Early Mobilisation After Stroke: Barriers and Enablers for Effective Use of Population-Level Evidence to Inform Individualised Clinical Decision-Making

Venesha Rethnam

ORCID: 0000-0002-0493-8536

September 2020

Submitted in total fulfilment of the requirements of the degree of

Doctor of Philosophy

Florey Institute of Neuroscience and Mental Health

Faculty of Medicine, Dentistry and Health Sciences

University of Melbourne

Abstract

Early mobilisation, defined as sitting out of bed, standing or walking early after stroke, is an important constituent of acute stroke unit care. However, the *multifaceted* and *complex* nature of early mobilisation evidence and interventions, clinical practice guideline development and reporting, and the characteristics of stroke patients and their outcomes challenges individualised decision-making. The overall *objective* of this thesis was to investigate the barriers and enablers in the effective use of population-level evidence to inform individual patient-level clinical decision-making for early mobilisation post-stroke.

The objective was achieved by conducting five studies: (1) a review of early mobilisation clinical practice guidelines as decision-support tools for individual patient-level decision-making; (2) a meta-analysis of individual participant data from EM trials, (3) an investigation of factors guiding early mobilisation decision-making by expert stroke clinicians, and investigation of how well (4) the utility-weighted modified Rankin scale, and (5) modified Rankin scale reflect post-stroke burden for an individual patient.

(1) Based on the findings from the review of clinical practice guidelines, the decision-support requirements were met to a varying degree by early mobilisation clinical practice guidelines. Four key recommendations were formed for the future development of clinical practice guidelines. These included more granular descriptions of patient and stroke characteristics to allow tailoring of decisions to individual patients; clarity about when clinical flexibility is appropriate; a detailed description of the intervention dose, and physical assessment criteria including safety

parameters. (2) The individual participant meta-analysis allowed the inclusion of individual patient-specific information and further strengthened earlier evidence from conventional meta-analyses that commencement of early mobilisation should only be considered after 24 h post-stroke. It also reinforced the importance of adequate reporting of early mobilisation interventions to ensure the applicability of evidence to individual patients. (3) Interviews with expert stroke clinicians revealed that more than 80% of stroke experts considered stroke type and severity, medical stability to be the most important factors contributing to decision-making about early mobilisation. Inadequate staffing, equipment and low level of staff expertise were barriers for early mobilisation. (4) Based on the investigations of the utility-weighted modified Rankin scale, high variability in individual patient-centred utility values between and within mRS categories, over time post-stroke, and using different derivation methods was found. This variability is not adequately reflected in the utility-weighted modified Rankin scale. (5) Quality of life and activities of daily living domains demonstrated patient-specific patterns of post-stroke burden across the mRS and UW-mRS.

From this research, gaps in the early mobilisation evidence base were identified that require future exploration of existing data and the development of new clinical trials to better support evidence-based clinical decision-making. Detailed areas of improvement for clinical practice guidelines as decision support tools were also identified to effectively translate population-level evidence to inform complex clinical decision-making at an individual patient-level. Finally, the identified multifaceted patterns of post-stroke burden across the utility-weighted modified Rankin scale and modified Rankin scale may facilitate appropriate assessment, articulation and interpretation of the outcomes for individual patient decision-making.

The collective work has substantially contributed to systematically identifying the barriers and enablers in using population-level evidence to inform patient-level clinical decision-making in this challenging clinical context.

Declaration

This is to certify that:

(i) the thesis comprises only my original work towards the Doctorate of Philosophy except where indicated in the Preface;

(ii) due acknowledgement has been made in the text to all other material used;

(iii) the thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Venesha Rethnam

Preface

This thesis includes original research undertaken in the Florey Department of Neuroscience and Mental Health, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne. The work presented does not include any research performed prior to enrolment of the PhD candidature toward the conferment of other qualifications.

All the work presented in this thesis has been published or is currently in submission for publication. I have contributed to more than half of the work towards publication, which has been verified by all co-authors. I acknowledge the contributions of all co-authors in preparing this work for submission.

1. **Rethnam V.**, Hayward K.S., Bernhardt J., & Churilov L. (2020). The Utility and Limitations of Clinical Practice Guidelines to Support Early Mobilisation Decision-Making After Stroke. Published in *Frontiers in Neurology* on 7 January 2021.

The text from Chapter 3 is taken from the accepted manuscript to *Frontiers in Neurology*. My contribution to this paper was the design of the study; obtaining human research ethics committee approval, appraisal and analysis of the data; conduct of the interviews; interpretation the findings; and drafting the manuscript. I acknowledge the valuable contributions of Leonid Churilov for the conception of the work; and Kate Hayward for independently appraising the data.

2. **Rethnam, V.**, Langhorne, P., Churilov, L., Hayward, K.S., Herisson, F., Poletto, S.R., Tong, Y., & Bernhardt, J. (2020). Early mobilisation post-stroke: a systematic review and meta-analysis of individual participant data. *Disability and Rehabilitation.* Copyright © 2020 Informa UK Limited. DOI: 10.1080/09638288.2020.1789229

The text from Chapter 4 is taken from the published article in *Disability and Rehabilitation* on 16 July 2020. My contribution to this paper was in obtaining local ethics approval; drafting and obtaining data sharing agreements; acquiring, appraising and analysing the data; interpretating the findings; and writing the manuscript. I acknowledge the valuable contributions of Julie Bernhardt and Peter Langhorne for the conception of the study and the development of the protocol; Leonid Churilov for providing statistical support; and Kate Hayward for independently appraising the data. I additionally acknowledge the generous contributions from the co-authors that obtained local approval to access their clinical trial data for this study. Permission has been granted by *Taylor & Francis Group* to include the Author Accepted Manuscript in Chapter 4. The Version of Record is under embargo for 12 months.

3. **Rethnam V.**, Hayward K.S., Churilov L., & Bernhardt J. (2020). Clinical and Systems of Care Factors Contributing to Individual Patient Decision-Making for Early Mobilisation Post-Stroke. Submitted to *Clinical Rehabilitation* on 18 January 2021.

The text from Chapter 5 is taken from a manuscript submitted to *Clinical Rehabilitation*. My contribution to this paper was the design of the study; obtaining human research ethics committee approval; acquisition, appraisal and analysis of the data; conduct of the interviews; interpretation the findings; and drafting the manuscript. I acknowledge the valuable contributions of Julie Bernhardt and Leonid Churilov for the conception of the work; Kate Hayward for independently appraising the data; and Hannah Johns for developing the data visualisation tool.

4. **Rethnam, V.**, Bernhardt J., Dewey H., Moodie M., Johns H., Gao L., Collier J., Ellery F., & Churilov L. (2020). Utility-weighted modified Rankin Scale: Still too crude to be a truly patient-centric primary outcome measure? *International Journal of Stroke*. 15(3), 268–277. Copyright © 2020 SAGE Publications. DOI: 10.1177/1747493019830583

The text from Chapter 6 is taken from the above article published by the *International Journal of Stroke* on 12 February 2020. My contribution to this paper was in designing the study; analysing the data; interpretating the findings; and writing the manuscript. I acknowledge the valuable contribution of Leonid Churilov for the conception of the work and providing statistical support. I additionally acknowledge the AVERT Trial Collaboration group for permitting access to the clinical trial data to conduct this work. Permission the has been granted by *SAGE Publications* to include the Author Accepted Manuscript in Chapter 6, and the final Version of Record in Appendix A.

5. **Rethnam V.**, Bernhardt J., Johns H., Hayward K.S., Collier J.M., Ellery F., Gao L., Moodie M., Dewey H., Donnan G.A., & Churilov L. (2020). Look Closer: The Multidimensional Patterns of Post-Stroke Burden behind the modified Rankin Scale. *International Journal of Stroke.* Copyright © 2020 SAGE Publications. DOI: 10.1177/1747493020951941

The text from Chapter 7 is taken from the above article published by the *International Journal of Stroke* on 27 August 2020. My contribution to this paper was in designing the study; analysing the data; interpretating the findings; and writing the manuscript. I acknowledge the valuable contribution of Leonid Churilov for the conception of the work and providing statistical support. I additionally acknowledge the AVERT Trial Collaboration group for permitting access to clinical trial data to conduct this work. Permission the has been granted by *SAGE Publications* to include the Author Accepted Manuscript in Chapter 7, and the final Version of Record in Appendix A.

Funding

2020	Miller Postgraduate Travel Award
2018	Olivia Newton-John Cancer Research Institute Travel Award
2018 – 2020	The Florey Institute of Neuroscience and Mental Health Annual Travel Award
2018 - 2020	The Florey Institute of Neuroscience and Mental Health Top-Up Scholarship
2017 – 2020	NHMRC Centre of Research Excellence in Stroke and Brain Recovery PhD Scholarship (# 1058635)

The Florey Institute of Neuroscience and Mental Health is supported by the Victorian government via the Operational Infrastructure Support Scheme.

Acknowledgements

The work in this thesis reflects the generosity, devotion and brilliance of the stroke research community, to whom I owe my sincere gratitude.

Firstly, I thank my supervisors, Professor Leonid Churilov and Professor Julie Bernhardt, for their guidance and support throughout my candidature. I acknowledge Leonid's unwavering dedication as my primary supervisor; his inconceivable brilliance; his ability to problem solve *everything*; and his gift to conceptualise and operationalise *complex* research. Most importantly, I thank Leonid for successfully ensuring my PhD was a gratifying and enjoyable experience. I am thankful for his warm encouragement in my growth, both professionally and personally. I can only wish that after four years, a speck of his brilliance has rubbed off on me. I remember our first meeting when I was an Honours student learning about ICC tests while simultaneously pretending not to panic by the sight of STATA. Within 30 minutes, Leonid took the weight of the world off my shoulders and instilled a sense of confidence that I never knew I had. He continues to do this every day, and I am profoundly grateful. I also thank my co-supervisor, Julie, as the work in this thesis is also a testament to her research acumen. She is an incredible researcher, a pioneer in this field and fearless leader. I thank Julie for teaching and showing me how to conduct rigorous and impactful research.

I thank the support of my Advisory Committee members, Professor Vincent Thijs and A/Professor Nawaf Yassi for challenging my thinking and research while also keeping me on track throughout my candidature. I am appreciative for their contributions to this thesis.

I acknowledge the enormous support and invaluable scientific inputs from Dr Kate Hayward to this thesis. I am lucky and thankful to have Kate's encouragement of my progress and future endeavours. Above all, I thank Kate for being a generous, kind and wonderful friend.

