, 28-41.

LABARERE, J., RENAUD, B. & FINE, M. J. 2014. How to derive and validate clinical prediction models for use in intensive care medicine.[Erratum appears in *Intensive Care Med.* 2014 Jun;40(6):925 Note: Bertrand, Renaud [corrected to Renaud, Bertrand]]. *Intensive Care Medicine*, 40, 513-27.

LANGHORNE, P., COUPAR, F. & POLLOCK, A. 2009. Motor recovery after stroke: a systematic review. *Lancet Neurology*, 8, 741-54.

LEAL, J., LUENGO-FERNANDEZ, R., GRAY, A., PETERSEN, S. & RAYNER, M. 2006. Economic burden of cardiovascular diseases in the enlarged European Union. *European Heart Journal*, 27, 1610-9.

LEBLOIS, A., BORAUD, T., MEISSNER, W., BERGMAN, H. & HANSEL, D. 2006. Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. *Journal of Neuroscience*, 26, 3567-83.

LECOUTURIER, J., RODGERS, H., MURTAGH, M. J., WHITE, M., FORD, G. A. & THOMSON, R. G. 2010. Systematic review of mass media interventions designed to improve public recognition of stroke symptoms, emergency response and early treatment. *BMC Public Health*, 10, 784.

LEE, J. S., HAN, M. K., KIM, S. H., KWON, O. K. & KIM, J. H. 2005. Fiber tracking by diffusion tensor imaging in corticospinal tract stroke: Topographical correlation with clinical symptoms. *Neuroimage*, 26, 771-6.

LEES, K. R., BATH, P. M., SCHELLINGER, P. D., KERR, D. M., FULTON, R., HACKE, W., MATCHAR, D., SEHRA, R., TONI, D. & EUROPEAN STROKE ORGANIZATION OUTCOMES WORKING, G. 2012. Contemporary outcome measures in acute stroke research: choice of primary outcome measure. *Stroke*, 43, 1163-70.

LEIRA, R., DAVALOS, A., SILVA, Y., GIL-PERALTA, A., TEJADA, J., GARCIA, M., CASTILLO, J. & STROKE PROJECT, C. D. G. O. T. S. N. S. 2004. Early neurologic deterioration in intracerebral hemorrhage: predictors and associated factors. *Neurology*, 63, 461-7.

LENNON, S. & JOHNSON, L. 2000. The modified rivermead mobility index: validity and reliability. *Disability & Rehabilitation*, 22, 833-9.

LISK, D. R., PASTEUR, W., RHOADES, H., PUTNAM, R. D. & GROTTA, J. C. 1994. Early presentation of hemispheric intracerebral hemorrhage: prediction of outcome and guidelines for treatment allocation. *Neurology*, 44, 133-9.

LOKK, J., ROGHANI, R. & DELBARI, A. 2011. Effect of methylphenidate and/or levodopa coupled with physiotherapy on functional and motor recovery after stroke--A randomized, double-blind, placebo-controlled trial. *Acta Neurologica Scandinavica*, 123, pp.

LONGSTRETH, W. T., JR., BERNICK, C., MANOLIO, T. A., BRYAN, N., JUNGREIS, C. A. & PRICE, T. R. 1998. Lacunar infarcts defined by magnetic resonance imaging of 3660 elderly people: the Cardiovascular Health Study. *Archives of Neurology*, 55, 1217-25.

LOUIS, E. D. 1997. The shaking palsy, the first forty-five years: a journey through the British literature. *Movement Disorders*, 12, 1068-72.

LU, W.-S., HUANG, S.-L., YANG, J.-F., CHEN, M.-H., HSIEH, C.-L. & CHOU, C.-Y. Convergent validity and responsiveness of the EQ-5D utility weights for stroke survivors. *Journal of Rehabilitation Medicine*, 48, 346-51.

LYNCH, J., MEAD, G., GREIG, C., YOUNG, A., LEWIS, S. & SHARPE, M. 2007. Fatigue after stroke: the development and evaluation of a case definition. *Journal of Psychosomatic Research*, 63, 539-44.

MACELLARI, F., PACIARONI, M., AGNELLI, G. & CASO, V. 2014. Neuroimaging in intracerebral hemorrhage. *Stroke*, 45, 903-8.

MAHONEY, F. I. & BARTHEL, D. W. 1965. Functional Evaluation: The Barthel Index. *Maryland State Medical Journal*, 14, 61-5.

MAKIN, S. D., TURPIN, S., DENNIS, M. S. & WARDLAW, J. M. 2013. Cognitive impairment after lacunar stroke: systematic review and meta-analysis of incidence, prevalence and comparison with other stroke subtypes. *Journal of Neurology, Neurosurgery & Psychiatry*, 84, 893-900.

MASE, G., ZORZON, M., BIASUTTI, E., TASCA, G., VITRANI, B. & CAZZATO, G. 1995. Immediate prognosis of primary intracerebral hemorrhage using an easy model for the prediction of survival. *Acta Neurologica Scandinavica*, 91, 306-9.

MAYER, S. A., BRUN, N. C., BEGTRUP, K., BRODERICK, J., DAVIS, S., DIRINGER, M. N., SKOLNICK, B. E., STEINER, T. & ON BEHALF OF THE RECOMBINANT ACTIVATED FACTOR VII INTRACEREBRAL HEMORRHAGE TRIAL INVESTIGATORS 2005a. Recombinant activated factor VII for acute intracerebral hemorrhage. *New England Journal of Medicine*, 352, 777-85.

MAYER, S. A., BRUN, N. C., BRODERICK, J., DAVIS, S., DIRINGER, M. N., SKOLNICK, B. E., STEINER, T. & EUROPE/AUSTRALASIA NOVOSEVEN, I. C. H. T. I. 2005b. Safety and feasibility of recombinant factor VIIa for acute intracerebral hemorrhage. *Stroke*, 36, 74-9.

MAYER, S. A., SACCO, R. L., SHI, T. & MOHR, J. P. 1994. Neurologic deterioration in noncomatose patients with supratentorial intracerebral hemorrhage. *Neurology*, 44, 1379-84.

MCHAFFIE, J. G., STANFORD, T. R., STEIN, B. E., COIZET, V. & REDGRAVE, P. 2005. Subcortical loops through the basal ganglia. *Trends in Neurosciences*, 28, 401-7.

MEAD, G., LYNCH, J., GREIG, C., YOUNG, A., LEWIS, S. & SHARPE, M. 2007. Evaluation of fatigue scales in stroke patients. *Stroke*, 38, 2090-5.

MEAD, G. E., LEWIS, S. C., WARDLAW, J. M., DENNIS, M. S. & WARLOW, C. P. 1999. Should computed tomography appearance of lacunar stroke influence patient management? *Journal of Neurology, Neurosurgery & Psychiatry*, 67, 682-4.

MEAD, G. E., LEWIS, S. C., WARDLAW, J. M., DENNIS, M. S. & WARLOW, C. P. 2000. How well does the Oxfordshire community stroke project classification predict the site and size of the infarct on brain imaging? *Journal of Neurology, Neurosurgery & Psychiatry*, 68, 558-62.

MEIJER, R., VAN LIMBEEK, J., KRIEK, B., IHNFELDT, D., VERMEULEN, M. & DE HAAN, R. 2004. Prognostic social factors in the subacute phase after a stroke for the discharge destination from the hospital stroke-unit. A systematic review of the literature. *Disability & Rehabilitation*, 26, 191-7.

MENDELOW, A. D., GREGSON, B. A., FERNANDES, H. M., MURRAY, G. D., TEASDALE, G. M., HOPE, D. T., KARIMI, A., SHAW, M. D., BARER, D. H. & INVESTIGATORS, S. 2005. Early surgery versus initial conservative treatment in patients with spontaneous supratentorial intracerebral haematomas in the International Surgical Trial in Intracerebral Haemorrhage (STICH): a randomised trial. *Lancet*, 365, 387-97.

MENDELOW, A. D., GREGSON, B. A., MITCHELL, P. M., MURRAY, G. D., ROWAN, E. N., GHOLKAR, A. R. & INVESTIGATORS, S. I. 2011. Surgical trial in lobar intracerebral haemorrhage (STICH II) protocol. *Trials [Electronic Resource]*, 12, 124.

MENDELOW, A. D., GREGSON, B. A., ROWAN, E. N., MURRAY, G. D., GHOLKAR, A., MITCHELL, P. M. & INVESTIGATORS, S. I. 2013. Early surgery versus initial conservative treatment in patients with spontaneous supratentorial lobar intracerebral haematomas (STICH II): a randomised trial.[Erratum appears in *Lancet*. 2013 Aug 3;382(9890):396]. *Lancet*, 382, 397-408.

MENDIS, S. 2010. Prevention and care of stroke in low- and middle-income countries; the need for a public health perspective. *International Journal of Stroke*, 5, 86-91.

MENDIS, S. 2013. Stroke disability and rehabilitation of stroke: World Health Organization perspective. *International Journal of Stroke*, 8, 3-4.

MEUTH, S. G., KLEINSCHNITZ, C., BROICHER, T., AUSTINAT, M., BRAEUNINGER, S., BITTNER, S., FISCHER, S., BAYLISS, D. A., BUDDE, T., STOLL, G. & WIENDL, H. 2009. The neuroprotective impact of the leak potassium channel TASK1 on stroke development in mice. *Neurobiology of Disease*, 33, 1-11.

MEYER-EBRECHT, D. 1994. Picture archiving and communication systems (PACS) for medical application. *International Journal of Bio-Medical Computing*, 35, 91-124.

MEYER, S., VERHEYDEN, G., BRINKMANN, N., DEJAEGER, E., DE WEERDT, W., FEYS, H., GANTENBEIN, A. R., JENNI, W., LAENEN, A., LINCOLN, N., PUTMAN, K., SCHUBACK, B., SCHUPP, W., THIJS, V. & DE WIT, L. 2015. Functional and motor outcome 5 years after stroke is equivalent to outcome at 2 months: follow-up of the collaborative evaluation of rehabilitation in stroke across Europe. *Stroke*, 46, 1613-9.

MICHELSEN, H. J., DE VRIES, J. & VAN HECK, G. L. 2003. Psychometric qualities of a brief self-rated fatigue measure: The Fatigue Assessment Scale. *Journal of Psychosomatic Research*, 54, 345-52.

MICHELSEN, H. J., DE VRIES, J. & VAN HECK, G. L. 2004. Examination of the dimensionality of fatigue: the construction of the Fatigue Assessment Scale (FAS). *European Journal of Psychological Assessment*, 20, 39-48.

MILLER, C. S., GRANDI, S. M., SHIMONY, A., FILION, K. B. & EISENBERG, M. J. 2012. Meta-analysis of efficacy and safety of new oral anticoagulants (dabigatran, rivaroxaban, apixaban) versus warfarin in patients with atrial fibrillation. *American Journal of Cardiology*, 110, 453-60.

MOONS, K. G., ALTMAN, D. G., REITSMA, J. B., IOANNIDIS, J. P., MACASKILL, P., STEYERBERG, E. W., VICKERS, A. J., RANSOHOFF, D. F. & COLLINS, G. S. 2015. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Annals of Internal Medicine*, 162, W1-W73.

MOONS, K. G., ALTMAN, D. G., VERGOUWE, Y. & ROYSTON, P. 2009a. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ*, 338, b606.

MOONS, K. G., ROYSTON, P., VERGOUWE, Y., GROBBEE, D. E. & ALTMAN, D. G. 2009b. Prognosis and prognostic research: what, why, and how? *BMJ*, 338, b375.

MOORHOUSE, P. & ROCKWOOD, K. 2008. Vascular cognitive impairment: current concepts and clinical developments. *Lancet Neurology*, 7, 246-55.

MUIR, K. W., FORD, G. A., MESSOW, C. M., FORD, I., MURRAY, A., CLIFTON, A., BROWN, M. M., MADIGAN, J., LENTHALL, R., ROBERTSON, F., DIXIT, A., CLOUD, G. C., WARDLAW, J., FREEMAN, J., WHITE, P. & INVESTIGATORS, P. 2017. Endovascular therapy for acute ischaemic stroke: the Pragmatic Ischaemic Stroke Thrombectomy Evaluation (PISTE) randomised, controlled trial. *Journal of Neurology, Neurosurgery & Psychiatry*, 88, 38-44.

NAMBU, A. 2008. Seven problems on the basal ganglia. *Current Opinion in Neurobiology*, 18, 595-604.

NANNETTI, L., PACI, M., PASQUINI, J., LOMBARDI, B. & TAITI, P. G. 2005. Motor and functional recovery in patients with post-stroke depression. *Disability & Rehabilitation*, 27, 170-5.

NASREDDINE, Z. S., PHILLIPS, N. A., BEDIRIAN, V., CHARBONNEAU, S., WHITEHEAD, V., COLLIN, I., CUMMINGS, J. L. & CHERTKOW, H. 2005. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53, 695-9.

NATIONAL ELECTRICAL MANUFACTURERS ASSOCIATION 2011. Digital Communications and Imaging in Medicine (DICOM). PS 3.1. Available online at <http://dicom.nema.org/>.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE 2013. *Stroke rehabilitation: long-term rehabilitation after stroke*. NICE Clinical Guideline 162. Available online at guidance.nice.org.uk/cg162

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE 2016. NICE pathways: stroke rehabilitation. Available online at <http://pathways.nice.org.uk/pathways/stroke>.

NHS ENGLAND 2015. 2015/16 Enhanced Tariff Option spreadsheet. Available online at <https://www.england.nhs.uk/?s=ETO+spreadsheet%2C+which+includes+prices+-+NHS+England&search=>.

NILSSON, O. G., LINDGREN, A., BRANDT, L. & SAVELAND, H. 2002. Prediction of death in patients with primary intracerebral hemorrhage: a prospective study of a defined population. *Journal of Neurosurgery*, 97, 531-6.

NOURI, F. M. & LINCOLN, N. B. 1987. An extended ADL scale for use with stroke patients. *Clinical Rehabilitation* 1, 301-305.

NUDO, R. J. 2008. Neuropsychology of motor skill learning. *Learning and memory: a comprehensive reference*, 3, 403-421.

NUTT, J. G. & FELLMAN, J. H. 1984. Pharmacokinetics of levodopa. *Clinical Neuropharmacology*, 7, 35-49.

NYS, G. M., VAN ZANDVOORT, M. J., DE KORT, P. L., JANSEN, B. P., KAPPELLE, L. J. & DE HAAN, E. H. 2005. Restrictions of the Mini-Mental State Examination in acute stroke. *Archives of Clinical Neuropsychology*, 20, 623-9.

O'BRIEN, J. T., ERKINJUNTTI, T., REISBERG, B., ROMAN, G., SAWADA, T., PANTONI, L., BOWLER, J. V., BALLARD, C., DECARLI, C., GORELICK, P. B., ROCKWOOD, K., BURNS, A., GAUTHIER, S. & DEKOSKY, S. T. 2003. Vascular cognitive impairment. *The Lancet Neurology*, 2, 89-98.

O'BRIEN, R. M. 2007. A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41, 673-690.

O'CONNOR, R. J., BEDEN, R., PILLING, A. & CHAMBERLAIN, M. A. 2011. What reductions in dependency costs result from treatment in an inpatient neurological rehabilitation unit for people with stroke? *Clinical Medicine*, 11, 40-3.

O'CONNOR, R. J., JACKSON, A., MAKOWER, S. G., COZENS, A. & LEVESLEY, M. 2014. A proof of concept study investigating the feasibility of combining iPAM robot assisted rehabilitation with functional electrical stimulation to deliver whole arm exercise in stroke survivors. *Journal of Medical Engineering & Technology*, 39, 411-8.

PALLANT, J. 2010. SPSS Survival Manual. McGraw -Hill; New York. 4th Edition.

PAOLUCCI, S., ANTONUCCI, G., GRASSO, M. G., BRAGONI, M., COIRO, P., DE ANGELIS, D., FUSCO, F. R., MORELLI, D., VENTURIERO, V., TROISI, E. & PRATESI, L. 2003. Functional outcome of ischemic and hemorrhagic stroke patients after inpatient rehabilitation: a matched comparison. *Stroke*, 34, 2861-5.

PAOLUCCI, S., BRAGONI, M., COIRO, P., DE ANGELIS, D., FUSCO, F. R., MORELLI, D., VENTURIERO, V. & PRATESI, L. 2008. Quantification of the probability of reaching mobility independence at discharge from a rehabilitation hospital in nonwalking early ischemic stroke patients: a multivariate study. *Cerebrovascular Diseases*, 26, 16-22.

PARKINSON, J. 1817. An essay on the shaking palsy. *Originally published as a monograph by Sherwood, Neely, and Jones (London, 1817). Reprint in: Journal of Neuropsychiatry & Clinical Neurosciences*, 2002, 14, 223-36; discussion 222.

PARRY-JONES, A. R., ABID, K. A., DI NAPOLI, M., SMITH, C. J., VAIL, A., PATEL, H. C., KING, A. T. & TYRRELL, P. J. 2013. Accuracy and clinical usefulness of intracerebral hemorrhage grading scores: a direct comparison in a UK population. *Stroke*, 44, 1840-5.

PARTRIDGE, J. G. & MORRIS, L. W. 1993. Recovery from physical disability after stroke: profiles for different levels of starting severity. *Clinical Rehabilitation*, 7, 210-217.

PASSERO, S., ULIVELLI, M. & REALE, F. 2002. Primary intraventricular haemorrhage in adults. *Acta Neurologica Scandinavica*, 105, 115-9.

PATEL, A. T., DUNCAN, P. W., LAI, S. M. & STUDENSKI, S. 2000. The relation between impairments and functional outcomes poststroke. *Archives of Physical Medicine & Rehabilitation*, 81, 1357-63.

PENDLEBURY, S. T., CUTHBERTSON, F. C., WELCH, S. J., MEHTA, Z. & ROTHWELL, P. M. 2010. Underestimation of cognitive impairment by mini-mental state examination versus the montreal cognitive assessment in patients with transient ischemic attack and stroke: a population-based study. *Stroke*, 41, 1290-3.

PENDLEBURY, S. T., MARIZ, J., BULL, L., MEHTA, Z. & ROTHWELL, P. M. 2012. MoCA, ACE-R, and MMSE versus the National Institute of Neurological Disorders and Stroke-Canadian Stroke Network Vascular Cognitive Impairment Harmonization Standards Neuropsychological Battery after TIA and stroke. *Stroke*, 43, 464-9.

PENHUNE, V. B. & STEELE, C. J. 2012. Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. *Behavioural Brain Research*, 226, 579-91.

PENTA, M., THONNARD, J. L. & TESIO, L. 1998. ABILHAND: a Rasch-built measure of manual ability. *Archives of Physical Medicine & Rehabilitation*, 79, 1038-42.

PHAN, T. G., KOH, M., VIERKANT, R. A. & WIJCK, E. F. 2000. Hydrocephalus is a determinant of early mortality in putaminal hemorrhage. *Stroke*, 31, 2157-62.

PITTOCK, S. J., MELDRUM, D., HARDIMAN, O., THORNTON, J., BRENNAN, P. & MORONEY, J. T. 2003. The Oxfordshire Community Stroke Project classification: correlation with imaging, associated complications, and prediction of outcome in acute ischemic stroke. *Journal of Stroke & Cerebrovascular Diseases*, 12, 1-7.

POHJASVAARA, T., LEPPAVUORI, A., SIIRA, I., VATAJA, R., KASTE, M. & ERKINJUNTTI, T. 1998. Frequency and clinical determinants of poststroke depression. *Stroke*, 29, 2311-7.

POLLOCK, A., BAER, G., POMEROY, V. & LANGHORNE, P. 2007. Physiotherapy treatment approaches for the recovery of postural control and lower limb function following stroke. *Cochrane Database of Systematic Reviews*, CD001920.

POLLOCK, A., FARMER, S. E., BRADY, M. C., LANGHORNE, P., MEAD, G. E., MEHRHOLZ, J. & VAN WIJCK, F. 2014. Interventions for improving upper limb function after stroke. *Cochrane Database of Systematic Reviews*, Issue 11. Art. No.: CD010820. DOI: 10.1002/14651858.CD010820.pub2.

POLLOCK, A., ST GEORGE, B., FENTON, M. & FIRKINS, L. 2012. Top ten research priorities relating to life after stroke. *Lancet Neurology*, 11, 209.

PORTENOY, R. K., LIPTON, R. B., BERGER, A. R., LESSER, M. L. & LANTOS, G. 1987. Intracerebral haemorrhage: a model for the prediction of outcome. *Journal of Neurology, Neurosurgery & Psychiatry*, 50, 976-9.

POTTER, G. M., DOUBAL, F. N., JACKSON, C. A., CHAPPELL, F. M., SUDLOW, C. L., DENNIS, M. S. & WARDLAW, J. M. 2010. Counting cavitating lacunes underestimates the burden of lacunar infarction. *Stroke*, 41, 267-72.

PRESTON, E., ADA, L., DEAN, C. M., STANTON, R. & WADDINGTON, G. 2011. What is the probability of patients who are nonambulatory after stroke regaining independent walking? A systematic review. *International Journal of Stroke*, 6, 531-40.

PUBLIC HEALTH ENGLAND 2018. Briefing document: First incidence of stroke. Estimates for England 2007 to 2016. Crown copyright. Available online at

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/678444/Stroke_incidence_briefing_document_2018.pdf.

PUBLIC HEALTH ENGLAND CAMPAIGNS RESOURCE CENTRE 2009. Act FAST. Available online at <http://campaigns.dh.gov.uk/category/act-fast/>.

PUTMAN, K., DE WIT, L., SCHUPP, W., BEYENS, H., DEJAEGER, E., DE WEERDT, W., FEYS, H., JENNI, W., LOUCKX, F. & LEYS, M. Inpatient stroke rehabilitation: a comparative study of admission criteria to stroke rehabilitation units in four European centres. *Journal of Rehabilitation Medicine*, 39, 21-6.

QUINN, T. J., DAWSON, J., WALTERS, M. R. & LEES, K. R. 2009a. Exploring the reliability of the modified rankin scale. *Stroke*, 40, 762-6.

QUINN, T. J., DAWSON, J., WALTERS, M. R. & LEES, K. R. 2009b. Functional outcome measures in contemporary stroke trials. *International Journal of Stroke*, 4, 200-5.

QUINN, T. J., DAWSON, J., WALTERS, M. R. & LEES, K. R. 2009c. Reliability of the modified Rankin Scale: a systematic review. *Stroke*, 40, 3393-5.

QUINN, T. J., LANGHORNE, P. & STOTT, D. J. 2011. Barthel index for stroke trials: development, properties, and application. *Stroke*, 42, 1146-51.

QUINN, T. J., LEES, K. R., HARDEMARK, H. G., DAWSON, J. & WALTERS, M. R. 2007. Initial experience of a digital training resource for modified Rankin scale assessment in clinical trials. *Stroke*, 38, 2257-61.

QURESHI, A. I., MENDELOW, A. D. & HANLEY, D. F. 2009. Intracerebral haemorrhage. *Lancet*, 373, 1632-44.

QURESHI, A. I., PALESCH, Y. Y., BARSAN, W. G., HANLEY, D. F., HSU, C. Y., MARTIN, R. L., MOY, C. S., SILBERGLEIT, R., STEINER, T., SUAREZ, J. I., TOYODA, K., WANG, Y., YAMAMOTO, H., YOON, B. W., INVESTIGATORS, A.-T. & THE NEUROLOGICAL EMERGENCY TREATMENT TRIALS, N. 2016. Intensive Blood-Pressure Lowering in Patients with Acute Cerebral Hemorrhage. *New England Journal of Medicine*, 375, 1033-43.

QURESHI, A. I., SAFDAR, K., WEIL, J., BARCH, C., BLIWISE, D. L., COLOHAN, A. R., MACKAY, B. & FRANKEL, M. R. 1995. Predictors of early deterioration and mortality in black Americans with spontaneous intracerebral hemorrhage. *Stroke*, 26, 1764-7.

RABIN, R. & DE CHARRO, F. 2001. EQ-5D: a measure of health status from the EuroQol Group. *Annals of Medicine*, 33, 337-43.

RANKIN, J. 1957. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. *Scottish Medical Journal*, 2, 200-15.

RASHID, P., LEONARDI-BEE, J. & BATH, P. 2003. Blood pressure reduction and secondary prevention of stroke and other vascular events: a systematic review. *Stroke*, 34, 2741-8.

RATIB, O., LIGIER, Y. & SCHERRER, J. R. 1994. Digital image management and communication in medicine. *Computerized Medical Imaging & Graphics*, 18, 73-84.

RAZZAQ, A. A. & HUSSAIN, R. 1998. Determinants of 30-day mortality of spontaneous intracerebral hemorrhage in Pakistan. *Surgical Neurology*, 50, 336-42; discussion 342-3.

REID, J. M., GUBITZ, G. J., DAI, D., KYDD, D., ESKES, G., REIDY, Y., CHRISTIAN, C., COUNSELL, C. E., DENNIS, M. & PHILLIPS, S. J. 2010. Predicting functional outcome after stroke by modelling baseline clinical and CT variables. *Age & Ageing*, 39, 360-6.