I wish to thank several people whose contribution has been vital for my research: Ms Fiona Ellery, Professor Geoffrey Donnan, Professor Helen Dewey, Professor Marj Moodie, Dr Janice Collier and Dr Lan Gao. I also wish to thank the international trialists and their respective teams for the generous contributions to my research: A/Professor Simone Poletto, Dr Fanny Herisson, Dr Yanna Tong, and especially Professor Peter Langhorne. I acknowledge the AVERT Trial Collaboration group for granting access to their dataset, and Dr Hannah Johns for allowing the use of her data visualisation tool (AVERT Atlas) in my research.

My PhD would not have been possible without the contributions of the stroke experts in my studies that shared their time, expertise and experiences. I also wish to express my humble gratitude to the participants, their carers and families who graciously dedicated their time to this research.

I thank my fabulous colleagues from the Florey Institute of Neuroscience and Mental Health. Thank you to the fantastic AVERT lab for their advice, support and laughs: Maree Honey, Sharon Kramer, Lilian Braighi Carvalho, Jan Chamberlain, Karen Borschmann, Toby Cumming, Liam Johnson, David Lawson, Tina Kaffenberger, Sarah-May Blaschke, and Michelle Shannon.

I acknowledge my second research family, the Clinical Cognitive Neuroscience Lab. I sincerely thank Professor Amy Brodtmann for investing in me to develop my clinical research skills. I thank the rest of the team for their endless amount of love and support: Carolina Restrepo, Sheila Patel, Elise Harrison, Gabby Jean, Laura Bird and Remika Mito. A special thank you to Dr Emilio Werden for the hours of mentoring (therapy sessions).

I wish to thank the exceptional and hilarious group of friends that I met during my candidature: Will Khan, Dearne Stewart, Karen Moss, Emily Dalton, Ruby Lipson-Smith, and Stanley Hung. To my friend, Elie Gottlieb, you drive me insane. We started this journey together, and I could not have asked for a better partner in crime.

This entire experience would not have been the same without my friends and family. I want to thank Jill Naylor, Louise Lindberg, Angel Chen, Rushani Epa and Elise Marc. In particular, I am indebted to Elise's immeasureable love and support. I am fortunate to be surrounded by intelligent, vivacious and hard-working women who inspire me every day.

Thank you to my wonderful family: Kavesha, Pradeep, and the McGrath's. I thank my partner, Liam, for his constant source of love and support, and for somehow always mitigating my concerns with very few words. Finally, thank you to my incredibly supportive Dad. I can only hope that I continue to make up for all the sacrifices you have made. You are the reason I am who I am today, and who I aspire to be in the future. Thank you for showing me that success is determined by hard work, courage and most importantly, compassion.

Table of Contents

Abstract **2**

Declaration ***5***

Preface **6**

Acknowledgements **10**

Abbreviations **17**

Chapter 1 ..... **18**

Introduction ***18***

1.1. Early mobilisation post-stroke: use of population-level evidence for individual patient decision-making 18

1.2. Research objective and questions 21

1.3. Outline of thesis chapters 23

Chapter 2 ..... **27**

Evidence-Based Clinical Decision-Making ***30***

2.1. Clinical decision-making 27

2.2. Evidence-based medicine 28

2.3. Clinical practice guidelines as decision support tools 31

2.4. Stroke recovery 39

2.5. Measuring post-stroke recovery 41

2.6. Early mobilisation of patients post-stroke 44

2.7 Summary and research gaps in the literature 55

Chapter 3 ..... **58**

Study 1: The Utility and Limitations of Clinical Practice Guidelines to Support Early Mobilisation Decision-Making After Stroke ***58***

3.1. Chapter overview 58

3.2. Introduction 60

3.3. Methods 63

3.4. Results 70

3.5. Discussion 77

3.6. Chapter contribution to the thesis objective 82

Chapter 4 ..... **83**

Study 2: Early Mobilisation Post-Stroke: A Systematic Review and Meta-Analysis of Individual Participant Data ***83***

4.1. Chapter overview 83

4.2. Introduction 85
4.3. Methods 86
4.4. Results 90
4.5. Discussion 101
4.6. Chapter contribution to the thesis objective 104

Chapter 5 105

Study 3: Clinical and Systems of Care Factors Contributing to Individual Patient Decision-Making for Early Mobilisation Post-Stroke 105

5.1. Chapter overview 105
5.2. Introduction 107
5.3. Methods 108
5.4. Results 111
5.5. Discussion 118
5.6. Chapter contribution to the thesis objective 123

Chapter 6 124

Study 4: Utility-Weighted Modified Rankin Scale: Still Too Crude to Be A Truly Patient-Centric Primary Outcome Measure? 124

6.1. Chapter overview 124
6.2. Introduction 126
6.3. Methods 127
6.4. Results 130
6.5. Discussion 137
6.6. Chapter contribution to the thesis objective 140

Chapter 7 141

Study 5: Look Closer: The Multidimensional Patterns of Post-Stroke Burden behind the modified Rankin Scale 141

7.1 Chapter overview 141
7.2. Introduction 143
7.3. Methods 144
7.4. Results 147
7.5. Discussion 153
7.6. Chapter contribution to the thesis objective 157

Chapter 8 158

General Discussion 158

8.1. Introduction 158
8.2. Summary of key findings and contributions 159
8.3. Limitations and future directions 168
8.4 Conclusions 176

References **178**

Appendices **179**

List of Tables

Table 2.1. Domains and items of the AGREE-REX....36

Table 2.2. Early mobilisation clinical practice guidelines published *prior* to AVERT....51

Table 2.3. Early mobilisation clinical practice guidelines published *after* AVERT....53

Table 3.1. Summary of AGREE-REX item scores....72

Table 3.2. Demographic characteristics of expert stroke clinicians....74

Table 4.1. Prespecified variables....88

Table 4.2. Summary of individual trial-level characteristics....91

Table 4.3. Early mobilisation dose characteristics....93

Table 4.4. Baseline characteristics of all patients....95

Table 4.5. Primary and secondary outcomes....96

Table 4.6. Risk of bias assessment....98

Table 4.7. SRRR recommendations and TIDieR checklist items....100

Table 5.1. Demographic characteristics of expert stroke clinicians....112

Table 5.2. List of clinical factors considered important to early mobilisation decision-making....113

Table 5.3. List of non-clinical factors influencing early mobilisation decision-making....115

Table 6.1. Characteristics of 2030 AVERT stroke patients....131

Table 6.2. ROC curve analyses between sequential mRS categories based on AQoL utility values at 3 and 12 months....133

Table 6.3. Comparison of UW-mRS values and derivation methods between AVERT and previously published studies....136

Table 7.1. AUC values (95% CI) for discrimination between sequential mRS categories for individual domains of post-stroke burden at 3 months....149

List of Figures

Figure 1.1. Thesis framework 22

Figure 2.1. Updated model for evidence-based clinical decisions. 29

Figure 2.2. Thesis framework with key gaps in the literature. 57

Figure 3.1. Conceptual framework integrating normative and empirical components. 64

Figure 3.2. Proportion of CPGs addressing AGREE-REX criteria 73

Figure 3.3. The areas of uncertainty and key recommendations for future CPGs 76

Figure 4.1. Modified PRISMA flowchart 92

Figure 4.2. Forest plots of subgroup analyses 97

Figure 5.1. Summary of characteristics representing responders and non-responders to EM decision-making 117

Figure 6.1. Distribution of AQoL utility values across mRS categories at 3 and 12 months. 132

Figure 6.2. Misclassification matrices at 3 and 12 months post-stroke 133

Figure 6.3. Distribution of change in AQoL-4D utility values over time 134

Figure 6.4. UW-mRS values from different studies applied to the AVERT dataset at 3 and 12 months (mean, 95% CI). 135

Figure 7.1. Distribution of AQoL-4D and BI domain scores stratified by mRS 148

Figure 7.2: Level of performance in (A) Household Tasks, (B) Mobility and (C) Self-care items at 3 months 150

Figure 7.3. Mean AQoL-4D and BI domain scores per mRS category 151

Figure 7.4. The most impacted domain (lowest mean score) to the least impacted domain (highest mean score) per mRS category. 152

Figure 7.5. mRS 0–2 vs. 3–5 patterns of post-stroke burden. 152

Figure 8.1. Thesis framework with contributions and future directions 160

Figure 8.2. Representation of the contribution of collective works from this thesis embedded within the EBM model and CPGs to support decision-making. 167

Abbreviations

ADL	Activities of Daily Living
AQoL-4D	Assessment of Quality of Life
AUC	Area Under the curve
AVERT	A Very Early Rehabilitation Trial
BI	Barthel Index
CART	Classification and Regression Tree
CI	Confidence Interval
CPG	Clinical Practice Guideline
ICF	International Classification of Functioning, Disability and Health
IQR	Interquartile Range
mRS	modified Rankin Scale
NHMRC	National Health and Medical Research Council
NIHSS	National Institute of Health Stroke Scale
OR	Odds Ratio
ROC	Receiver Operator Characteristic
rtPA	Recombinant Tissue Plasminogen Activator
SD	Standard Deviation
SEVEL	Stroke and Early Vertical Positioning
UW-mRS	Utility-weighted modified Rankin Scale
VERIS-Brazil	Very Early Rehabilitation in Acute Ischemic Stroke
VERITAS	Very Early Rehabilitation or Intensive Telemetry After Stroke
WHO	World Health Organization

Chapter 1

Introduction

1.1. Early mobilisation post-stroke: use of population-level evidence for individual patient decision-making

Evidence-based medicine is fundamental to modern clinical decision-making and healthcare delivery.[1] The movement emphasises the importance of clinical decisions to prudently balance best available research and clinical expertise with the values and preferences of patients.[2,3] Clinical decisions are a *contextual, continuous,* and *evolving* process that involves evidence *generation*, *evaluation*, and *interpretation* to select an *evidence-based* decision that is tailored to the patient or setting.[4] Yet, generated evidence from randomised controlled trials (RCT), often expressed as an average population-level treatment effect, does not directly support decision-making at an individual patient-level.[5,6] Clinical practice guidelines (CPGs) are decision support tools that aim to optimise this translation and the *accuracy, efficacy and adaptability* of clinical decisions.[7] In this thesis, the translation of RCT evidence to inform individualised decision-making about early mobilisation post-stroke was employed as the exemplar.

Early mobilisation (EM), defined as sitting out of bed, standing or walking early after stroke, is an important constituent of acute stroke unit care.[8] Since poor functional recovery is associated with immobility-related complications, the rationale for EM is to prevent and reduce prolonged bedrest during hospitalisation.[8] Contrary to this theoretical belief, the publication of the single largest RCT of 2104 stroke patients, A

Very Early Rehabilitation Trial (AVERT), demonstrated poorer disability outcomes (modified Rankin Scale) at 3 months post stroke in the EM group compared to usual care.[9] Globally, CPGs began recommending against *intensive* out-of-bed activity starting within 24 hours post-stroke.[10] While a cautious approach within 24 hours from stroke onset is increasingly recommended; EM is still commonly recommended to start between 24-48 hours post-stroke, with little best practice guidance on clinical decisions following first initiation of activity.[10] There is also concern that the results from AVERT may lead to protocols that prolong bedrest, particularly in under-resourced settings, which may cause unwarranted harm.[11] The variability in the interpretation of EM evidence, and consequently, the variability in CPG recommendations necessitate clinicians to rely on clinical judgements and expertise to inform their decisions about EM. The degree of uncertainty in clinical decision-making to deliver best practice care related to this intervention presented an ideal research area for this thesis.