RESTEMEYER, C., WEILLER, C. & LIEPERT, J. 2007. No effect of a levodopa single dose on motor performance and motor excitability in chronic stroke. A double-blind placebo-controlled cross-over pilot study. *Restorative Neurology & Neuroscience*, 25, 143-50.

ROMAN, G. C., ERKINJUNTTI, T., WALLIN, A., PANTONI, L. & CHUI, H. C. 2002. Subcortical ischaemic vascular dementia. *Lancet Neurology*, 1, 426-36.

ROSSER, N., HEUSCHMANN, P., WERSCHING, H., BREITENSTEIN, C., KNECHT, S. & FLOEL, A. 2008a. Levodopa improves procedural motor learning in chronic stroke patients. *Archives of Physical Medicine & Rehabilitation*, 89, 1633-41.

ROSSER, N., HEUSCHMANN, P., WERSCHING, H., BREITENSTEIN, C., KNECHT, S., FLOEL, A., ROSSER, N., HEUSCHMANN, P., WERSCHING, H., BREITENSTEIN, C., KNECHT, S. & FLOEL, A. 2008b. Levodopa improves procedural motor learning in chronic stroke patients. *Archives of Physical Medicine & Rehabilitation*, 89, 1633-41.

ROTHWELL, P. M. 2005. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet*, 365, 176-86.

RUIZ-SANDOVAL, J. L., CHIQUETE, E., ROMERO-VARGAS, S., PADILLA-MARTINEZ, J. J. & GONZALEZ-CORNEJO, S. 2007. Grading scale for prediction of outcome in primary intracerebral hemorrhages. *Stroke*, 38, 1641-4.

SACCO, R. L., KASNER, S. E., BRODERICK, J. P., CAPLAN, L. R., CONNORS, J. J., CULEBRAS, A., ELKIND, M. S., GEORGE, M. G., HAMDAN, A. D., HIGASHIDA, R. T., HOH, B. L., JANIS, L. S., KASE, C. S., KLEINDORFER, D. O., LEE, J. M., MOSELEY, M. E., PETERSON, E. D., TURAN, T. N., VALDERRAMA, A. L., VINTERS, H. V., ON BEHALF OF THE AMERICAN HEART ASSOCIATION STROKE COUNCIL, COUNCIL ON CARDIOVASCULAR SURGERY AND ANESTHESIA, COUNCIL ON CARDIOVASCULAR RADIOLOGY AND INTERVENTION, COUNCIL ON CARDIOVASCULAR AND STROKE NURSING, COUNCIL ON EPIDEMIOLOGY AND PREVENTION, COUNCIL ON PERIPHERAL VASCULAR DISEASE & AND THE COUNCIL ON NUTRITION PHYSICAL ACTIVITY AND METABOLISM 2013. An updated definition of stroke for the 21st century: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, 44, 2064-89.

SACCO, S., MARINI, C., TONI, D., OLIVIERI, L. & CAROLEI, A. 2009. Incidence and 10-year survival of intracerebral hemorrhage in a population-based registry. *Stroke*, 40, 394-9.

SANCHEZ-BLANCO, I., OCHOA-SANGRADOR, C., LOPEZ-MUNAIN, L., IZQUIERDO-SANCHEZ, M. & FERMOSO-GARCIA, J. 1999. Predictive model of functional independence in stroke patients admitted to a rehabilitation programme. *Clinical Rehabilitation*, 13, 464-75.

SANDERCOCK, P., COUNSELL, C., TSENG, M. C. & CECCONI, E. 2014. Oral antiplatelet therapy for acute ischaemic stroke. *The Cochrane Database of Systematic Reviews*, Article number CD000029.

SANDERCOCK, P., WARDLAW, J. M., LINDLEY, R. I., DENNIS, M., COHEN, G., MURRAY, G., INNES, K., VENABLES, G., CZLONKOWSKA, A., KOBAYASHI, A., RICCI, S., MURRAY, V., BERGE, E., SLOT, K. B., HANKEY, G. J., CORREIA, M., PEETERS, A., MATZ, K., LYRER, P., GUBITZ, G., PHILLIPS, S. J., ARAUZ, A., GROUP, I. S. T. C., SANDERCOCK, P., WARDLAW, J. M., LINDLEY, R. I., DENNIS, M., COHEN, G., MURRAY, G., INNES, K., VENABLES, G., CZLONKOWSKA, A., KOBAYASHI, A., RICCI, S., MURRAY, V., BERGE, E., SLOT, K. B., HANKEY, G. J., CORREIA, M., PEETERS, A., MATZ, K., LYRER, P., GUBITZ, G., PHILLIPS, S. J. & ARAUZ, A. 2012. The benefits and harms of intravenous thrombolysis with recombinant tissue plasminogen activator within 6 h of acute ischaemic stroke (the third international stroke trial [IST-3]): a randomised controlled trial. *Lancet*, 379, 2352-63.

SAUR, D., RONNEBERGER, O., KUMMERER, D., MADER, I., WEILLER, C. & KLOPPEL, S. 2010. Early functional magnetic resonance imaging activations predict language outcome after stroke. *Brain*, 133, 1252-64.

SAVER JL, JOHNSTON KC, HOMER D, WITYK R, W, K., TRUSKOWSKI LL & EC, H. 1999. Infarct Volume as a Surrogate or Auxiliary Outcome Measure in Ischemic Stroke Clinical Trials. *Stroke*, 30, 293-298.

SAVER, J. L. & ALTMAN, H. 2012. Relationship between neurologic deficit severity and final functional outcome shifts and strengthens during first hours after onset. *Stroke*, 43, 1537-41.

SAVER, J. L., GOYAL, M., BONAFE, A., DIENER, H. C., LEVY, E. I., PEREIRA, V. M., ALBERS, G. W., COGNARD, C., COHEN, D. J., HACKE, W., JANSEN, O., JOVIN, T. G., MATTLE, H. P., NOGUEIRA, R. G., SIDDIQUI, A. H., YAVAGAL, D. R., BAXTER, B. W., DEVLIN, T. G., LOPES, D. K., REDDY, V. K., DU MESNIL DE ROCHEMONT, R., SINGER, O. C., JAHAN, R. & INVESTIGATORS, S. P. 2015. Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke. *New England Journal of Medicine*, 372, 2285-95.

SCHEIDTMANN, K., FRIES, W., MULLER, F. & KOENIG, E. 2001. Effect of levodopa in combination with physiotherapy on functional motor recovery after stroke: a prospective, randomised, double-blind study. *Lancet*, 358, 787-90.

SCHMAHMANN, J. D., MACMORE, J. & VANGEL, M. 2009. Cerebellar stroke without motor deficit: clinical evidence for motor and non-motor domains within the human cerebellum. *Neuroscience*, 162, 852-61.

SCHMAHMANN, J. D. & PANDYA, D. N. 2008. Disconnection syndromes of basal ganglia, thalamus, and cerebrocerebellar systems. *Cortex*, 44, 1037-66.

SCHWARZ, S., HAFNER, K., ASCHOFF, A. & SCHWAB, S. 2000. Incidence and prognostic significance of fever following intracerebral hemorrhage. *Neurology*, 54, 354-61.

SEMANS, S. 1967. The Bobath concept in treatment of neurological disorders; a neuro-developmental treatment. *American Journal of Physical Medicine*, 46, 732-88.

SHAYA, M., DUBEY, A., BERK, C., GONZALEZ-TOLEDO, E., ZHANG, J., CALDITO, G. & NANDA, A. 2005. Factors influencing outcome in intracerebral hematoma: a simple, reliable, and accurate method to grade intracerebral hemorrhage.[Erratum appears in Surg Neurol. 2005 Sep;64(3):279]. *Surgical Neurology*, 63, 343-8; discussion 348.

SHUM, S. T., CHIU, J. K., TSANG, C. P., WONG, C. H., TSANG, R. C., MA, S. L., HO, Y. L. & TSANG, P. 2014. Predicting walking function of patients one month poststroke using

modified Rivermead mobility index on admission. *Journal of Stroke & Cerebrovascular Diseases*, 23, 2117-21.

SIVAN, M., GALLAGHER, J., MAKOWER, S., KEELING, D., BHAKTA, B., O'CONNOR, R. J. & LEVESLEY, M. 2014. Home-based Computer Assisted Arm Rehabilitation (hCAAR) robotic device for upper limb exercise after stroke: results of a feasibility study in home setting. *Journal of Neuroengineering & Rehabilitation*, 11, 163.

SLADE, A., TENNANT, A. & CHAMBERLAIN, M. A. 2002. A randomised controlled trial to determine the effect of intensity of therapy upon length of stay in a neurological rehabilitation setting. *Journal of Rehabilitation Medicine*, 34, 260-6.

SLOAN, M. A., PRICE, T. R., PETITO, C. K., RANDALL, A. M., SOLOMON, R. E., TERRIN, M. L., GORE, J., COLLEN, D., KLEIMAN, N., FEIT, F. & ET AL. 1995. Clinical features and pathogenesis of intracerebral hemorrhage after rt-PA and heparin therapy for acute myocardial infarction: the Thrombolysis in Myocardial Infarction (TIMI) II Pilot and Randomized Clinical Trial combined experience. *Neurology*, 45, 649-58.

SOMMERFELD, D. K. & VON ARBIN, M. H. 2001. Disability test 10 days after acute stroke to predict early discharge home in patients 65 years and older. *Clinical Rehabilitation*, 15, 528-34.

SONDE, L. & LOKK, J. 2007. Effects of amphetamine and/or L-dopa and physiotherapy after stroke - a blinded randomized study. *Acta Neurologica Scandinavica*, 115, 55-9.

STAUB, F. & BOGOUSLAVSKY, J. 2001. Fatigue after stroke: a major but neglected issue. *Cerebrovascular Diseases*, 12, 75-81.

STEINER, T., DIRINGER, M. N., SCHNEIDER, D., MAYER, S. A., BEGTRUP, K., BRODERICK, J., SKOLNICK, B. E. & DAVIS, S. M. 2006. Dynamics of intraventricular hemorrhage in patients with spontaneous intracerebral hemorrhage: risk factors, clinical impact, and effect of hemostatic therapy with recombinant activated factor VII. *Neurosurgery*, 59, 767-73; discussion 773-4.

STEYERBERG, E. W., MOONS, K. G., VAN DER WINDT, D. A., HAYDEN, J. A., PEREL, P., SCHROTER, S., RILEY, R. D., HEMINGWAY, H., ALTMAN, D. G. & GROUP, P. 2013. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Medicine / Public Library of Science*, 10, e1001381.

STINEAR, C. M., BARBER, P. A., PETOE, M., ANWAR, S. & BYBLOW, W. D. 2012. The PREP algorithm predicts potential for upper limb recovery after stroke. *Brain*, 135, 2527-35.

STINEAR, C. M. & WARD, N. S. 2013. How useful is imaging in predicting outcomes in stroke rehabilitation? *International Journal of Stroke*, 8, 33-7.

STOLL, G., KLEINSCHNITZ, C. & NIESWANDT, B. 2008. Molecular mechanisms of thrombus formation in ischemic stroke: novel insights and targets for treatment. *Blood*, 112, 3555-62.

STOLTZFUS, J. C. 2011. Logistic regression: a brief primer. *Academic Emergency Medicine*, 18, 1099-104.

STROKE UNIT TRIALISTS COLLABORATION 2013. Organised inpatient (stroke unit) care for stroke. *Cochrane Database of Systematic Reviews*, 9, CD000197.

SUBRAMANIAN, S. K., MASSIE, C. L., MALCOLM, M. P. & LEVIN, M. F. 2010. Does provision of extrinsic feedback result in improved motor learning in the upper limb poststroke? A systematic review of the evidence. *Neurorehabilitation & Neural Repair*, 24, 113-24.

SULTER, G., STEEN, C. & DE KEYSER, J. 1999. Use of the Barthel index and modified Rankin scale in acute stroke trials. *Stroke*, 30, 1538-41.

TAKAHASHI, K., DOMEN, K., SAKAMOTO, T., TOSHIMA, M., OTAKA, Y., SETO, M., IRIE, K., HAGA, B., TAKEBAYASHI, T. & HACHISUKA, K. Efficacy of Upper Extremity Robotic Therapy in Subacute Poststroke Hemiplegia: An Exploratory Randomized Trial. *Stroke*, 47, 1385-8.

TANG, P. F., KO, Y. H., LUO, Z. A., YEH, F. C., CHEN, S. H. & TSENG, W. Y. 2010. Tract-specific and region of interest analysis of corticospinal tract integrity in subcortical ischemic stroke: reliability and correlation with motor function of affected lower extremity. *Ajnr: American Journal of Neuroradiology*, 31, 1023-30.

TENNANT, A. & CONAGHAN, P. G. 2007. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis & Rheumatism*, 57, 1358-62.

THE AMERICAN PSYCHIATRIC ASSOCIATION 2012. Diagnostic and Statistical Manual of Mental Disorders, Version V. Available online at <http://www.dsm5.org>.

THE EUROPEAN PARLIAMENT AND COUNCIL 2001. Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use. Amended by Regulation (EC) No 1901/2006 of the European Parliament and of the Council of 12 December 2006; Regulation (EC) No 596/2009 of the European Parliament and of the Council of 18 June 2009. Available online at http://ec.europa.eu/health/files/eudralex/vol-1/dir_2001_20/dir_2001_20_en.pdf.

THE IST COLLABORATIVE GROUP 2015. Association between brain imaging signs, early and late outcomes, and response to intravenous alteplase after acute ischaemic stroke in the third International Stroke Trial (IST-3): secondary analysis of a randomised controlled trial. *Lancet Neurology*, 14, 485-96.

THE ROYAL COLLEGE OF PHYSICIANS 2010. *Medical rehabilitation in 2011 and beyond. Report of a working party.*

THE ROYAL COLLEGE OF PHYSICIANS 2016. National clinical guideline for stroke, fifth edition. Prepared by the Intercollegiate Stroke Working Party. Available online at: <https://www.rcplondon.ac.uk/guidelines-policy/stroke-guidelines>.

THE ROYAL COLLEGE OF RADIOLOGISTS 2011. Standards and recommendations for the reporting and interpretation of imaging investigations by non-radiologist medically qualified practitioners and teleradiologists.

THE STATIONARY OFFICE 1998. Data Protection Act 1998. Published by the Controller of Her Majesty's Stationery Office and Queen's Printer of Acts of Parliament. Available online at <http://www.legislation.gov.uk>.

THE STATIONERY OFFICE 2006. The Medicines for Human Use (Clinical Trials) Amendment Regulations 2006 Number 1928. Published by Her Majesty's Stationery Office

and Queen's Printer of Acts of Parliament. Available online at http://www.legislation.gov.uk/ukxi/2006/1928/pdfs/ukxi_20061928_en.pdf.

THE STROKE ASSOCIATION 2013. Feeling overwhelmed: life after stroke. Available online at http://www.stroke.org.uk/sites/default/files/files/Feeling%20Overwhelmed_Final_2_5mb.pdf.

THE STROKE ASSOCIATION 2016. State of the Nation: stroke statistics, January 2016. Available online at https://www.stroke.org.uk/sites/default/files/stroke_statistics_2015.pdf.

THE STROKE ASSOCIATION 2017a. Current, future and avoidable costs of stroke in the UK. Executive summary Part 1: Burden of stroke in the next 20 years and potential returns from increased spending on research. Available online at https://www.stroke.org.uk/sites/default/files/costs_of_stroke_in_the_uk_report_-_executive_summary_part_1.pdf.

THE STROKE ASSOCIATION 2017b. Current, future and avoidable costs of stroke in the UK. Executive summary Part 2: Societal costs of stroke in the next 20 years and potential returns from increased spending on research. Available online at https://www.stroke.org.uk/sites/default/files/costs_of_stroke_in_the_uk_report_-_executive_summary_part_2.pdf.

THE STROKE ASSOCIATION 2017c. State of the Nation: stroke statistics, January 2017. Available online at https://www.stroke.org.uk/sites/default/files/state_of_the_nation_2017_final_1.pdf.

THE WORLD HEALTH ORGANISATION 2010. International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) Version for 2010. Chapter V: Mental and behavioural disorders.

THE WORLD HEALTH ORGANISATION 2011. World Report on Disability. Available online at http://www.who.int/disabilities/world_report/2011/en/index.html.

THE WORLD HEALTH ORGANISATION 2014a. Ageing and Life Course. Available online at <http://www.who.int/ageing/about/facts/en/>.

THE WORLD HEALTH ORGANISATION 2014b. Global Health Estimates 2014 Summary Tables: DALY by Cause, Age, and Sex 2000-2012. Available online at http://www.who.int/healthinfo/global_burden_disease/en/.

THE WORLD HEALTH ORGANISATION 2014c. Global Health Estimates 2014 Summary Tables: Deaths by Cause, Age, and Sex 2000-2012. Available online at http://www.who.int/healthinfo/global_burden_disease/en/

THE WORLD MEDICAL ASSOCIATION 2013. WMA Declaration of Helsinki - Ethical principles for research involving human subjects. Adopted by the 64th WMA General Assembly, Fortaleza, Brazil, October 2013. Available online at [http://www.wma.net/en/30publications/10policies/b3/index.html.pdf?print-media-type&footer-right=\[page\]/\[toPage\]](http://www.wma.net/en/30publications/10policies/b3/index.html.pdf?print-media-type&footer-right=[page]/[toPage]).

THOMAS, C. & BAKER, C. I. 2013. Teaching an adult brain new tricks: a critical review of evidence for training-dependent structural plasticity in humans. *Neuroimage*, 73, 225-36.

TOWNEND, E., TINSON, D., KWAN, J. & SHARPE, M. 2010. 'Feeling sad and useless': an investigation into personal acceptance of disability and its association with depression following stroke. *Clinical Rehabilitation*, 24, 555-64.

TUHRIM, S., DAMBROSIA, J. M., PRICE, T. R., MOHR, J. P., WOLF, P. A., HIER, D. B. & KASE, C. S. 1991. Intracerebral hemorrhage: external validation and extension of a model for prediction of 30-day survival. *Annals of Neurology*, 29, 658-63.

UNITED NATIONS 2009. United Nations Convention on the Rights of Persons with Disabilities. Available online at <http://www.un.org/disabilities/convention/conventionfull.shtml>.

UYTTENBOOGAART, M., STEWART, R. E., VROOMEN, P. C., DE KEYSER, J. & LUIJCKX, G. J. 2005. Optimizing cutoff scores for the Barthel index and the modified Rankin scale for defining outcome in acute stroke trials. *Stroke*, 36, 1984-7.

VAN DE PORT, I. G., KWAKKEL, G., SCHEPERS, V. P. & LINDEMAN, E. 2006a. Predicting mobility outcome one year after stroke: a prospective cohort study. *Journal of Rehabilitation Medicine*, 38, 218-23.

VAN DE PORT, I. G., KWAKKEL, G., VAN WIJK, I. & LINDEMAN, E. 2006b. Susceptibility to deterioration of mobility long-term after stroke: a prospective cohort study. *Stroke*, 37, 167-71.

VAN DE PORT, I. G. L., KWAKKEL, G., VAN WIJK, I. & LINDEMAN, E. 2006c. Susceptibility to deterioration of mobility long-term after stroke: a prospective cohort study. *Stroke*, 37, 167-71.

VAN SWIETEN, J. C., HIJDRA, A., KOUDSTAAL, P. J. & VAN GIJN, J. 1990. Grading white matter lesions on CT and MRI: a simple scale. *Journal of Neurology, Neurosurgery & Psychiatry*, 53, 1080-3.

VAN SWIETEN, J. C., KOUDSTAAL, P. J., VISSER, M. C., SCHOUTEN, H. J. & VAN GIJN, J. 1988. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke*, 19, 604-7.

VAN VLIET, P. M. & WULF, G. 2006. Extrinsic feedback for motor learning after stroke: what is the evidence? *Disability & Rehabilitation*, 28, 831-40.

VAN WIJK, I., ALGRA, A., VAN DE PORT, I. G., BEVAART, B. & LINDEMAN, E. 2006. Change in mobility activity in the second year after stroke in a rehabilitation population: who is at risk for decline? *Archives of Physical Medicine & Rehabilitation*, 87, 45-50.

VEERBEEK, J. M., VAN WEGEN, E. E., HARMELING-VAN DER WEL, B. C., KWAKKEL, G. & INVESTIGATORS, E. 2011. Is accurate prediction of gait in nonambulatory stroke patients possible within 72 hours poststroke? The EPOS study. *Neurorehabilitation & Neural Repair*, 25, 268-74.

VERDELHO, A., HENON, H., LEBERT, F., PASQUIER, F. & LEYS, D. 2004. Depressive symptoms after stroke and relationship with dementia: A three-year follow-up study. *Neurology*, 62, 905-11.

VERMEER, S. E., LONGSTRETH, W. T., JR. & KOUDSTAAL, P. J. 2007. Silent brain infarcts: a systematic review. *Lancet Neurology*, 6, 611-9.

VICKERS, A. J. & CRONIN, A. M. 2010. Everything you always wanted to know about evaluating prediction models (but were too afraid to ask). *Urology*, 76, 1298-301.

VOGT, G., LAAGE, R., SHUAIB, A., SCHNEIDER, A. & COLLABORATION, V. 2012. Initial lesion volume is an independent predictor of clinical stroke outcome at day 90: an analysis of the Virtual International Stroke Trials Archive (VISTA) database. *Stroke*, 43, 1266-72.

VON KUMMER, R. 1998. Effect of training in reading CT scans on patient selection for ECASS II. *Neurology*, 51, S50-2.

WADE, D. T. & HEWER, R. L. 1987. Functional abilities after stroke: measurement, natural history and prognosis. *Journal of Neurology, Neurosurgery & Psychiatry*, 50, 177-82.

WANDEL, A., JORGENSEN, H. S., NAKAYAMA, H., RAASCHOU, H. O. & OLSEN, T. S. 2000. Prediction of walking function in stroke patients with initial lower extremity paralysis: the Copenhagen Stroke Study. *Archives of Physical Medicine & Rehabilitation*, 81, 736-8.

WANG, T. J., MASSARO, J. M., LEVY, D., VASAN, R. S., WOLF, P. A., D'AGOSTINO, R. B., LARSON, M. G., KANNEL, W. B. & BENJAMIN, E. J. 2003. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community: the Framingham Heart Study. *JAMA*, 290, 1049-56.

WARD, N. S. & COHEN, L. G. 2004. Mechanisms underlying recovery of motor function after stroke. *Archives of Neurology*, 61, 1844-8.

WARDLAW, J. M. 2005. What causes lacunar stroke? *Journal of Neurology, Neurosurgery & Psychiatry*, 76, 617-9.

WARDLAW, J. M. 2008. What is a lacune? *Stroke*, 39, 2921-2.

WARDLAW, J. M., CARPENTER, T., SAKKA, E., MAIR, G., COHEN, G. & SHULER, K. 2014a. Imaging perfusion deficits, artery patency and thrombolysis safety and efficacy in acute ischaemic stroke. An observational study of the effects of advanced imaging methods in the Third International Stroke Trial (IST3), a randomised controlled trial. *Efficacy Mech Eval*, 1, DOI: 10.3310/eme01010.

WARDLAW, J. M., DORMAN, P. J., LEWIS, S. C. & SANDERCOCK, P. A. 1999. Can stroke physicians and neuroradiologists identify signs of early cerebral infarction on CT? *Journal of Neurology, Neurosurgery & Psychiatry*, 67, 651-3.

WARDLAW, J. M., FARRALL, A. J., PERRY, D., VON KUMMER, R., MIELKE, O., MOULIN, T., CICCONE, A., HILL, M. & ACUTE CEREBRAL, C. T. E. O. S. S. S. G. 2007. Factors influencing the detection of early CT signs of cerebral ischemia: an internet-based, international multiobserver study. *Stroke*, 38, 1250-6.

WARDLAW, J. M., LEWIS, S. C., DENNIS, M. S., COUNSELL, C. & MCDOWALL, M. 1998. Is visible infarction on computed tomography associated with an adverse prognosis in acute ischemic stroke? *Stroke*, 29, 1315-9.

WARDLAW, J. M. & MIELKE, O. 2005. Early signs of brain infarction at CT: observer reliability and outcome after thrombolytic treatment--systematic review. *Radiology*, 235, 444-53.

WARDLAW, J. M., MURRAY, V., BERGE, E. & DEL ZOPPO, G. J. 2014b. Thrombolysis for acute ischaemic stroke. *Cochrane Database of Systematic Reviews*, CD000213.

WARDLAW, J. M. & SELLAR, R. 1994. A simple practical classification of cerebral infarcts on CT and its interobserver reliability. *Ajnr: American Journal of Neuroradiology*, 15, 1933-9.

WARDLAW, J. M., SMITH, C. & DICHGANS, M. 2013a. Mechanisms of sporadic cerebral small vessel disease: insights from neuroimaging. *Lancet Neurology*, 12, 483-97.