While some guidance about EM are outlined in CPGs, tailoring decisions to heterogenous stroke patient population with inherently different recovery patterns is made more complex in the absence of evidence from RCTs, including AVERT, that are sufficiently powered to detect subgroup-treatment effects. Systematic reviews and meta-analyses of RCTs may detect these effects, however, conventional aggregated methodology inhibits such capabilities, and does not effectively consider the complex and multifaceted nature of stroke rehabilitation trials and intervention reporting.[12,13] This might explain the findings from the recently published Cochrane Review which found that EM did not reduce the proportion of people who survived or made a good recovery after their stroke despite the inclusion of AVERT.[14] Exploring methods of RCT data aggregation at an individual patient level shows promise as a way to

overcome some of the limitations with current systematic review methods. This may provide more precise estimates of effects that better support individualised clinical decision-making in EM.

While the publication of AVERT provided important high level evidence about EM, it also raised other questions.[10] How these results, and the current evidence base, have influenced clinicians in their EM decision-making is not known. Understanding how clinicians alter their decisions based on their interpretation of the risks and benefits of EM would help uncover the ways in which decision support tools could be developed and improved. Clinical decisions should also occur with due consideration of the patient's values and preferences.[2,3] The use of patient-centred outcomes provides a way of measuring health outcomes from the patient's perspective, and helps to better communicate the benefits, risks and costs of an intervention to inform decision-making at an individual patient-level. The modified Rankin Scale is an extensively employed and endorsed primary outcome measure of post-stroke disability. The International Classification Functioning, Disability and Health (ICF) Core Set for Stroke highlights the necessity for RCT outcomes, such as the mRS, to describe a more representative spectrum of problems in the functioning of patients post-stroke.[15] The utility-weighted mRS (UW-mRS) has been proposed as a solution to the current limitation of the mRS, by enhancing the measure to reflect patient perspectives of quality of life.[16,17] The ability of the UW-mRS and mRS to represent post-stroke burden for an individual patient is yet to be investigated. However, given the extensive use of the mRS as a primary outcome of post-stroke burden, and that the mRS is not likely to be challenged for this top position, the opportunity to examine the degree to which the mRS could be considered 'patient-centred' warrants further exploration.

In summary, the use of population-level evidence to inform individualised decision-making is complicated by the multifaceted and complex nature of EM evidence and interventions, CPG development and reporting, and the characteristics of stroke patients and their outcomes. This largely unexplored research area formed the focus of this thesis.

1.2. Research objective and questions

The overall *objective* of this thesis was to investigate the barriers and enablers in the effective use of population-level evidence to inform individual patient-level clinical decision-making for EM post-stroke.

Based on the principles of evidence-based clinical decision-making, which emphasises the importance of evidence *generation*, *evaluation*, and *interpretation* to select an *evidence-based* decision that is *tailored to the patient* or setting, and the identified key gaps in the literature detailed in Chapter 2, four sets of research questions emerged to achieve this objective:

1. *What are the utility and limitations of current CPGs as decision support tools that facilitate or inhibit the use of population-level evidence to support individual patient-level clinical decision-making for EM post-stroke?*

   Given current uncertainties in the EM field, by answering this question, we can understand how to improve the development of CPGs to better support clinical decisions.

2. *How can we use an individual participant data meta-analysis of EM RCTs to investigate patient-specific estimates of subgroup and dose effects of EM?*

   Exploration of EM evidence at an individual patient-level will allow appropriate tailoring of evidence-based clinical decisions to the patient.

3. *Based on individual interpretation of the potential risks vs. benefits of EM to specific patients, what clinical and non-clinical factors do clinicians consider to tailor their decisions?*

   This will identify important factors in individual patient-level decision-making that needs to be substantiated by empirical investigations to better support clinical decision-making.

4. *How well do the mRS and UW-mRS reflect post-stroke burden for an individual patient, and to what degree might they be considered 'patient-centred'?*

   This will ensure appropriate selection of outcomes in empirical investigations and CPG development so that clinical decisions can be appropriately tailored to the patient and context.

The identified research questions are grouped into four research themes and are addressed by individual studies as shown in Figure 1.1.

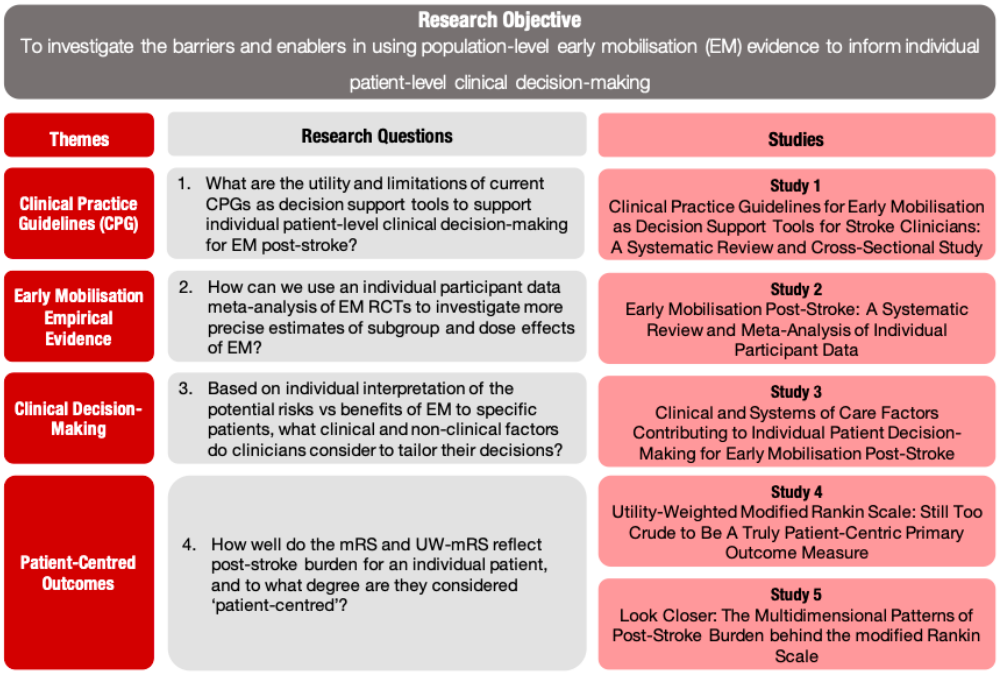


Figure 1.1. Thesis framework. Four primary research themes with corresponding research questions and related studies are depicted in this framework.

Overall, answering these four sets of questions will provide opportunities to optimise the translation of population-level EM evidence to inform individual patient-level clinical decision-making post-stroke. Each of the studies that address these questions are outlined in Figure 1.1 and are detailed below.

1.3. Outline of thesis chapters

Chapter 2 provides a review of the literature relevant to the objective of this thesis and identifies the major gaps that the empirical and analytical work in this thesis aims to address. The chapter begins by exploring how clinical decisions are made within the context of evidence-based practice. The role of CPGs as decision support tools to facilitate the translation of best available evidence to inform individualised decision-making was discussed before introducing the burden of stroke, measures of post-stroke recovery, and a brief history of EM practice and its role in acute stroke unit care. Finally, the changes in EM decision-making and CPGs due to emerging evidence was described. Each subsequent chapter contains evidence reviews relevant to the questions, and details the research conducted to address the objective of this thesis.

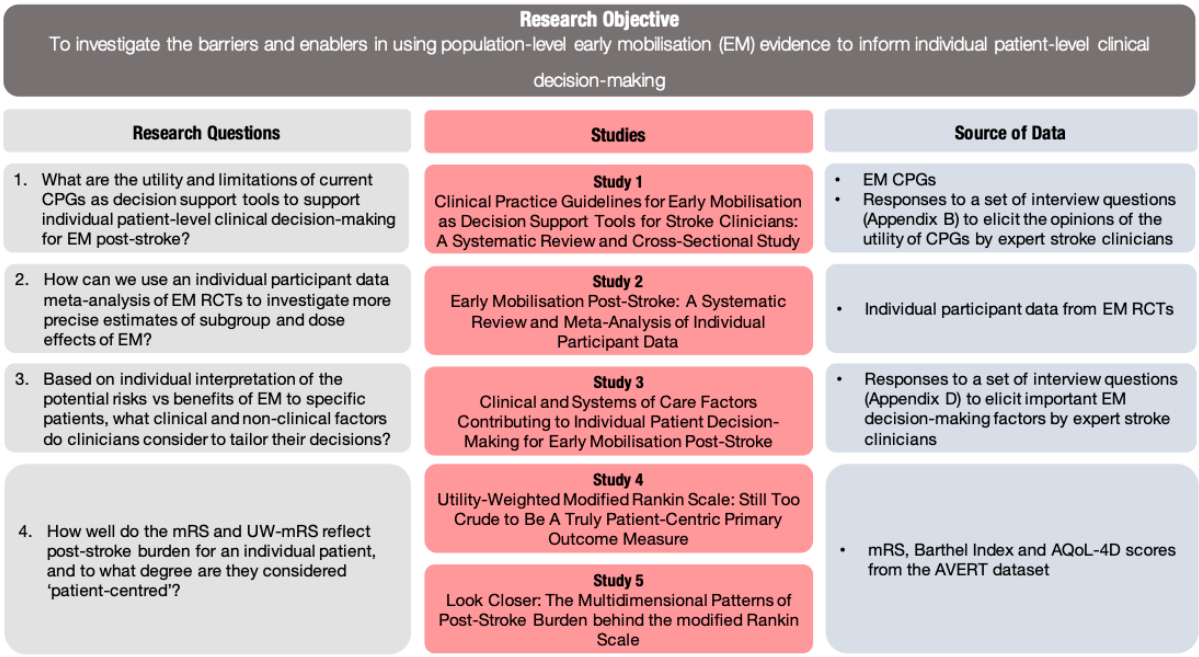


Figure 1.2: Thesis framework with primary sources of data. The sources of data are mapped to the research questiond and studies. Studies 1 and 3 were designed to answer two distinct and independent research questions (Research questions 1 and 3) as shown in Figure 1.2. A single sample of expert stroke clinicians were interviewed using two distinct sets of interview questions (see Appendix C and E, respectively). For ease of data collection, single one-on-one sessions were arranged with the expert stroke clinicians to collect two distinct sets of data.