WARDLAW, J. M., SMITH, E. E., BIESELS, G. J., CORDONNIER, C., FAZEKAS, F., FRAYNE, R., LINDLEY, R. I., O'BRIEN, J. T., BARKHOF, F., BENAVENTE, O. R., BLACK, S. E., BRAYNE, C., BRETELER, M., CHABRIAT, H., DECARLI, C., DE LEEUW, F. E., DOUBAL, F., DUERING, M., FOX, N. C., GREENBERG, S., HACHINSKI, V., KILIMANN, I., MOK, V., OOSTENBRUGGE, R., PANTONI, L., SPECK, O., STEPHAN, B. C., TEIPEL, S., VISWANATHAN, A., WERRING, D., CHEN, C., SMITH, C., VAN BUCHEM, M., NORRVING, B., GORELICK, P. B., DICHGANS, M. & NEUROIMAGING, S. T. F. R. V. C. O. 2013b. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurology*, 12, 822-38.

WARDLAW, J. M., VON KUMMER, R., FARRALL, A. J., CHAPPELL, F. M., HILL, M., PERRY, D. & WADIA, P. 2010. A large web-based observer reliability study of early ischaemic signs on computed tomography. The Acute Cerebral CT Evaluation of Stroke Study (ACCESS). *PLoS ONE [Electronic Resource]*, 5, e15757.

WERNICKE, C. 1970. The aphasic symptom-complex: a psychological study on an anatomical basis [translation of: Wernicke, C. Der aphasische Symptomencomplex: Eine Psychologische Studie auf anatomischer Basis. Breslau, Max Cohn & Weigert, 1874, pp 15, 16, 18, 19, 21, 22, 24, 25, 43-45.] *Archives of Neurology*, 22, 280-282.

WERRING, D. J., GREGOIRE, S. M. & CIPOLOTTI, L. 2010. Cerebral microbleeds and vascular cognitive impairment. *Journal of the Neurological Sciences*, 299, 131-5.

WHITELEY, W., LINDLEY, R., WARDLAW, J., SANDERCOCK, P. & GROUP, I. S. T. C. 2006. Third international stroke trial. *International Journal of Stroke*, 1, 172-6.

WILSON, J. T., HAREENDRAN, A., GRANT, M., BAIRD, T., SCHULZ, U. G., MUIR, K. W. & BONE, I. 2002. Improving the assessment of outcomes in stroke: use of a structured interview to assign grades on the modified Rankin Scale. *Stroke*, 33, 2243-6.

WILSON, J. T., HAREENDRAN, A., HENDRY, A., POTTER, J., BONE, I. & MUIR, K. W. 2005. Reliability of the modified Rankin Scale across multiple raters: benefits of a structured interview. *Stroke*, 36, 777-81.

WINTERMARK, M., ALBERS, G. W., ALEXANDROV, A. V., ALGER, J. R., BAMMER, R., BARON, J. C., DAVIS, S., DEMAERSCHALK, B. M., DERDEYN, C. P., DONNAN, G. A., EASTWOOD, J. D., FIEBACH, J. B., FISHER, M., FURIE, K. L., GOLDMAKHER, G. V., HACKE, W., KIDWELL, C. S., KLOSKA, S. P., KOHRMANN, M., KOROSHETZ, W., LEE, T. Y., LEES, K. R., LEV, M. H., LIEBESKIND, D. S., OSTERGAARD, L., POWERS, W. J., PROVENZALE, J., SCHELLINGER, P., SILBERGLEIT, R., SORENSEN, A. G., WARDLAW, J., WU, O. & WARACH, S. 2008. Acute stroke imaging research roadmap. *Stroke*, 39, 1621-8.

WISE, R. A. 2004. Dopamine, learning and motivation. *Nature Reviews Neuroscience*, 5, 483-94.

WOLTERS, F. J., PAUL, N. L. M., LI, L., ROTHWELL, P. M. & OXFORD VASCULAR, S. 2015. Sustained impact of UK FAST-test public education on response to stroke: a population-based time-series study. *International Journal of Stroke*, 10, 1108-14.

WORLD HEALTH ORGANISATION 2001. International Classification of Functioning, Disability and Health. *Geneva*.

WORLD HEALTH ORGANIZATION 2009. Package of Essential Noncommunicable (PEN) Disease Interventions for Primary Health Care in Low Resource Settings.

WU, S., MEAD, G., MACLEOD, M. & CHALDER, T. 2015. Model of understanding fatigue after stroke. *Stroke*, 46, 893-8.

XUE, M. & YONG, V. W. 2008. Matrix metalloproteinases in intracerebral hemorrhage. *Neurological Research*, 30, 775-82.

YOUNG, F. B., LEES, K. R., WEIR, C. J. & INVESTIGATORS, G. I. T. S. C. A. 2005. Improving trial power through use of prognosis-adjusted end points. *Stroke*, 36, 597-601.

YOUNG, W. B., LEE, K. P., PESSIN, M. S., KWAN, E. S., RAND, W. M. & CAPLAN, L. R. 1990. Prognostic significance of ventricular blood in supratentorial hemorrhage: a volumetric study. *Neurology*, 40, 616-9.

ZOROWITZ, R. D., SMOUT, R. J., GASSAWAY, J. A. & HORN, S. D. 2005. Neurostimulant medication usage during stroke rehabilitation: the Post-Stroke Rehabilitation Outcomes Project (PSROP). *Topics in stroke rehabilitation*, 12, 28-36.

List of abbreviations

ACA	Anterior cerebral artery territory
Act FAST	Facial drooping, Arm weakness, Speech slurred - time to call 999 (public health campaign)
AF	Atrial Fibrillation
AISCT	Acute ischaemic stroke classification template
ASPECTS	Alberta Stroke Program Early CT Score
BI	Barthel Index
BOLD	Blood oxygen level dependent signal
BRIC	University of Edinburgh Brain Research Imaging Centre
CD	Compact Disc
CE-CT	Contrast-enhanced computerised tomography scanning
CLRN	Comprehensive local research network
C-RMI	Clinician-completed Rivermead Mobility Index
CT	Computerised tomography scanning
CTRU	The Clinical Trials Research Unit, University of Leeds
DALY	Disability-Adjusted Life Year
DARS	Dopamine Augmented Rehabilitation in Stroke (clinical trial)
DICOM	Digital Imaging and Communication in Medicine
DMEC	Data monitoring and ethics committee
DTI	Diffusion tensor imaging
DWI	Diffusion-weighted imaging
FAS	Fatigue assessment scale

fMRI	Functional Magnetic Resonance Imaging
GCS	Glasgow Coma Scale
GHQ	General Health Questionnaire
GHQ-12	Twelve-item version of the General Health Questionnaire
GOS	Glasgow Outcome Scale
HWS	"Haemorrhage with scan available" (analysis sub-group)
ICF	The World Health Organisation International Classification of Functioning, Disability, and Health
ICH	Intracerebral haemorrhage (here used specifically in reference to primary haemorrhage, not traumatic haemorrhage)
ISRCTN	International standard randomised controlled trial number database
IWS	"Primary infarction, with scan available" (analysis sub-group)
JMP	Dr Jeremy Macmullen-Price, consultant neuroradiologist, Leeds Teaching Hospitals NHS Trust
JP	Dr John Pearn, Clinical Research Fellow to the DARS trial
LACS	Lacunar stroke syndrome
LTHT	Leeds Teaching Hospitals NHS Trust
MCA	Middle cerebral artery
MMSE	Mini mental-state examination
MoCA	Montreal cognitive assessment
MRI	Magnetic resonance imaging
M-RMI	Modified version of the Rivermead Mobility Index
mRS	Modified Rankin scale
MSK-SSP	Musculoskeletal signs, symptoms, and pain manikin
NEADL	Nottingham extended activities of daily living scale

NHS	National Health Service (United Kingdom)
NIHSS	National Institute of Health Stroke Scale,
NOAC	Non-vitamin K oral anticoagulants
OCSP	Oxford Community Stroke Project clinical classification of stroke syndromes
PACS	Partial anterior circulation stroke syndrome
pBI	Postal version of the Barthel Index
PCA	Posterior cerebral artery
PI	Principle Investigator
POCS	Posterior circulation stroke syndrome
PSF	Post-stroke fatigue
RMA	Rivermead Motor Assessment
RMI	Rivermead Mobility Index
RS	Rankin scale (original version)
SBP	Systolic blood pressure
SR-RMI	Patient self-reported Rivermead Mobility Index
STICH	Surgical Trial in Intracerebral Haemorrhage (clinical trial)
TACS	Total anterior circulation stroke syndrome
TP	Dr Tufail Patankar, consultant interventional neuroradiologist, Leeds Teaching Hospitals NHS Trust
TSC	Trial steering committee
UK	United Kingdom
USB	Universal Serial Bus
VAS	Visual Analogue Scale

VBM	Voxel-based morphometry
VCI	Vascular cognitive impairment
VIF	Variance Inflation Factor
VISTA	Virtual International Stroke Trials Archive
VISTA-Rehab	Stroke rehabilitation trials sub-section of the Virtual International Stroke Trials Archive,
WHO	World Health Organisation

Appendix A. Table summarising prognostic models for predicting outcome following ICH

Table A.1. Summary of prognostic models for predicting outcome following ICH.

Author	Type of ICH	Outcome predicted	Predictors included in multivariate model	Results
Hemphill et al. (2001)	Supratentorial and infratentorial haemorrhage	30-day mortality	ICH Score: GCS (score 3-4, 5-12, 13-15); age (<80, ≥80); haematoma volume (<30ml, ≥30ml); intraventricular haemorrhage (yes/no)	30-day mortality was 13% for score of 1; 26% for score of 2; 72% for score of 3; 97% for score of 4 (p<0.0005 for trend). No patients with a score of 0 died. No patient attained the maximum score of 6.
Cho et al. (2008)	Basal ganglia haemorrhage	Six-month mortality, and Glasgow Outcome Scale and Barthel Index at one year	Modified ICH Score (MICH): GCS (score 13-15, 5-12, 3-4); haematoma volume (≤20ml, 21-50ml, ≥50ml); presence of intraventricular haemorrhage or hydrocephalus (yes/no)	For prediction of good functional outcome: score ≥2. At score of 0-1, conservative management achieved a better functional outcome than surgical intervention. Score 3-4: six-month mortality was higher for conservative treatment than surgical management.

Lisk et al. (1994)	Hemispheric haemorrhage	Model 1: good outcome (mRS 0-2) Model 2: poor outcome (mRS 5 and death)	Model 1: age; haemorrhage diameter; presence of intraventricular haemorrhage Model 2: GCS; age; gender; haemorrhage volume Both models generated a predicted probability of achieving the specified outcome.	Model 1: predicted probability of ≥ 0.32 results in a sensitivity of 100% and specificity of 91% for good outcome. Model 2: predicted probability of > 0.60 results in a sensitivity of 62.1% and specificity of 95% for poor outcome.
Ruiz-Sandoval et al. (2007)	Supratentorial and infratentorial haemorrhage	Glasgow Outcome Scale at 30 days. In-hospital and 30-day mortality	ICH Grading Scale (ICH-GS): Age (<45, 45-64, ≥ 65); GCS on admission (13-15, 9-12, 3-8); ICH location (supratentorial or infratentorial); haematoma volume (for supratentorial ICH: <40ml, 40-70ml, >70ml); for infratentorial ICH: <10ml, 10-20ml, >20ml); intraventricular extension (yes/no).	Model explained 44.2% of variance in in-hospital mortality, 43.8% of variance in 30-day mortality, and 33.2% of variance in functional outcome. Sensitivity for in-hospital mortality 78.2%, and 30-day mortality 78.5%.
Tuhrim et al. (1991)	Supratentorial haemorrhage	30-day survival. Barthel Index, Activities of Daily Living Score	Pulse pressure (≤ 40 , 41-65, ≥ 65); GCS (≤ 8 , ≥ 9); volume of haematoma (<27cc, 27-72cc, >72cc); intraventricular extension (yes/no)	Model correctly classified 94% of patients as dead or alive at 30 days, and 95% of patients as having a good or poor outcome.

Cheung and Zou (2003)	Supratentorial and infratentorial haemorrhage	30-day mortality. 30-day Modified Rankin Scale score	Model 1 (30-day mortality): NIHSS score; pulse pressure; subarachnoid extension; intraventricular extension. Model 2 (favourable outcome): low admission temperature; low NIHSS	For 30-day morality: sensitivity of model 91.3%, specificity 72.7%. Positive predictive value 50.0%, negative predictive value 96.6% For favourable outcome: sensitivity 70.0%, specificity 91.7%, positive predictive value 84.8%, negative predictive value 82.1%.
Godoy et al. (2006)	Supratentorial and infratentorial haemorrhage	30-day mortality. 6-month Glasgow Outcome Scale score.	Two modified ICH (mICH) scores were created: Version A: GCS (13-15, 5-12, 3-4); ICH volume (≥ 30 cc, < 30 cc); intraventricular haemorrhage (yes/no); Graeb's score (1-4, 5-8, ≥ 9); infratentorial origin (yes/no); age (< 80 , ≥ 80). Version B: as above, but excluding "infratentorial origin" item.	Version A: cutoff score of 4 best predicted mortality (sensitivity 0.79, specificity 0.80, positive predictive value 0.68, negative predictive value 0.88). Score of 3 best predicted a good outcome (sensitivity 0.85, specificity 0.73, positive predictive value 0.62, negative predictive value 0.9). Version B: cut-off score of 3 was optimum for predicting both mortality (sensitivity 0.85, specificity 0.73, positive predictive value 0.62, negative predictive value 0.90) and good outcome (sensitivity 0.70, specificity 0.90, positive predictive value 0.92, negative predictive value 0.66).