Chapters 3-7 provide the results from the five studies (Studies 1-5) undertaken to achieve the research objective and the sources of data for each study are summarised in Figure 1.2. Each study is presented as the *final author accepted* or *submitted* version of the manuscript. Therefore, consistent formatting and numbering of headings, tables and figures was maintained throughout all chapters. Each chapter begins with a chapter overview to introduce the the accepted or submitted journal manuscript followed by a concluding section to briefly summarise the contribution of the study. These contributions are subsequently discussed in greater detail in the General Discussion (Chapter 8).

Chapter 3, contains a ***submitted manuscript (under review)*** titled *Clinical Practice Guidelines for Early Mobilisation as Decision Support Tools for Stroke Clinicians: A Systematic Review and Cross-Sectional Study*.***** Study 1 investigated the utility and limitations of CPGs for EM post-stroke as decision support tools for expert stroke clinicians. A two-pronged approach was used: (1) a systematic review of EM CPGs, and (2) a cross-sectional study to interview expert stroke clincians using a set of interview questions to elicit their opinions on the utility of EM CPGs.

Chapter 4 contains a ***published manuscript*** titled *Early Mobilisation Post-Stroke: A Systematic Review and Meta-Analysis of Individual Participant Data.* In Study 2, a meta-analysis using individual participant data from six RCTs was conducted to investigate the safety and efficacy of EM compared to usual care, and to investigate subgroup and dose effects of EM.

Chapter 5 contains a ***submitted manuscript*** titled *Clinical and Systems of Care Factors Contributing to Individual Patient Decision-Making for Early Mobilisation Post-Stroke*.* Study 3 reports the results from a cross-sectional study of expert stroke clinicians which aimed to identify patient and systems of care factors involved in individual patient decision-making, and high-interest patient subgroups.

Chapter 6 contains a ***published manuscript*** titled *Utility-Weighted Modified Rankin Scale: Still Too Crude to Be A Truly Patient-Centric Primary Outcome Measure.* In Study 4, the construct validity of the utility-weighted modified Rankin Scale (UW-mRS) as a patient-centred outcome measure was investigated. Data from the large, international dataset from AVERT was used to conduct this study.

Chapter 7 contains a ***published manuscript*** titled *Look Closer: The Multidimensional Patterns of Post-Stroke Burden behind the modified Rankin Scale.* In Study 5, quality

of life and activities of daily living domains were used to demonstrate patient specific patterns of post-stroke burden across the mRS and UW-mRS. Data from the large, international dataset from AVERT was used to conduct this study.

***NB:** Studies 1 and 3 were designed to answer two distinct and independent research questions (Research questions 1 and 3) as shown in Figure 1.2. A single sample of expert stroke clinicians were interviewed using two distinct sets of interview questions (see Appendix C and E, respectively). For ease of data collection, single one-on-one sessions were arranged with the expert stroke clinicians to collect two distinct sets of data.

Chapter 8 contains the *General Discussion*, which summarises the findings from the thesis. In this chapter, the key implications of the findings, the overall limitations, and the future directions that may improve the translation of evidence to inform individualised decision-making for EM post-stroke are discussed.

Chapter 2

Literature Review

2.1. Clinical decision-making

As described in Chapter 1, clinical decision-making is a *contextual*, *continuous*, and *evolving* process, where evidence is *generated*, *evaluated*, and *interpreted* to select an *evidence-based* choice of action that is tailored to the patient and their preferences.[4] Decisions involve a number of complex psychological processes and factors such as those relevant to the decision itself (context or framing), situational factors (psychosocial context or systems of care factors) and the characteristics of the decision-maker (discipline, level of expertise, values and preferences).[18]

There are several theories of clinical decision-making or clinical reasoning. One theory is that decisions involve a dual process system: *intuitive* processing is largely autonomous generally acquired through repeated experience and the biggest source of cognitive failure, while *analytic* processing is conscious, logical and resource intensive.[19,20] The ways in which people attend to, organise and remember information will affect how that information will be used to impact the final decision outcome.[21,22] Another well-known decision theory involves "hypothetico-deductive" reasoning: the process of information gathering, hypothesis generation, hypothesis testing and reflection.[23]

There is no coherent theory of decision-making, but ultimately, it is based on the clinician's current set of resources (evidence, past experiences, skills and knowledge, patient characteristics and preferences) and the possible consequences of their choice

(how will this affect the outcome of the patient)–the "ability to collect the *right* resources and to take the *right* action for the *right* patient at the right time and for the *right* reason".[24] This embodies the principles of evidence-based medicine (EBM).

2.2. Evidence-based medicine

In the early 1990s, the EBM movement began in response to the growing demand for medical practice to be informed by trustworthy, empirically- collected evidence. The term, "evidence-based medicine", was first coined in 1991 with the objective to educate clinicians on how to assess the credibility of research evidence, understand the results of empirical investigations, and determine the appropriate strategies to apply the acquired evidence to inform clinical decision-making.[3] Three epistemological principles exist for EBM decision-making: (1) not all evidence is created equal, (2) clinical decisions should be informed by a collection of high-quality evidence, and (3) effective individualised decision-making requires the consideration of patients' values and preferences within the given environment and context.[1,25] EBM does not merely support clinical decision-making, it improves the consistency and delivery of care, yields better health outcomes, and assist with ethical dilemmas.[26]

2.2.1 The current model of evidence-based decisions

EBM is described as "the integration of best research evidence with clinical expertise and patient values".[27] The conceptual model of EBM that reflects this definition is depicted in Figure 2.1 The model is intended to be prescriptive (how clinical decisions should be made) rather than descriptive (how clinical decision are made), and is composed of four elements: clinical state and circumstances, patients' preferences and actions, research evidence and clinical expertise.[28]

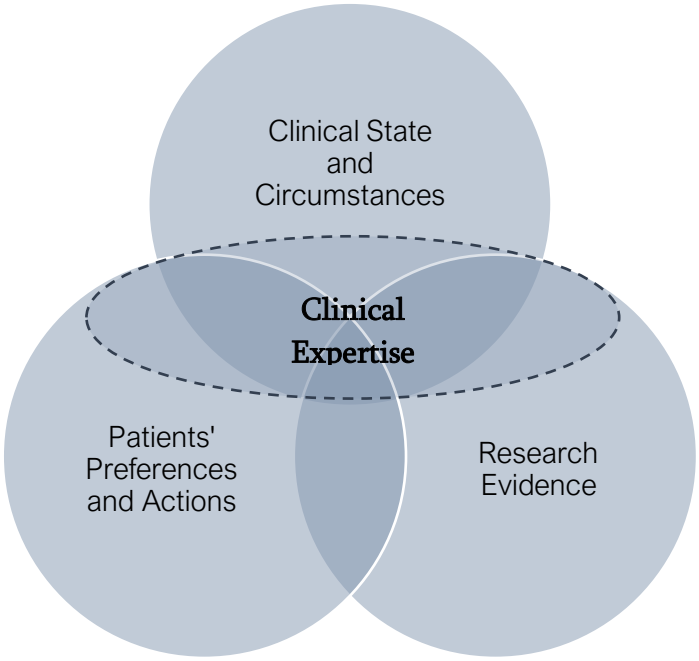


Figure 2.1. Updated model for evidence-based clinical decisions. This model provides a guide for how clinical decisions should be made by depicting three key elements in clinical decisions and the role of clinical expertise in integrating these elements. The model also acknowledges that different weights can be assigned to each decision-making component (depicted visually by varying the sizes of the circles).[28]

The first element, *patients' clinical state and circumstances*, is thought to be a primary component of clinical decision-making. This element emphasises that decisions *should* vary from patient to patient according to individual clinical circumstances. The second element, *patient preference and actions*, relates to the crucial role of individual patient preferences in clinical decision-making. This element also acknowledges the patients' adherence to the clinicians' advice as an important indicator of the decision-making outcome. The third element, *research evidence*, involves gathering, evaluating and interpreting cumulated best-available evidence to inform clinical decision-making. It assists in the quantitative interpretation of the patient's clinical circumstances. The last element, *clinical*

expertise, relates to the general skills involved in clinical practice and the past experiences of individual clinicians. More than that, it is the integrative decision-making element which judiciously negotiates between trade-offs in the patient's clinical state and circumstance, best-available research evidence, and the patient's preferences and actions. Clinical expertise is distinct from clinician preferences, which generally contribute to unwanted variances in clinical decisions.[19,20] This decision-making model provides a conceptual rather than practical framework, but nonetheless, it has had a substantial role in the development of several schemas and decision support tools. It also informs a standard for best practice and a benchmark for clinical practice evaluation.[29]

2.2.2 The impact of evidence based medicine: rise of clinical practice guidelines

The EBM movement has seen the introduction of many EBM-based initiatives[1] including the Cochrane Collaboration's systematic review methodology,[30] standards for the conduct and reporting of research,[31] and approaches in rating evidence quality.[32,33] In particular, the epistemological principles of EBM decision-making which emphasise the need to systematically acquire and appraise evidence while combining the considerations of the clinical context, clinical expertise and patients' values and preferences, saw the rise of CPGs.[1] The need for CPGs was also recognised by the US Institute of Medicine due to substantial variation in clinical practice.[34] In an era of publication overload, a clinicians' capacity to stay abreast of best available evidence is not only challenging, it is impractical.[35,36] CPGs provide a mechanism for communicating "pre-processed" evidence-based recommendations to

clinicians in order to facilitate the effective implementation of EBM decision-making.[1]

Despite the contributions of EBM, there are some criticisms of the model that should be considered for the remainder of this thesis. Firstly, EBM has been criticised for seemingly discouraging clinical reasoning and focusing too heavily on formulaic approaches to clinical decision-making.[37,38] However, supporters of EBM have argued that the model has promoted the need to consider patient values when making preference-sensitive decisions.[1] Another criticism is the strict adherence to the evidence hierarchy however, as Djulbegovic et al argue, "no critic ever suggested that reliable evidence should not be a key to effective problem solving and decision making."