Berwaerts et al. (2000)	Any ICH related to use of oral anticoagulants (included post-traumatic ICH)	In-hospital mortality; functional recovery by time of discharge (none, partial, full).	<p>Model 1: predicts in-hospital mortality. Variables included: type of haemorrhage (presumably refers to supratentorial/ infratentorial location, but the authors do not specify what is meant by this); admission within 12 hours of onset of symptoms; and GCS score <14.</p> <p>Model 2 (radiological predictors of in-hospital mortality). Variables included: diameter of ICH (continuous variable); presence of ischaemic change (yes/ no).</p> <p>Model 3 (in-hospital mortality): presence of ischaemia on CT; intraventricular haemorrhage; displacement of midline; location in posterior fossa; haemorrhage diameter (10mm, ≤30mm, 30-50mm, >50mm).</p>	<p>GCS score alone predicted in-hospital mortality with sensitivity of 46% and specificity of 83%.</p> <p>Model 1: sensitivity of 74%, specificity of 100% for prediction of in-hospital mortality.</p> <p>Model 2: predicted in-hospital mortality with a sensitivity of 73%, specificity of 89%, positive predictive value of 80%, negative predictive value of 85%. The sign for the categorical predictor “presence of ischaemic change” was <i>negative</i>, suggesting that in this sample the presence of ischaemic change protective against mortality.</p> <p>Values of sensitivity/ specificity were not presented for model 3. Odds ratios (with 95% CI) were: ischaemia on scan OR 0.053 (0.003-0.979); intraventricular blood OR 10.21 (0.49-211.81); displacement of midline OR 2.89 (0.04-212.98); location in posterior fossa OR 1.71 (0.06-45.39); haemorrhage diameter OR 93 (OR 0.70-5.33).</p>
-------------------------	---	--	---	--

Fogelholm et al. (1997)	First-ever supratentorial ICH	28-day survival	Level of consciousness (alert, somnolent/disorientated, unconscious/comatose); mean arterial pressure on first day (continuous variable); spread of haemorrhage in to the subarachnoid space (yes/no); shift of midline structures (yes/no); blood glucose level on admission (continuous variable); vomiting on admission (yes/no).	Model predicted 28-day survival with a sensitivity of 78%, specificity of 90%, positive predictive value of 80%, and a negative predictive value of 85%.
Hallevy et al. (2002)	Spontaneous supratentorial ICH	mRS at discharge (cut-off score of ≤ 3 for good outcome and ≥ 4 for poor outcome).	Age (<60 , ≥ 61); severity of hemiparesis (none-moderate [Medical Research Council grade 3-5/5]; severe [grade 0-2/5]); level of consciousness (alert versus drowsy or comatose); presence of mass effect (yes/ no); size of haematoma (small versus medium or large according to radiology opinion); presence of intraventricular extension (yes/ no).	Percentages of patients achieving good outcome: 82% for score of 0-1; 53.7% for a score of 2; 23.3% for a score of 3; 0% for a score of 4-6.
Mase et al. (1995)	Primary ICH	30-day mortality	Intraventricular extension of haemorrhage (yes/no); haematoma "size" (used in this context to refer to volume, not	Model correctly identified 93% of patients who survived and 88% of patients who died.

			maximal dimension: <8.3cc, 8.3-20.4cc, 20.5-47.7cc, >47.7cc); GCS score	
Nilsson et al. (2002)	Primary ICH	30-day and one-year mortality	<p>Model 1 (30-day mortality): level of consciousness on admission (alert, [GCS 14-15] drowsy [GCS 8-13], comatose (GCS 3-7)); haematoma volume (<30cc; 30-60cc; >60cc); history of heart disease prior to ICH (yes/no).</p> <p>Model 2 (one-year mortality): level of consciousness on admission (alert, [GCS 14-15] drowsy [GCS 8-13], comatose (GCS 3-7)); age (0-54; 55-74; ≥75); haematoma location (lobar, central, cerebellar, brainstem).</p>	<p>Patients who were drowsy on admission had a 5.2-fold increase in the odds of death at 30days compared with those who were alert (95% CI 2.3-11.6). Those who were unconscious had a 42-fold increase in the odds of death at 30days (95% CI 15.6-113.3). Haematoma volume of >60cc conferred a 3.2-fold increase in the odds of 30-day mortality compared with haematomas of <30ml (95% CI 1.5-9).</p> <p>At one year, patients who were drowsy on admission had a 3.6-fold increase in the hazard of death compared with those who were alert (95% CI 2.1-6.1). Those who were unconscious had a 15.2-fold increase in the hazard of death at 30days (95% CI 8.8-26.3). Age of ≥75 conferred a 3.3-fold increase in the hazard of death compared with those in the 0-54 age group (95% CI 1.6-6.9).</p>
Passero et al. (2002)	Primary intraventricular haemorrhage	In-hospital mortality	GCS (≤8, >8); early hydrocephalus (yes/no).	Odds of death were 4.67-fold greater for patients with GCS ≤8 compared with those with GCS >8

				(95% CI 1.22-17.92). Odds of death were increased 4.93-fold for those with hydrocephalus versus those without (95% CI 1.13-21.59)
Phan et al. (2000)	Deep ICH. Divided in to: caudate/ thalamic ICH ("medial" group) and putaminal ICH ("lateral" group)	30-day mortality mRS (good outcome <2; moderate to poor outcome 2-4; dependency 5; death 6)	Model 1 (30-day mortality for both medial groups): GCS (≤ 8 , > 8); presence of hydrocephalus on visual inspection (yes/no). Model 2 (30-day mortality in medial group): only GCS (≤ 8 , > 8) was a significant predictor variable. Model 3 (30-day mortality in lateral group): GCS (≤ 8 , > 8); presence of hydrocephalus on visual inspection (yes/no).	Model 1: sensitivity of 55% and specificity of 90%. Positive predictive value 70%, negative predictive value 83%. Model 2: sensitivity and specificity not quoted. Patients with GCS ≤ 8 had 16.5-fold increase in the odds of 30-day mortality compared with patients with GCS > 8 (95% CI 3.7-73.4). Model 3: sensitivity of 57% and specificity of 91%. Positive predictive value 73%, negative predictive value 84%.
Portenoy et al. (1987)	Spontaneous supratentorial ICH	Good outcome (no deficit to moderately dependent) versus poor outcome (severely dependent, persistent vegetative state, or death).	Model for prediction of good (versus poor) outcome: GCS; size of haematoma (a calculated index value based on maximal haematoma dimensions. Categorised as: 4-19; 20-28; 29-44; 45-81); intraventricular extension (yes/no).	Model correctly identified 87% of patients with good outcomes and 88% of patients with poor outcomes.

Qureshi et al. (1995)	Spontaneous ICH	Deterioration from an initial GCS of >12 by ≥ 4 points within the first 24 hours of admission; in-hospital mortality.	<p>Model 1 (prediction of early deterioration): presence of intraventricular extension (yes/no); ICH volume (<30cc, ≥ 30cc).</p> <p>Model 2 (prediction of mortality): presence of intraventricular extension (yes/no); ICH volume (<30cc, ≥ 30cc); initial GCS (≤ 12, >12).</p>	<p>Model 1: odds ratio for early deterioration with (versus without) ventricular extension was 4.67 (95% CI 1.30-16.72). Odds ratio early deterioration with ICH volume ≥ 30cc (versus <30cc) 6.78 (95% CI 1.89-24.35).</p> <p>Model 2: odds ratio for death with (versus without) intraventricular extension 4.23 (95% CI 1.82-9.82); odds ratio for mortality with ICH volume ≥ 30cc (versus <30cc) 6.66 (95% CI 2.85-15.58); odds ratio for mortality with initial GCS ≤ 12 (versus >12) 3.23 (95% CI 1.46-7.14)</p>
Razzaq and Hussain (1998)	Spontaneous ICH	30-day mortality	<p>Model 1 (clinical predictors): GCS score (≥ 12, 9-11, ≤ 8); paresis (yes/no); 7th nerve palsy (yes/no).</p> <p>Model 2 (CT predictors): ventricular enlargement (yes/no); haematoma size (<3cm, >3cm); intraventricular extension of haemorrhage (yes/no).</p> <p>Model 3 (combined clinical and CT predictors): GCS score (≥ 12, 9-11, ≤ 8); history of hypertension</p>	<p>Model 1: patients with GCS ≤ 8 had 7.6-fold increase in the odds of dying compared with patients with GCS ≥ 12.</p> <p>Model 2: patients with intraventricular extension of the haemorrhage had a 3.5-fold increase in the odds of death compared with those without. Those with ventricular enlargement had a 2.6-fold increase in the odds of death compared with those without.</p> <p>Model 3: patients with GCS ≤ 8 had 11-fold increase in</p>

			(yes/no); pulse pressure (continuous); haematoma size (<3cm, >3cm); intraventricular extension of haemorrhage (yes/no).	the odds of death compared with patients with GCS ≥ 12 . Patients with hypertension had a 3-fold increase in the odds of death compared with those without. Each unit increase in pulse pressure conferred a 1.3-fold increase in the odds of death. Those with a haematoma size of >3cm had a 2.9-fold increase in the odds of death compared with those with a haematoma of <3cm. Patients with intraventricular extension of the haemorrhage had a 2.9-fold increase in the odds of death compared with those without.
Schwarz et al. (2000)	Spontaneous supratentorial ICH	Glasgow Outcome Scale at 72 hours. Poor outcome 1-2; moderate or favourable outcome 3-5	Model for prediction of poor outcome: ventricular haemorrhage (yes/no); haematoma enlargement or secondary haemorrhage (yes/no); initial GCS (14-15; 8-13; 3-7); haematoma volume (<25ml; 25-60ml; >60ml); duration of elevation in blood pressure (<24hrs; 24-48hrs; >48hrs); duration of elevated blood glucose levels >11mmol/l (never; <24hrs; >24hrs); duration of	Presence of ventricular haemorrhage conferred a five-fold increase in the odds of a poor outcome. Haematoma expansion increased the odds of a poor outcome seven-fold. An initial GCS of 3-7 increased the odds of an unfavourable outcome by a factor of 18.34 compared with an initial GCS of 14-15. A haematoma volume of 25-60ml increased the odds of an unfavourable outcome by a factor of 1.75 compared with a volume of <25ml; haematoma volumes of >60ml conferred a 13.9-

			temperature >37.5C (<24hrs; 24-48hrs; >48hrs).	fold increase in the odds of an unfavourable outcome. Elevation in blood pressure for >48hrs conferred an 8.42-fold increase in the risk of an unfavourable outcome compared to a rise in blood pressure of <24hrs duration. Elevated blood glucose levels for >48hrs increased the odds of an unfavourable outcome by a factor of 13.53 compared with those who had never had high blood glucose levels. A body temperature of >37.5C for >48hrs conferred a 13.52-fold increase in the odds of an unfavourable outcome when compared with those who had an elevated body temperature for <24hrs.
Shaya et al. (2005)	Non-traumatic first-time ICH.	Glasgow Outcome Scale at 6 months	Haematoma volume (<20ml 1 point; 20-50ml 2 points; >50ml 3 points); hydrocephalus (1 point if present); focal neurological deficit (1 point if present)	Score derived from model was used to predict six-month GOS (no detail provided on how this was accomplished). Spearman correlation coefficient between observed and predicted values was 0.76 (p=0.0001).

Appendix B. The AISCT template for coding ischaemic stroke lesions

The Acute Ischaemic Stroke Classification Template (AISCT) was developed by Professor Wardlaw et al, of the University of Edinburgh Brain Research Imaging Centre (BRIC), Neuroimaging Sciences, Edinburgh (www.bric.ed.ac.uk) (Wardlaw and Sellar, 1994). The BRIC is part of the SINAPSE (Scottish Imaging Network–A Platform for Scientific Excellence) collaboration (www.sinapse.ac.uk) funded by the Scottish Funding Council and the Chief Scientist Office. The AISCT is available to download from:

<http://www.bric.ed.ac.uk/research/imageanalysis.html#ais>

The AISCT was used in DARS with the permission of Professor Wardlaw (private communication with author).

B.1. AISCT template for coding acute ischaemic change in the MCA territory

The AISCT classifies acute ischaemic change within the MCA territory using the eight template diagrams shown below (© Professor J.M. Wardlaw, University of Edinburgh, Brain Research Imaging Centre; reproduced here with permission).

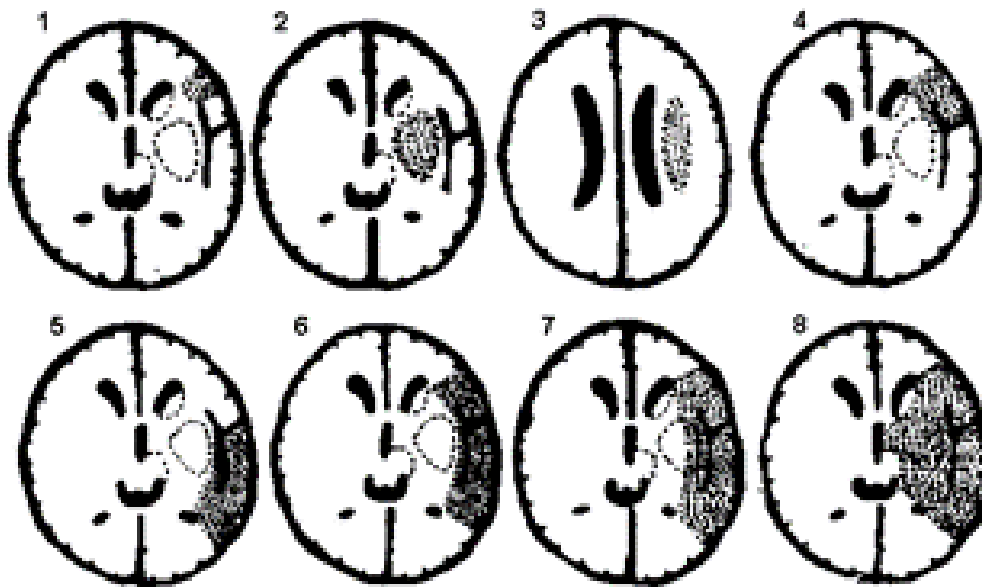


Table B.1. Key for template images coding ischaemic change in the MCA territory

Template	Finding represented
1	Small cortical infarct
2	Basal ganglia infarct (>2×2×2cm)
3	Infarct of the white matter lateral to the ventricle (>2×2×2cm)
4	Infarct of anterior half of the peripheral MCA territory
5	Infarct of the posterior half of the peripheral MCA territory
6	Infarct of whole of peripheral MCA territory
7	Infarct of posterior half of MCA territory plus lateral part of basal ganglia
8	Infarct of whole MCA territory

B.2. AISCT template for lacunar lesions and border zone ischaemia

The AISCT classifies lacunar lesions and border-zone ischaemia using the template diagrams shown below (© Professor J.M. Wardlaw, University of Edinburgh, Brain Research Imaging Centre; reproduced here with permission).

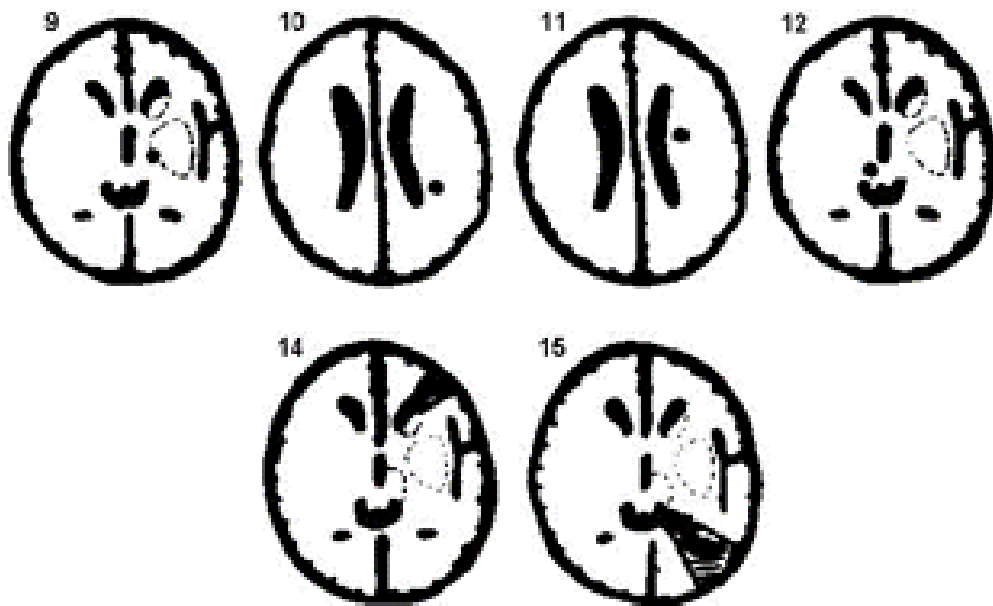


Table B.2. Key for template images coding lacunar lesions and border-zone ischaemic change

Template	Finding represented
9	Lacune in internal capsule/ lentiform nucleus
10	Lacune in internal border zone
11	Lacune in centrum semiovale
12	Lacune in thalamus
13	Lacune in brainstem, including pons (not shown)
14	Anterior border zone infarction
15	Posterior border zone infarction

B.3. AISCT template for tissue swelling

The AISCT classifies the extent of tissue swelling in acute ischaemic stroke using the template diagrams shown below (© Professor J.M. Wardlaw, University of Edinburgh, Brain Research Imaging Centre; reproduced here with permission).

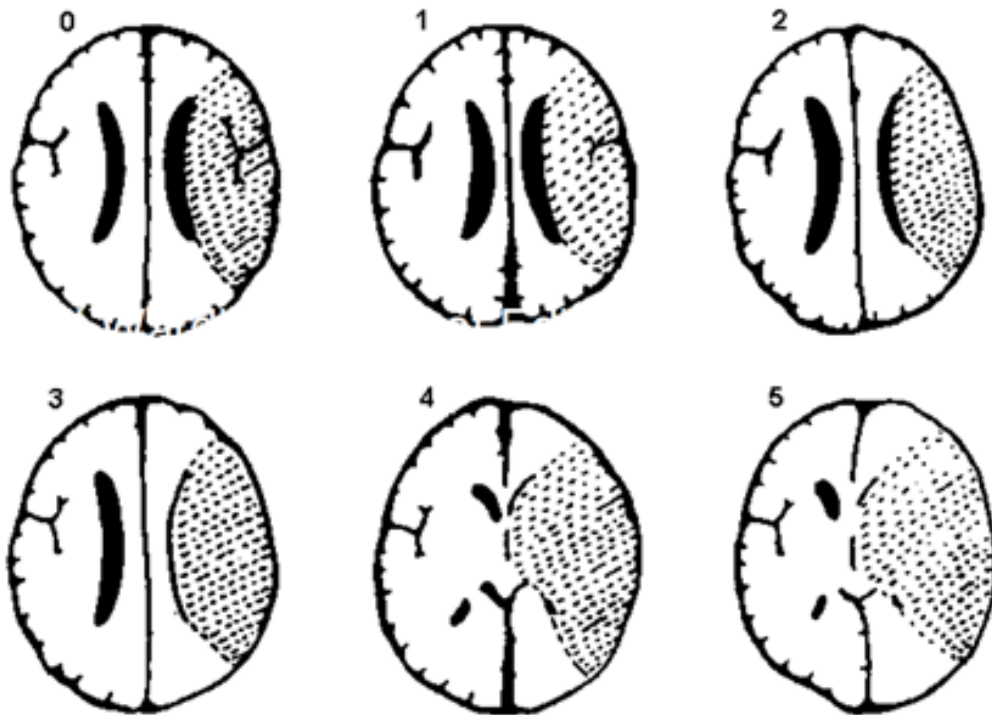


Table B.3. Key for template images coding extent of tissue swelling in acute ischaemic stroke

Template	Finding represented
0	No swelling
1	Effacement of sulci overlying the infarct
2	1+minor effacement of adjacent lateral ventricle
3	1+complete effacement of lateral ventricle
4	1+effacement of lateral and third ventricle
5	4+shift of midline away from the side of the ventricle
6	5+effacement of basal cisterns (not shown)

Appendix C. Summary of standardised outcome measures used in the DARS trial

A range of measures were used in the DARS trial, to capture not only the primary outcome (mobility) but also other functional outcomes (such as upper limb function and independence in activities of daily living), general disability, and other variables that could modify a patient's response to rehabilitation (such as musculoskeletal pain, fatigue, and depression). These outcome measures are summarised below in Table C.1.

Abbreviations used in this table:

- PO: Primary Outcome
- EM: Effect modifier
- SO: Secondary outcome
- A/P: Activities/participation
- I: Impairment
- SR: Patient self-report
- CR: Clinician report
- RI: Researcher interview
- Y: Yes
- N: No

Table C.1. Standardised outcome measures used in the DARS trial.

Outcome measure	Analysis	Construct	ICF domain	Method	T ₀	T ₁	T ₂	T ₃
Rivermead Mobility Index (RMI) (Collen et al., 1991)	PO	Mobility	A/P	SR, CR	Y	Y	Y	Y
General Health Questionnaire, 12-item version (GHQ-12) (Goldberg and Hillier, 1979)	EM	Psychological morbidity	I	SR	Y	Y	Y	Y
Musculoskeletal signs/ symptoms and pain manikin (MSK-SSP) (Keenan et al., 2006)	EM	Musculoskeletal pain	I	SR	Y	Y	Y	Y
Fatigue Assessment Scale (FAS) (Michielsen et al., 2003)	EM	Fatigue	I	SR	N	Y	Y	Y
Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005)	EM	Cognitive function	I	RI	Y	Y	Y	Y
Postal version of the Barthel Index (Gompertz et al., 1994)	SO	Independence in activities of daily living	A/P	SR	Y	Y	Y	Y
Nottingham Extended Activities of Daily Living Scale (NEADL) (Nouri and Lincoln, 1987)	SO	Independence in activities of daily living	A/P	SR	Y	Y	Y	Y
ABILHAND scale (Penta et al., 1998)	SO	Hand function	A/P	SR	Y	Y	Y	Y
EuroQol EQ-5D (Rabin and de Charro, 2001)	SO	Health-related quality of life	A/P	SR	Y	Y	Y	Y
Modified Rankin Scale (mRS) (Bonita and Beaglehole, 1988)	SO	Global disability	A/P	RI	N	Y	Y	N

Appendix D. The Rivermead Mobility Index

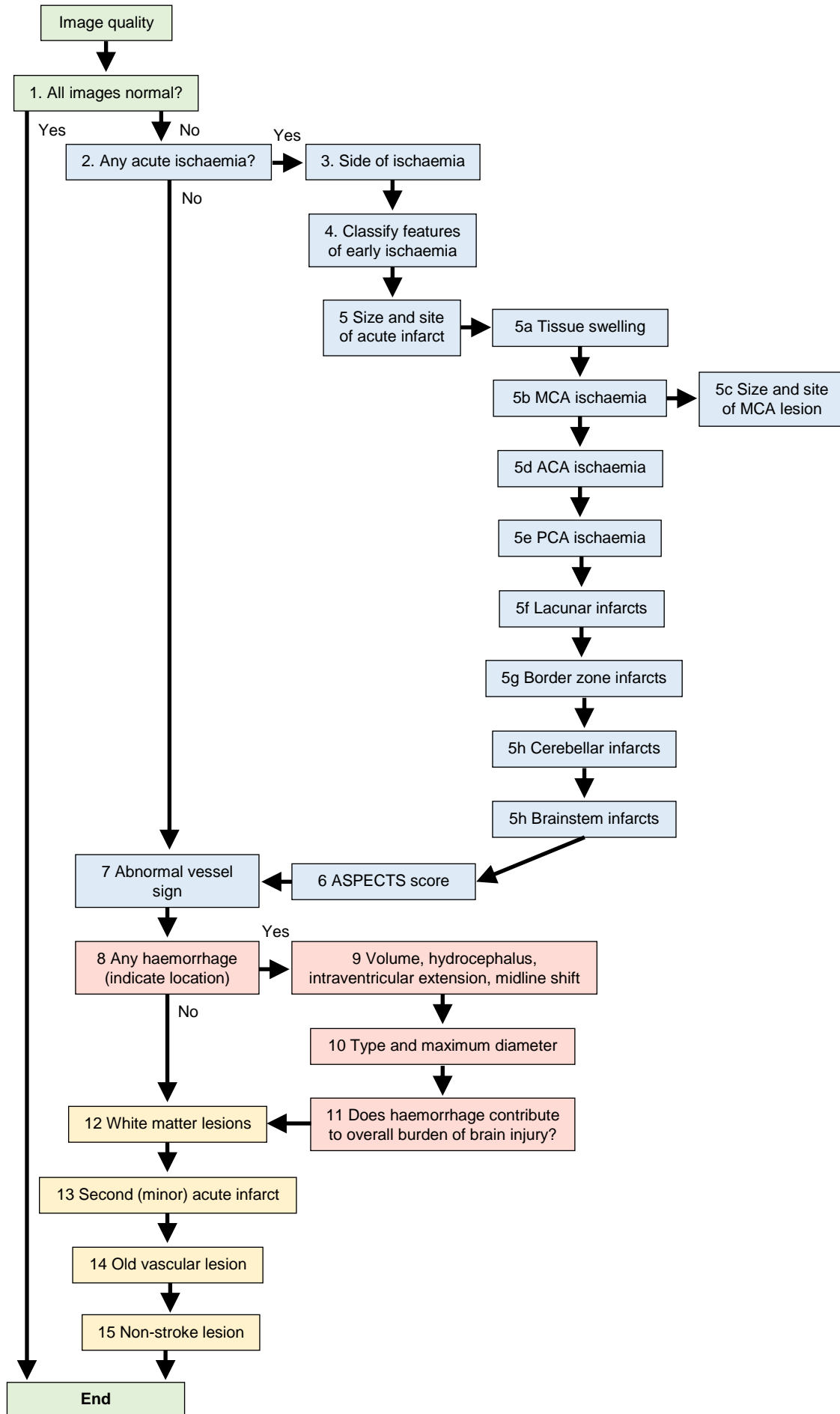
The RMI (Collen et al., 1991), reproduced in full below, was both an inclusion/exclusion criterion and the key outcome measure for DARS. The RMI was used here as both a professionally-completed (C-RMI) and as a patient self-report (SR-RMI) measure. Both were identical in every respect, other than the manner in which they were completed.