2.3. Clinical practice guidelines as decision support tools

When faced with such complex and important decisions, decision support tools provide clinical experts with strategies to improve the decision-making process and outcome.[39] This approach is primarily about problem solving rather than accumulating theoretical knowledge.[40] It is based on the principles of design science, which prioritises utility (to serve human purposes) rather than the truth (to understand reality).[41] In clinical practice, the success of a clinical decision support tool is contingent on how well the tool: (1) reflects the nature of the clinicians' decision-making process, (2) can be integrated into clinical practice.[42] By providing evidence-based decision support to clinicians, CPGs improve quality of healthcare and patient outcomes, reduce unwarranted variation in clinical decision-making, optimise healthcare resource use, provide guidance for consumers, and identify gaps in knowledge to better support clinical decision-making.[26,43-47]

2.3.1 What are clinical practice guidelines

The US Institute of Medicine defines CPGs as "statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options".[48] Based on this definition, CPG development involves two components: (1) a systematic review of evidence based priority clinical questions that focuses on the strength of the evidence base to inform decision-making, and (2) the formulation of recommendations based on the evidence base and value judgement to provide a description of the relative value placed in outcomes, benefits and harms and resource use. CPGs can be informed by a range of evidence-based and consensus-based resources including RCTs, systematic reviews, observational studies, physiological experiments, case series, case reports, expert opinions or the experience of individual clinicians.[49,50] As the first EBM epistemological principle states, *not all evidence is of the same quality*. In response to this, the EBM-initiated GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) approach was developed to provide a sophisticated framework for rating the quality of evidence by assessing study design, risk of bias (study strengths and limitations), precision, consistency (variability in results between studies), directness (applicability), publication bias, magnitude of effect, and dose-response gradients, and grading the strength of recommendations.[33,51-53] Taking evidence at any level, quality, relevance or strength and using it to derive recommendations that support evidence-based clinical decision-making "depends on judgement, experience and good sense".[54] As a result, guideline development groups include content experts e.g. researchers, clinicians and health professionals; methodologists; guideline development or implementation experts; end-users e.g. specific types of health workers or policy

makers; and consumers e.g. patients or carers, who formally consider the scientific evidence, expert opinions, clinical experience, and other relevant information so that CPG recommendations reflect a core set of desirable attributes.[7,55-58] These attributes include validity, reliability/reproducibility, clinical applicability, clinical flexibility, clarity, involving a multidisciplinary process, scheduled review of new data, and meticulous documentation of guideline development procedures.[7]

2.3.2 Standards for CPG development

Evidence is not created equal—neither are CPGs. A study reviewing 431 guidelines found that 88% gave no information on searches for published studies, 82% did not give any explicit grading of the strength of recommendations, and 67% did not report any description of the type of stakeholders.[59] Since there are currently over 3749 documents from 96 organisation and 87 countries present in the library of the Guidelines International Network,[60] CPG development processes can vary.[61,62] Generally, the non-linear development process involves 4 broad phases: (1) preparing for CPG development – topic selection, defining scope and purpose, adapting existing CPGs, forming a multidisciplinary guideline development group, and involving consumers, (2) systematically reviewing evidence – formulate clinical questions, systematically searching, including and/or excluding identified research, and research appraisal, (3) drafting the CPG – developing recommendations, implementation strategy, public consultation, summary version of CPG and (4) reviewing CPG – impact evaluation, and plans for revision and update.[61] Several guideline development handbooks are utilised to outline the methodology by which CPGs should be developed. Handbooks are produced by numerous international organisations such as the National Health and Medical Research Council of Australia

(NHMRC), the Council of Europe, the National Institute for Health and Clinical Excellence in the UK, the New Zealand Guidelines Group, the Scottish Intercollegiate Guideline Network, and the World Health Organization (WHO).[54,56,63-65] A study comparing the advice and guidance of six handbooks found a strong agreement on several guideline development elements.[61] While several systematic reviews have found tasks relevant to ethical issues, piloting, updating CPGs and implementation strategies to be poorly described.[62,66,67]

2.3.3 Clinical practice guideline appraisal tools

In parallel to guideline development handbooks, tools for their appraisal are also available to support the evaluation of the quality of guidelines by the end user.[68] There are over 40 CPG appraisal tools with many assessing whether the literature search and evaluation, and synthesis and presentation of the evidence follow the principles of EBM.[68] In particular, the Appraisal of Guidelines for Research & Evaluation (AGREE) Enterprise, has had significant contributions to the science and advancement of CPGs. This includes the development of several evaluation and reporting instruments such as the AGREE II[69] (methodological rigour and transparency in CPG development), My AGREE PLUS[70] (online AGREE II appraisal platform), AGREE-HS[71] (quality of health systems guidance), and CheckUp[72] (completeness of reporting in updated CPGs). The AGREE II tool is the most validated and extensively used tool to evaluate the quality of CPGs, and provide a methodological framework for developing and reporting CPGs.[69,73,74] The tool consists of 23 items grouped into 6 domains (scope and purpose, stakeholder involvement, rigor of development, clarity of presentation, applicability, and editorial independence).[69] Over the last two decades, systematic reviews have shown an

increase in the general quality of CPGs, yet the domains rigour of development, editorial independence and clinical applicability are often unsatisfactorily met.[75-78] Several studies reveal clinical applicability is consistently the worst performing quality domain,[79-83] which is a particularly concerning inadequacy since the success of a decision support tool is contingent on how well the tool can be integrated into clinical practice.[42]

2.3.3 Assessing the clinical applicability of CPGs

The validity, reproducibility and utility of guidelines are influenced by more than just the availability of good evidence or rigorous development strategies.[84-86] Evidence, which is often expressed as an average treatment effect at a population-level, cannot directly support clinical decision-making at an individual patient-level.[5,6] Research of *clinical applicability* and *implementability* of CPGs to support individualised decision-making is crucial in achieving the principles of EBM.[87] In consideration of this, the AGREE Enterprise created and employed the Guideline Implementability for Decision Excellence Model (GUIDE-M).[88] This model informed the recent development of the Appraisal of Guideline Research and Evaluation-Recommendations Excellence (AGREE-REX) tool[89] to assess the clinical credibility and implementability of CPGs.

The AGREE-REX tool[89] consists of nine items structured within three theoretical domains: Clinical Applicability, Values and Preferences, and Implementability (Table 2.1). All nine AGREE-REX items contain an operational definition of the item and a number of item-specific criteria (ranging between 2 to 10) for CPGs to be evaluated (Appendix B).

Table 2.1. Domains and items of the AGREE-REX

Domains	Items
Clinical Applicability	1. Evidence 2. Applicability to Target Users 3. Applicability to Patients/Populations
Values and Preferences	4. Values and Preferences of Target Users 5. Values and Preferences of Patients/Populations 6. Values and Preferences of Policy/Decision-Makers 7. Values and Preferences of Guideline Developers
Implementability	8. Purpose 9. Local Application and Adoption

In order to support clinical decision-making at an individual patient-level, CPGs need to provide recommendations that are based on best-available evidence, consider the clinical applicability to the patient and setting, and explicitly state the options for, and implication of, different care options, i.e. trade-off between harms vs. benefits, to support clinical judgements.[7,67,88] This degree of decision support can be assessed using four specific AGREE-REX domains:

1. **Evidence:** thorough review of the quality and results of the available evidence.

2. **Applicability to Target Users:** the degree to which the recommendations are applicable to the guideline's target users' practice context.

3. **Applicability to Patients:** the extent to which the anticipated outcomes of the recommended action are relevant for, and valued by, the intended patients/populations.

4. **Values and Preferences of Target Users:** the relative importance that the target users of the guidelines place on the outcomes of interest (e.g., survival, adverse effects, quality of life, cost, convenience). These are important to consider since they influence the acceptability and adoption into practice.

Even if the evidence base is trustworthy, and the development of the CPGs are of high methodological quality, the utility of CPGs to improve the translation of population-level evidence to individualised decision-making is contingent on how well CPGs address these domains.

2.3.4 Barriers in using CPGs to support individualised decision-making

While the use of CPGs can improve patient outcomes, CPGs are not frequently utilised.[90,91] Effective use of CPGs is driven by the creation and communication of CPG content that involves target user involvement, evidence synthesis, and ensuring clinical applicability and implementation feasibility.[92] Moreover, the rate and extent of adoption is determined by the interaction between the characteristics of the CPG, the intended CPG target users, and the particular context of clinical practice.[93] A non-applicable CPG for a specific patient subgroup or clinical setting will have poor adherence while an ambiguous CPG is less likely to be followed. In a sample of Dutch general practitioners, the perceived barriers in the uptake of CPGs were related to lack of applicability or lack of evidence (68% of recommendations), environmental factors such as organisational constraints (52%), lack of knowledge regarding the guideline recommendations (46%), and guideline factors such as unclear or ambiguous guideline recommendations (43%).[94]

The known inadequacies of many CPGs to be clinically applicable[79-83] may be a critical barrier in using CPGs as decision support tools, and potentially contribute to the observed non-adherence. As the principles of EBM suggest, using population-level evidence (measured as an average treatment effect) is not conducive to decision-making for individual patients or subgroups that deviate from the population average. Any differences in treatment effects often reflects the *heterogeneity* of patients, their

responsiveness to treatment, *vulnerability* to adverse effects and *utility* for different outcomes.[95-97] Effective use of CPG not only depends on the applicability to clinical practice, but it also is dependent on how well the recommendations respond to patient variability. This is particularly evident in the underutilisation of CPGs for patients who are older and/or have co-morbidities that are most likely to benefit from treatment, but most susceptible to the potential harms.[47,98-101]

The validity and reliability of evidence alone is not enough for CPGs to have their intended effects.[102] The specificity of recommendations is also important. Shekelle et al.[103] found that physicians receiving non-specific CPGs ordered fewer diagnostic tests for appropriate clinical vignettes than physicians receiving no guidelines. While compared to physicians receiving non-specific guidelines, physicians receiving specific guidelines ordered significantly more diagnostic tests for appropriate vignettes and significantly fewer diagnostic tests for inappropriate vignettes. In this case, specific guidelines may change physician decision-making for the better, while non-specific guidelines may change physician behaviour for the worse.

Alternatively, barriers in the uptake of CPGs to support individualised decision-making may be due to the perceived rigidity of CPGs, insufficient clinical flexibility or loss of clinical autonomy.[36,37,101,104] Indeed CPGs do not necessarily apply to everyone, however, if the magnitude of harms, benefits or cost differs for different groups of patients, CPGs are required to specify the options for and implications of different practice options. The current evidence can be leveraged by conducting subgroup analyses of RCTs and conventional meta-analyses to identify effect modifiers to support individualised decision-making.[105] However, RCTs are often underpowered to detect subgroups and conventional meta-analysis methods cannot

unpack the granularity of RCT data. Individual patient meta-analysis can offer a solution by providing individualized estimates, however it is infrequently performed, requires data sharing by trialists, and is resource-intensive.[105,106]

In summary, these key barriers can be interpreted as a *lack of specific and adaptable CPGs to respond effectively to variability in the clinical context (patient and clinician) to appropriately inform individualised clinical decision-making*. The methods that improve the acceptance and utility of CPGs require feedback and involvement from end-users of CPGs.[107] The *decision task, decision-making process and value judgements of CPG target users need to be understood and reflected in the design of CPGs* to ensure the development and delivery of evidence-based recommendations are able to close the gap between how healthcare is delivered and the scientific evidence that supports it.[53,108,109]

In the remaining sections of this chapter, the burden of stroke will be briefly outlined followed by a summary of extensively utilised and newly proposed acute stroke outcome measures. These sections will collectively provide an in-depth discussion of the clinical context that serves as the exemplar to explore the barriers and enablers in translating of RCT evidence to inform individualised decision-making.