At the time of enrolment, all patients had a C-RMI score of ≤ 6 , and had answered “no” to question 7 (highlighted below in red). When used as an outcome measure, the SR-RMI was dichotomised as “able to walk 10m or more with an aid if necessary but no standby help” (“yes” or “no”).

1. Do you turn over from your back to your side without help?
2. From lying in bed, are you able to get up to sit on the edge of the bed on your own?
3. Could you sit on the edge of the bed without holding on for 10 seconds?
4. Can you (using hands and an aid if necessary) stand up from a chair in less than 15 seconds, and stand there for 15 seconds
5. Observe patient standing for 10 seconds without any aid.
6. Are you able to move from bed to chair and back without any help?
- 7. Can you walk 10 metres with an aid if necessary but with no standby help?**
8. Can you manage a flight of steps alone, without help?
9. Do you walk around outside alone, on pavements?
10. Can you walk 10 metres inside with no calliper, splint or aid and no standby help?
11. If you drop something on the floor, can you manage to walk 5 metres to pick it up and walk back?
12. Can you walk over uneven ground (grass, gravel, dirt, snow or ice) without help?
13. Can you get in and out of a shower or bath unsupervised, and wash yourself?
14. Are you able to climb up and down four steps with no rail but using an aid if necessary?
15. Could you run 10 metres in 4 seconds without limping? (A fast walk is acceptable.)


Appendix E. Flow diagram summarising the process of image analysis

Figure E.1. Summary of the process of imaging analysis in the DARS trial.



Appendix F. CT Image Interpretation Case Report Form

The CT Image Interpretation Case Report Form. One form was completed for every scan analysed. This protocol was developed by Wardlaw et al for use in the Third International Stroke Trial (Wardlaw and Sellar, 1994).



CT Image Interpretation
Page 1 of 6

Participant Initials

Date of Birth

Participant ID

To be completed during/following review of scans

Date of reading Date of scan

Scan quality Good Moderate Poor

Clinical information

Reader name

Type of scan

CT plain	<input type="checkbox"/> Yes	<input type="checkbox"/> No
CTA	<input type="checkbox"/> Yes	<input type="checkbox"/> No
CECT	<input type="checkbox"/> Yes	<input type="checkbox"/> No

Please tick Yes or No. Please do not leave blanks. Thank you.

	Yes	No	
1. Are all the scan sequences completely normal?	<input type="checkbox"/>	<input type="checkbox"/>	If Yes stop here and sign the bottom of this form

ISCHAEMIC CHANGES

	Yes	No	
2. Is there any sign of acute ischaemic change on any sequence? If in doubt as to whether acute or old, code as acute.	<input type="checkbox"/>	<input type="checkbox"/>	If No go to Q7

	Right	Left	
3. Which side of the brain shows ischaemic change?	<input type="checkbox"/>	<input type="checkbox"/>	Tick Right and Left if both

4. Classify signs of ischaemic change in the main lesions (if more than one recent lesion). (see examples)

	Yes	No	
a) Loss of grey/white matter cortex definition	<input type="checkbox"/>	<input type="checkbox"/>	N/A
b) Loss of basal ganglia outline	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Hypodensity present (i.e. more than in 'a' or 'b' so that the lesion appears less dense than white matter)	<input type="checkbox"/>	<input type="checkbox"/>	

Completed by Date Form continues on next page ►►

Please complete and return top copy to CTRU, University of Leeds (please see Investigator Site File for full contact details)

For office use only	Computerised <small>Date Initials</small>	Verified/Checked <small>Date Initials</small>
---------------------	--	--

Version 1.0 17/10/2013



CT Image Interpretation
Page 2 of 6

Participant Initials			
Date of Birth	Day	Month	Year
Participant ID	Trail Site	Trail No	

5. CLASSIFY SITE AND SIZE OF ACUTE ISCHAEMIC LESION ON PLAIN CT (see examples)

- a) Describe the degree of tissue swelling
Tick one only
- None
 - Effacement of the sulci overlying the infarct
 - 2+ Minor effacement of the adjacent lateral ventricle
 - 2+ Complete effacement of the lateral ventricle
 - 2+ Effacement of the lateral and third ventricle
 - 5+ Shift of the midline away from the side of the ventricle
 - 6+ Effacement of the basal cisterns
- b) Are there any early ischaemic changes in the middle cerebral artery (MCA) territory?
- MCA territory is not affected → Go to Q5d
 - Less than 33% of MCA territory } → Continue to Q5c
 - More than 33% of MCA territory }
- c) Classify site and size of ischaemic lesion of MCA territory
- Small cortical
 - Basal ganglia striatocapsular
 - Lateral to ventricle striatocapsular
 - Anterior cortical MCA territory
 - Posterior cortical MCA territory
 - Whole of cortical MCA territory
 - Whole of cortical MCA territory with lateral part of basal ganglia
 - Whole MCA territory
- d) Is the anterior cerebral artery (ACA) territory affected?
- ACA territory not affected
 - Less than 50% of ACA territory
 - More than 50% of ACA territory
 - Complete ACA territory
- e) Is the posterior cerebral artery (PCA) territory affected?
- PCA territory not affected
 - Less than 50% of PCA territory
 - More than 50% of PCA territory
 - Complete PCA territory

Completed by Date

Day	Month	Year
-----	-------	------

Form continues on next page ▶▶

Please complete and return top copy to CTRU, University of Leeds (please see Investigator Site File for full contact details)

For office use only	Computerised	Verified/Checked
	Date	Initials
	Initials	Date
		Initials



CT Image Interpretation
Page 3 of 6

Participant Initials			
Date of Birth	Day	Month	Year
Participant ID	Trial Site		Trial No

- f) Please specify any small subcortical infarctions (i.e. lacunar) that are present
- None
 - Internal capsule / lentiform
 - Internal border zone
 - Centrum semiovale
 - Thalamus
- g) Are there any cortical border zone infarctions?
- None
 - Anterior border zone
 - Posterior border zone
- h) Are there any cerebellar infarctions?
- None
 - Lacunar infarction
 - Less than 50% of hemisphere
 - More than 50% of hemisphere
- i) Are there any brainstem infarctions?
- None
 - Lacunar infarction
 - Less than 50% of brainstem
 - More than 50% of brainstem

6. ASPECT SCORE LESION

Please enter '1' for all abnormal areas, '0' for normal areas, 'U' for unscorable areas.
(*Unscorable' = areas not included on CTP)

	Caudate (C)	Lentiform (L)	Insula (I)	Internal Capsule (IC)	MCA1 (M1)	MCA2 (M2)	MCA3 (M3)	MCA4 (M4)	MCA5 (M5)	MCA6 (M6)
Swelling and/or hypoattenuation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Completed by Date

Day	Month	Year

Form continues on next page ►►

Please complete and return top copy to CTRU, University of Leeds (please see Investigator Site File for full contact details)

For office use only	Computerised		Verified/Checked	
	Date	Initials	Date	Initials



CT Image Interpretation

Page 4 of 6

Participant Initials			
Date of Birth	Day	Month	Year
Participant ID	Trial Site		Trial No

7. CT HYPERATTENUATED/ABNORMAL VESSEL SIGN

Is there a hyperdense artery sign (i.e. potentially due to thrombus, rather than calcification)?

Tick all that apply

- None
- Middle cerebral artery (MCA) main stem
- Insular MCA
- Internal carotid artery (ICA)
- Anterior cerebral artery (ACA)
- Posterior cerebral artery (PCA)
- Basilar artery (BA)
- Vertebral artery (VA)

Completed by Date

Day	Month	Year

Form continues on next page ►►

Please complete and return top copy to CTRU, University of Leeds (please see Investigator Site File for full contact details)

For office use only	<i>Computerised</i>		<i>Verified/Checked</i>	
Date	Date	Initials	Date	Initials



CT Image Interpretation

Page 5 of 6

Participant Initials	_ _
Date of Birth	_ _ _ _ _ _
Participant ID	_ _ _ _ _ _ _ _
	_ _ _ _ _ _ _ _

8. HAEMORRHAGE

Yes No

Is there any haemorrhage anywhere? Yes No **If No go to Q12**

If Yes, indicate location:

<input type="checkbox"/> Frontal	<input type="checkbox"/> Occipital	<input type="checkbox"/> Brain stem
<input type="checkbox"/> Temporal	<input type="checkbox"/> Basal ganglia	<input type="checkbox"/> Cerebellum
<input type="checkbox"/> Parietal	<input type="checkbox"/> Internal capsule	

9. If Yes to question 8, please score intracerebral haemorrhage:

Number of CT slices where haemorrhage is visible	_	Height (A)	_	.	_	mm	Midline shift	_	.	_	mm
CT slice thickness	_	Length (B)	_	.	_	mm					
	mm	Breadth (C)	_	.	_	mm					
Volume ((A × B × C)/2)		_ _ _ _	.	_	mm ³						

Yes No

Presence of hydrocephalus? Yes No

Intraventricular extension? Yes No

10. Classify haemorrhage

(If more than one haemorrhage, tick all present – indicate order of significance):

Order
(Insert 1 (most important), 2, 3 (least important) to indicate your estimate of the order of clinical importance)

Size of Haematoma
(Tick box for maximum diameter)

	Yes No		Order	Size of Haematoma			
	Yes	No		<3 cm	3–5 cm	5–8 cm	>8 cm
a) Petechial haemorrhage	<input type="checkbox"/>	<input type="checkbox"/>	_	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Significant haemorrhagic transformation of infarct <i>(i.e. underlying infarct still visible)</i>	<input type="checkbox"/>	<input type="checkbox"/>	_	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Parenchymal haematoma <i>(i.e. no infarct visible)</i>	<input type="checkbox"/>	<input type="checkbox"/>	_	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Parenchymal haematoma clearly remote from infarct	<input type="checkbox"/>	<input type="checkbox"/>	_	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Subdural haematoma	<input type="checkbox"/>	<input type="checkbox"/>	_	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Subarachnoid haemorrhage	<input type="checkbox"/>	<input type="checkbox"/>	_	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Extradural haemorrhage	<input type="checkbox"/>	<input type="checkbox"/>	_	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Yes No

11. In your opinion, is the haemorrhage a major component of the infarct which is likely to have worsened mass effect or involved more brain in the damage present and so worsened symptoms, or if remote from the infarct, likely to have contributed significantly to the burden of brain damage? Yes No

Completed by Date

Day	_	_	Year	_	_
-----	---	---	------	---	---

Form continues on next page ►►

Please complete and return top copy to CTRU, University of Leeds (please see Investigator Site File for full contact details)

For office use only	<i>Computerised</i>	<i>Verified/Checked</i>
Date	_ _ _ _	Date
	Initials	Initials

“Job done.”

- Bipin Bhakta