2.4. Stroke recovery

Stroke is a major public health concern. It is the second leading cause of mortality and disability worldwide.[110] Stroke is defined as a clinical syndrome persisting for more than 24 hours with sudden focal or global neurological disturbances of vascular aetiology.[111] Cerebral stroke can be divided into ischaemic stroke (IS) and intracerebral haemorrhage (ICH). Approximately 80% of cerebral stroke in the Western population are ischaemic events, while the remaining 20% are a result of

haemorrhage.[112-114] Sudden death within 24 hours of stroke onset occurs in 20% of the population while two thirds have residual neurological deficits that impair daily function and reduce quality of life. [115,116]

Prevention and improved control of risk factors alongside the introduction of several acute stroke therapies have led to decreased mortality rates. Yet, this decline is conversely met with an increase in the prevalence of disability post-stroke, and a growing demand for rehabilitation that is largely unmet. In 2017, more than 475,000 Australians were estimated to have residual neurological deficits that persistently impair daily function.[117] Stroke burden is not limited to biopsychosocial disability; it imposes a significant economic burden. Loss of productivity (approx. AU$3 billion) and healthcare expenditure (approx. AU$881 million) cost Australia an estimated AU$5 billion in 2012.[118] Given the immense burden, efforts must be made to optimise recovery post-stroke.

The effects of stroke may be determined by the site and size of the initial stroke lesion, and by the extent of subsequent recovery.[119] The Stroke Recovery and Rehabilitation Roundtable taskforce defines the term "recovery" to mean: "(1) the change (mostly improvement) of a given outcome that is achieved by an individual between two (or more) timepoints, or (2) the mechanism underlying this improvement in terms of behavioural restitution or compensation strategies".[120] However, the mechanism of recovery is multifaceted and complex potentially involving a combination of spontaneous biological recovery and learning dependent processes,[121] therefore, individual patient recovery patterns are heterogeneous in nature.[122] Nonetheless, any improvement in any domain of post-stroke burden is important to capture and is a "sign of ongoing recovery".[120]

2.5. Measuring post-stroke recovery

As previously described, since stroke is a significant cause of disability and measuring mortality alone does not capture post-stroke burden, measures of functional recovery are favoured. The burden of stroke on patients can include a combination of neurological deficits, inabilities to perform activities of daily living and self-care tasks, neuropsychological dysfunction and neuropsychiatric deficits.[15,123,124] As a result, outcome measures should describe a representative spectrum of post-stroke burden while also capturing meaningful changes relevant to patients.[125] This is particularly important in the selection of outcome for RCTs that often provide evidence that informs clinical decisions and the development CPGs. However, in many acute stroke trials, including AVERT, important outcome measures are typically scored by clinicians or researchers, and only capture a limited aspect of post-stroke burden concerning impairments and an inability to perform activities of daily living.[15,123,124] There is however, no perfect scale and the choice of outcome measure will depend on the research question, stroke population and available resources.

2.5.1. The modified Rankin scale

The modified Rankin scale (mRS) is the most widely favoured[126,127] and extensively endorsed[128-131] measure of "global disability" and primary endpoint for many acute stroke trials.[132] The mRS includes six ordered grades of functional outcome and a death state. The scale is favoured because of its ease of use and intuitive administration, and ability to demarcate effective and ineffective treatments in adequately powered samples. While a single-point shift on the scale is clinically relevant,[133,134] its ordinal nature may make it less responsive or sensitive to change

compared to stroke scales.[135] There are also limitations concerning the reproducibility of scores and its limited power to detect treatment effects when dichotomized, typically into good outcome (mRS 0-1 or 0-2) and poor outcome (mRS >1 or >2).[132,136] The Optimising the Analysis of Stroke Trials (OAST) collaboration recommends preserving the ordinal nature over a dichotomous approach due to its demonstrated superiority however, this approach still raises questions about the clinical interpretability.[137]

While all mRS categories consistently and explicitly refer to a generic degree of post-stroke disability, some categories are varyingly described in terms of the presence of symptoms, ability to perform activities of daily living, walking ability and self-care. Furthermore, the simple descriptions of each mRS category do not include other disability conditions such as post-stroke cognitive decline, mood disorder and fatigue.[138] Despite this, the mRS remains the recommended measure of global disability,[128,130,131,133] and it is thought that this measure "should be retained in its present form for comparability with existing treatment comparisons",[133] including within the field of EM.

2.5.2 Patient-centred outcomes

Disparities between the preferences of patients and physicians have been previously demonstrated.[139-141] As such, patient-centred approaches to best practice have been a growing area of interest.[142-144] Patient- or person-centred care is broadly defined as "care that is respectful of, and responsive to, the preferences, needs and values of the individual patient".[145] As part of this movement, attention has been paid to the identification and development of more patient-centred outcome measures for inclusion in research and clinical practice.[146-148] Outcomes can be developed by

patients (patient-generated), with the participation of patients (patient-centred), ranked or valued by the patient (patient-valued), or without any patient participation or assessed by patients as irrelevant or of no interest or value (patient-irrelevant).[146] Patient-centred outcomes aim to value the patients' perspective of various health outcomes over the perspectives of clinicians and researchers and provide a method to optimise the communication of relevant benefits, risks and costs associated with different clinical decisions. However, it should be noted that there are practical and methodological reasons that trialists have preferred "hard" clinical outcomes to patient-reported outcome measures. There can be disadvantages in the use of patient-reported measure such as the accuracy and reliance on self-appraisal of functioning rather than objective assessments of performance.[149] Nonetheless, it is important to capture stroke burden and health status from a patient's perspective without interpretation by a clinician.

Various generic and stroke-specific patient-centred outcomes have been developed to improve the understanding of the multiple domains affected by stroke beyond physical functioning. Health related quality of life is one measure that can serve to assess aspects of "participation" and capture patients' experiences of stroke recovery. Work by Ali et al.[150] investigated how well standard outcome measures reflect quality of life using the Virtual International Stroke Trials Archive (VISTA) dataset. They found that of the traditional acute stroke outcome measures (mRS, National Institutes of Health Stroke Scale and Barthel Index), the strongest association was found between quality of life and the mRS at 3 months.

2.5.3. Utility-weighted mRS

Utility, centred on health-related quality of life, is a widely accepted patient-centred outcome measure of a patient's preference for a specific health outcome.[151,152] The utility-weighted mRS (UW-mRS) has therefore been proposed and endorsed by the Stroke Therapy Academic Industry Roundtable (STAIR) as an approach to create a more patient-centred outcome measure by reflecting patient perception of quality of life,[16,17,153] and to improve the interpretability and statistical efficiency of the mRS in acute stroke RCTs.[17,154] UW-mRS values are obtained by assigning a mean utility weight to each mRS category based on patient, clinician or societal valuation of various disability health states.[155] Investigations have shown similar efficiency compared to the ordinal mRS, and greater efficiency compared to the dichotomised mRS.[17,154] The perceived benefits of the UW-mRS is reflected in the recent approval as a co-primary outcome measure by the Food and Drug Administration in a thrombectomy trial.[156] Validity, reliability, responsiveness to change, acceptability and feasibility are key criteria for an assessment scale.[138] However, the construct validity of the UW-mRS as a patient-centred outcome measure has yet to be established and this is necessary to enable appropriate communication of relevant benefits, risks and costs associated with different clinical decisions. Further exploration of the 'patient-centredness' of this measure was a focus for this thesis.

2.6. Early mobilisation of patients post-stroke

It is well-established that the delivery of care in organised, multidisciplinary acute stroke units is effective in reducing death and long-term dependency.[157-161] Over the last decade, understanding the components of care that might contribute to the success of acute stroke unit care has grown in interest.[8] Early mobilisation (EM), defined as

sitting out of bed, standing or walking early after stroke, has been considered important in this care model for many years.[162] The susceptibility of patients to inactivity during hospitalisation[163,164] and the increased risk of immobility-related complications such as infections, falls, thromboembolism leading to poor functional recovery forms the central theoretical rationale for EM.[9,165-167] It is also hypothesised that EM may help promote brain recovery, however this is largely based on preclinical studies and theories of brain remodelling in the acute stages post-stroke.[168-172] Equally, some concerns exist about EM which include increased risk of falls,[165] impairment of cerebral blood flow and perfusion due to hemodynamic changes,[173,174] and risk of further bleeding in haemorrhagic strokes[175] or after thrombolysis.[176]

2.6.1 The evidence base behind early mobilisation

EM was first recommended by several local CPGs in Norway and Sweden in the late 1980s. It became a prominent area of research after the publication of a Norwegian RCT in 1991.[158] The results of the trial demonstrated a marked reduction in death and disability in patients managed in a stroke care unit (with a focus on early mobilisation) when compared with general ward care. This trial was included in the landmark Stroke Unit Cochrane review, which further confirmed the benefits of organised stroke unit care.[177] In 1994, the American Heart Association published *Guidelines for the Management of Patients With Acute Ischemic Stroke* which strongly recommended EM.[178]

Since then, RCTs and subsequent systematic reviews have investigated the safety and efficacy of EM. Findings however have been inconsistent.[14,179] In several trials investigating EM starting within 24 to 72 hours of stroke onset compared to usual

stroke unit care (later mobilisation and/or monitoring) no significant effects on complication rates, mortality or global disability (modified Rankin Scale) were found. An individual participant data meta-analysis of two RCTs conducted in 2010, demonstrated that EM patients had significantly greater odds of independence compared with standard care patients (adjusted odds ratio (OR), 3.11; 95% confidence interval (CI), 1.03–9.33).[180] These inconsistent findings might be attributed to varied intervention protocols and small sample sizes.

Thus, uncertainty about the practice of EM from these early, small trials, was greatly impacted by the findings from the single largest EM RCT, A Very Early Rehabilitation Trial (AVERT; n=2104), conducted in 56 acute stroke units in five countries.[9] AVERT demonstrated fewer patients in the EM group had a favourable functional outcome (modified Rankin Scale score 0-2) than those in the usual care group (n=480 [46%] vs. n=525 [50%]; adjusted OR 0·59–0·59–0·004). Interestingly, no effect on immobility associated complications or walking recovery was found. When interpreting these results, it should be noted that over the course of the trial, a gradual decrease in time to first mobilisation was observed in the usual care group (median of 28 minutes per year). As such, the time difference in first mobilisation between groups was only 4 hours: 18.5 hours (median; IQR, 12.8–22.3) vs. 22.4 hours (16.5–29.3). Moreover, 92% of EM patients and 59% of usual care patients were mobilized within the first 24 hours; possibly suggesting that the often-larger dose of therapy received in the EM group had a greater influence on functional outcome at 3 months post-stroke. Prespecified subgroup analyses of efficacy suggested that patients with severe stroke in the EM group had reduced odds of a favourable outcome at 3 months.[181] Patients with intracerebral haemorrhage also appeared to have signal of reduced odds of a favourable outcome and increased odds of mortality

at 3 months post-stroke. However, these subgroup results were not statistically significant.

The variability of EM delivery in usual stroke unit care and the potential subgroup-effects raises the questions about clinical decisions: (1) What is the optimal mobilisation strategy in the acute period post-stroke to inform individualised clinical decision-making? (3) Which patients are likely to benefit or be harmed by EM?

To provide practical guidance for clinicians on the timing, frequency, and amount of mobilisation, the AVERT investigators performed a pre-specified "dose-response" analysis.[181] This analysis suggested shorter, more frequent mobilisation early after acute stroke is associated with greater odds of favourable outcome at 3 months after controlling for age and stroke severity.[181] However, the investigators cautioned against over-interpretation of these additional analyses. Thus, the interpretation and translation of this multifaceted EM evidence to inform individualised decision-making is complex and challenging.

AVERT was a landmark trial that challenged conventional beliefs and still fuels debate about the importance of EM in the acute stroke care.[182] Some have argued the "danger of falling into the opposite extreme of immobilisation and inactivity" based on these findings.[11] While others speculate about the impact of upright posture on cerebral perfusion on exacerbating ischemia and leading to poor outcome[182-184] or the amplification of therapy-induced cellular stress and increased expression of pro-inflammatory mediators.[185,186] AVERT has nonetheless provided robust evidence that highlights the complexity of EM decision-making involved in acute stroke care. Whether CPG recommendations informed by AVERT should apply to all individuals after stroke is uncertain.[187]

Most recently, the updated EM Cochrane review (nine RCTs, 2958 participants) found EM within 48 hours of stroke onset did not show a difference in the number of patients who survived or had a good outcome post-stroke compared to usual care.[14] However, due to the aggregated nature of the meta-analysis, predictive factors, i.e., stroke severity and age, could not be investigated producing uncertainty in the estimations of effect, and an inability to explore subgroup–treatment effects.

It is obvious the field of EM research is burdened by surprising and/or inconsistent findings that challenge the utility of EM evidence to inform individualised clinical decision-making. The gaps in knowledge concerning the optimal dose of EM (time, amount, frequency), and how to identify patients whom EM might benefit (responders) or harm (non-responders) needed to be investigated.

2.6.2 Early mobilisation clinical decision-making post-stroke

Several studies have investigated the factors influencing EM decision-making; however, a majority were conducted prior to the publication of AVERT. A survey of nurses, physicians, and physiotherapists (n=99) in 2007 emphasised the lack of agreement about what constitutes EM and who should initiate it.[188] Most nurses (62%) and physicians (67%) thought that EM should begin within either 12 or 24 hours of stroke onset, while most physiotherapists (76%) thought that EM should commence after the first 24 hours. The opinions of clinicians concerning the benefit or harms of EM shifted over time even when the evidence based did not. In 2011, a survey of 202 survey of stroke professionals was conducted that demonstrated 60% had concerns about EM in patients with intracerebral haemorrhages.[175] This is noteworthy given there was no evidence of benefits or harms at the time, but the survey results reflect the subgroups effects (albeit non-significant) seen in AVERT four years later. In

contrast, 77% of professionals thought patients with ischemic stroke should be mobilised within 24 hours of stroke onset.[189] There were also differences in the perceived importance of factors affecting decisions about EM based on the health professionals discipline, i.e. doctors were most concerned about the patients' physical ability to mobilize, while physiotherapists were mostly concerned about the patients' medical stability. Arguably, the most interesting outcome from the survey was that only 19% of professionals required a large RCT or a systematic review to be convinced of the benefits of EM.

It is not surprising that the risk of symptomatic intracerebral haemorrhage post-thrombolysis may influence EM decision-making. A case-cross over study of 54 clinicians using hypothetical vignettes revealed that neurological decline, neurological decline with symptomatic intracerebral haemorrhage, infection of uncertain cause, severe chest infection, severe stroke, drowsiness, and confusion significantly influenced decisions to mobilise patients treated with thrombolysis early post-stroke.[190] In the absence of decision support or guidance, there seems to be a tendency to take a conservative approach, even when there is little or no evidence to support a conservative approach.[191-193]

Since many of these studies were published before AVERT, we currently do not have a sound understanding about how clinicians make individualised clinical decisions about which patients should be mobilised, particularly in the absence of clinical trials that are powered to detect subgroup effects. Therefore, any variances in clinical decision-making is likely due to differences in the *interpretation* and *understanding* of the *heterogeneity of acute stroke factors*, and the theorised or experience of the

benefits and harms of EM to different patient subgroups. This highlights the need for clinically applicable CPGs to provide decision support in these areas.

2.6.3 The use of CPGs to support early mobilisation decision-making

Despite the inconclusive and inconsistent evidence base, and prior to the publication of AVERT, many CPGs recommended patients to be mobilised as early and as frequently as possible.[194-197] A review of 30 acute stroke guidelines worldwide found 22 (73%) contained EM recommendations with most recommending EM within 24 hours of stroke onset, and only 8 (36%) defined the intervention.[8] In 13 CPGs, the primary justification of these recommendations was to prevent immobility-related complications rather than to promote recovery. In this review, the authors highlighted the variation in the definition of EM (e.g. standing and sitting out of bed, walking, actively engaging in activities of daily living) and the timing of EM (after stabilization, within 24 hours, within 72 hours). In Table 2.2, CPGs and recommendations related to EM prior to the publication of AVERT in 2015 from Bayley et al are summarised.

Table 2.2. Early mobilisation clinical practice guidelines published *prior* to AVERT

Developers	Recommendations
UK, National Institute of Clinical Excellence[198]	Early mobilisation and optimum positioning of people with acute stroke. Early mobilisation is considered a key element of acute stroke care. Sitting up will help to maintain oxygen saturation and reduce the likelihood of hypostatic pneumonia. People with acute stroke should be mobilized as soon as possible (when their clinical condition permits) as part of an active management program in a specialist stroke unit. People with acute stroke should be helped to sit up as soon as possible (when their clinical condition permits).
Norway, Behandling og rehabilitering ved hjerneslag - National guidelines for treatment and rehabilitation by stroke[199]	All stroke patients should quickly be mobilized out of bed, and all of the interdisciplinary team B (level 2a) should contribute to mobilize patients as early and as often as possible. It has been common clinical practice to await mobilisation when systolic BP > 220 mmHg or severe dizziness. In these cases, it is recommended to measure BP lying and sitting, and to cancel mobilisation for SBP-decrease of >30 mmHg and concurrent symptoms. Patients treated with beta-blockers and/or who have hypovolemia should be monitored especially carefully. Mobilisation consists of the patient returns to a sitting, standing, or walking position, depending on the functional level. All members of the multidisciplinary team will help the patient be mobilized as early and as often as possible. Mobilisation must be conducted 00:10 day 7 days a week (175; 223; 224) (Level 2a)
Scotland, SIGN[200]	Stroke patients should be mobilized as early as possible after stroke (Level B).
European Stroke Organisation[201]	1. Early mobilisation of the patient according to out of bed within 24-h principle as soon as the stroke has stabilized, and the general medical condition and stroke severity make this possible. 2. Early mobilisation and physical therapy at least once daily (7/7) and, if resources make it possible, preferably twice a day, adapted to the medical condition of the patient.
World Stroke Organisation[202,203]	All patients admitted to hospital with acute stroke should start to be mobilized early (between 24 and 48 h of stroke onset) if there are no contraindications Contraindications to early mobilisation include, but are not restricted to, patients who have had an arterial puncture for an interventional procedure, unstable medical conditions, low oxygen saturation, and lower limb fracture or injury.
Nursing Research Institute Stroke Coalition[204]	Early mobilisation: Within the first 24 h for neurologically and hemodynamically stable patients is safe and feasible (class IIa: level of evidence B). Early mobilisation (within 52h) is associated with fewer complications. Patients with stable neurological and hemodynamic presentation can be mobilized to out of bed chair sitting even if level of consciousness is depressed i.e. stupor, obtundation, and lethargy (class IIa: level of evidence C)

Source: Adapted from Bayley et al, 2017[10]

Current CPGs predominantly reflect the findings from AVERT.[10] In Table 2.3, CPGs and recommendations related to EM *after* to the publication of AVERT from Bayley et al are summarised with the addition of the Australian CPGs[10] published in 2017. The consistent proposal is to recommend against [intensive] out-of-bed activity starting within 24 hours post-stroke,[10] with significant uncertainties about best practice care beyond the first episode of mobilisation. Some exceptions do apply e.g. the Heart and Stroke Foundation CPG states that mobilisation may be reasonable for some patients with acute stroke in the very early timeframe and NICE recommends the use of clinical judgement for those requiring little to no assistance.[205,206] It appears that clinical decisions regarding the dose of therapy and whom this therapy should be given to is complex, multifactorial and not sufficiently addressed in current CPGs. As previously discussed, clinicians are frequently faced with difficult clinical decisions and there is often uncertainty in the value of different options, and consequently clinical practice can vary. Decision-making is further complicated by the uncertainty in using population-level research evidence to inform individualised decision-making particularly when subgroup-effects do not exist to inform the applicability of that data to certain patients. *It is important to understand the factors which influence the decisions of clinicians so that clinical guidelines reflect decision-making in practice*. There is an immense need to formulate well-defined acute EM CPGs that can are clinically applicable and implementable to support clinical practice.

Table 2.3. Early mobilisation clinical practice guidelines published *after* AVERT

Developers	Recommendations
Canada, Heart and Stroke Foundation [206].	All patients with stroke should receive rehabilitation therapy as early as possible once they are determined to be rehabilitation ready and they are medically able to participate in active rehabilitation (Level A) within an active and complex stimulating environment (Level C). Frequent, out-of-bed activity in the very early time frame (within 24 h of stroke onset) is not recommended (Evidence Level B). Mobilisation may be reasonable for some patients with acute stroke in the very early time frame and clinical judgment should be used (Evidence Level C). All patients admitted to hospital with acute stroke should start to be mobilized early (between 24 and 48 h of stroke onset) if there are no contraindications (Evidence Level B). Contraindications to early mobilisation include, but are not restricted to, patients who have had an arterial puncture for an interventional procedure, unstable medical conditions, low oxygen saturation, and lower limb fracture or injury.
US, American Heart Association/ American Stroke Association [207]	It is recommended that stroke survivors receive rehabilitation at intensity commensurate with anticipated benefit and tolerance. (Level B) High-dose, very early mobilisation within 24 h of stroke onset can reduce the odds of a favorable outcome at 3 months and is not recommended (Level A)
UK, Royal College of Physicians Intercollegiate Stroke Working Party/ National Institute for Health and Care Excellence [205].	Patients with difficulty moving after stroke should be assessed as soon as possible within the first 24 h of onset by an appropriately trained healthcare professional to determine the most appropriate and safe methods of transfer and mobilisation. Patients with difficulty moving early after stroke who are medically stable should be offered frequent, short daily mobilizations (sitting out of bed, standing, or walking) by appropriately trained staff with access to appropriate equipment, typically beginning between 24 and 48 h of stroke onset. Mobilisation within 24 h of onset should only be for patients who require little or no assistance to mobilize.
Australia, National Stroke Foundation [208]	For stroke patients, starting intensive out-of-bed activities within 24 hours of stroke onset is not recommended. For patient with mild and moderate stroke, frequent, short session of out-of-bed activity should be provided, but optimal timing within the 48-hour post-stroke time period is unclear.

Source: Adapted from Bayley et al, 2017.[10]

2.6.4 Summary of challenges in clinical decision support for EM

In the preceding sections, the challenges clinicians often face in evidence-based decision-making have been outlined. The need for decision support stems from the significant challenges faced by stroke clinicians when translating population-level evidence from recovery and rehabilitation trials to individualised clinical decision-making. The inherent heterogeneity of stroke patient population and their recovery patterns are not explicitly covered by CPG recommendations. Stroke-specific factors such as stroke type, location and severity, and clinical factors such as age, sex and co-morbidities vary widely in the stroke population and contribute to variable recovery patterns. This adds to the complexity and ambiguity of decision-making in early mobilisation practices and potential inconsistencies in the delivery of care. Rarely does clinical trial evidence provide all the answers to give clinicians the information they need to confidently deliver treatment. In the case of EM, signal of harm in response to intervention reported in AVERT, and concerns over potential harms without mobilisation being undertaken during acute care, make the EM field particularly challenging for clinicians right now. At the most simplistic level, in the face of uncertainty, clinicians could choose to do nothing (leave the patients resting in bed for days or weeks) or something (mobilising the patient at a time, dose and intensity of their choosing). In AVERT, time to first mobilisation in the usual care group steadily decreased; around 60% of patients eventually started out-of-bed therapy within 24 h of stroke onset. The authors speculated whether it was a consequence of contamination, or a response to changes in attitudes to EM over time, response to the recommendations in clinical guidelines at the time, or both.[9] Nonetheless, understanding how clinicians internally and externally negotiate risks and benefits for different patients can provide important insights into how to optimise

EM decision-making for clinicians and better support risk communication with patients. It is evident that clear and clinically applicable CPGs for early mobilisation that balance the risks of harms vs. benefit are required to effectively support decision-making at an individual patient level, standardise care, and ultimately optimize patient outcomes. Overall, the use of population-level evidence to inform individualised decision-making is complicated by the multifaceted and complex nature of EM evidence and interventions, CPG development and reporting, and the characteristics of stroke patients and their outcomes. These critical research areas are the major focus of this thesis.

2.7 Summary and research gaps in the literature

As articulated in Section 1.2, the overall objective of this thesis was to investigate the barriers and enablers to the effective use of population-level to inform individual patient-level clinical decision-making for EM post-stroke. Based on the review of the literature provided earlier in this chapter, eight key gaps within four distinct research themes emerged (Figure 2.2):

Clinical practice guidelines

The variability and ambiguity of EM evidence and the heterogeneity of stroke patients requires complex decision-making. Clinically credible and applicable CPGs are needed to support and optimise clinical decisions.

Key gap 1: since the publication of AVERT, the utility and limitations of current CPGs to support individualized clinical decision-making have not been investigated.

Key gap 2: the key areas of improvement for the reporting of CPGs to better support EM decision-making are not known.

Early mobilisation evidence

Findings from EM RCTs have been largely inconclusive and inconsistent, with some trials showing benefit, while the single, largest RCT (AVERT) showed harm. The recent Cochrane review found no difference between EM and usual care despite the inclusion of AVERT and could not shed light on these uncertainties, yet the practice of EM continues. There appears to be differences in the response-patterns of patients to EM based on subgroup and dose analyses, but many are not statistically significant.

Key gap 3: most RCTs are not powered to look at subgroup-treatment effects to inform applicability of evidence to subgroups or individual patients.

Key gap 4: conventional meta-analysis methods cannot unpack the granularity of RCT data at an individual patient-level to investigate subgroup-treatment and dose effects of EM.

Clinical decision-making

Prior to AVERT, several factors such as stroke type, medical stability and administration of thrombolysis influenced EM clinical decision-making.

Key gap 5: since the publication of AVERT, the clinical and non-clinical factors that influence EM decision-making are not known.

Key gap 6: the way clinicians tailor their decisions based on their own interpretation of potential risks vs. benefits of EM to individual patients has not been investigated.

Patient-centered outcomes

The principles of EBM and CPGs advocate for the use of patient-centred outcomes. In stroke research, the UW-mRS is a newly proposed measure to improve the patient-centric and ordinal nature of the mRS.

Key gap 7: the construct validity of the UW-mRS as a patient-centred outcome measure of post-stroke burden has not been investigated.

Key gap 8: the ability of the mRS to capture important patient-centred outcomes of post-stroke burden is not known.

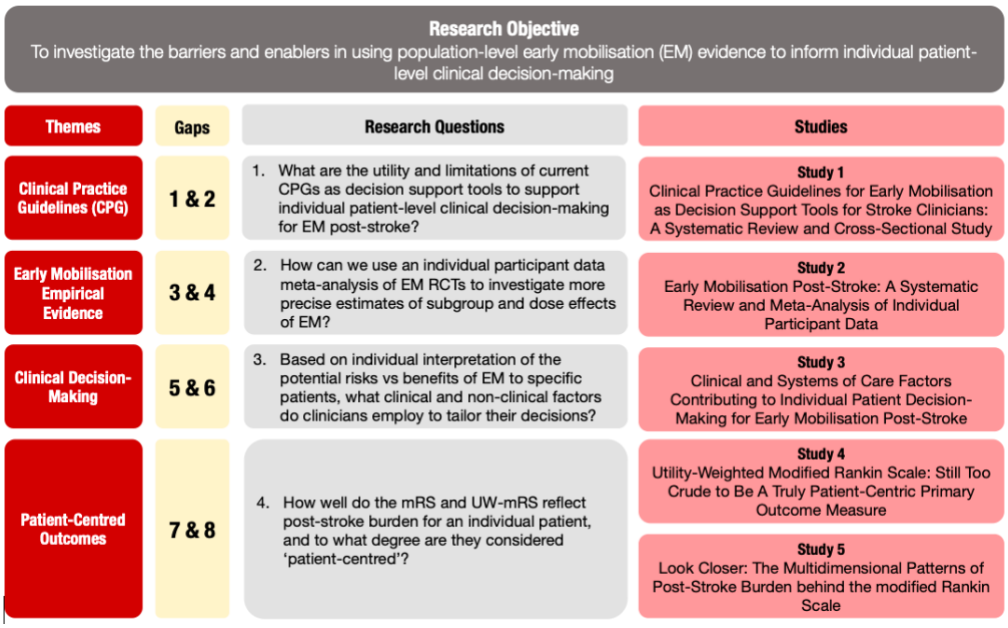


Figure 2.2. Thesis framework with key gaps in the literature. Four primary research themes with corresponding key gaps, research questions and related studies are depicted in this framework.

This concludes the introductory part of this thesis. In these first two introductory chapters, the necessary clinical and methodological background as well as the key gaps in knowledge were provided to form the foundation for the remainder of this thesis. The following sections (Chapters 3 to 7) constitute the novel research work of this thesis.

Chapter 3

Study 1: The Utility and Limitations of Clinical Practice Guidelines to Support Early Mobilisation Decision-Making After Stroke

3.1. Chapter overview

In this chapter, the utility of current CPGs as decision support tools to facilitate the use of population-level evidence to support individual patient-level decision-making for EM post-stroke was investigated (Research question 1, Figure 1.1). The *rationale* for Study 1 stems from the variability of EM evidence, and the heterogeneity of stroke patient population identified in the literature (Chapter 2), which requires complex clinical decision-making. Therefore, an important research area for this thesis was to investigate the utility and limitations for clinically credible and applicable CPGs to support decision-making in this challenging EM context.

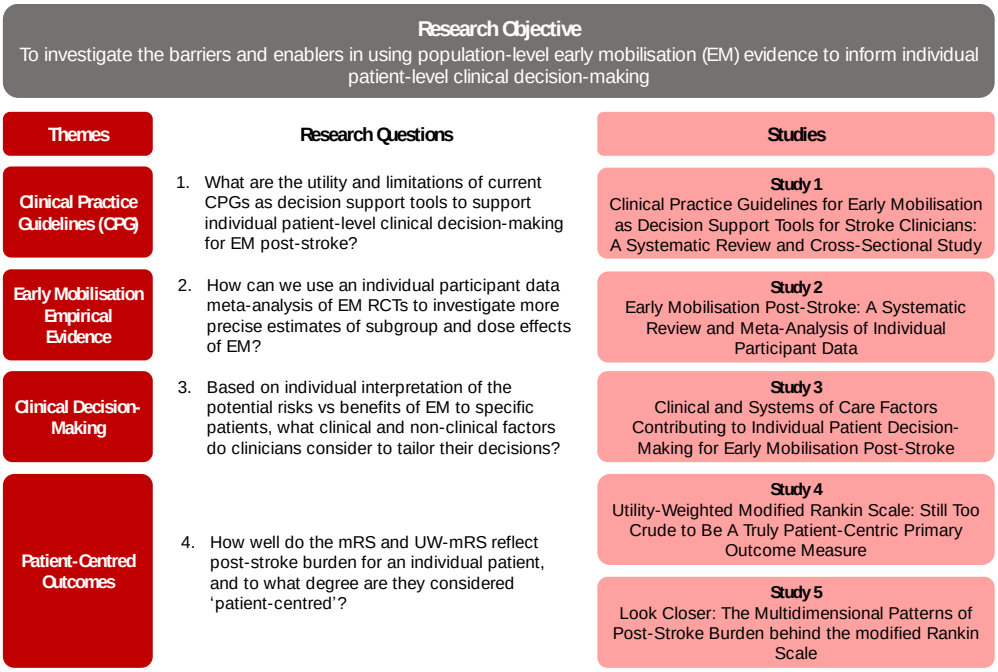


Figure 1.1. Thesis Framework.

Study 1 investigated how well CPGs for early mobilisation after stroke support individual patient-level decision-making for expert stroke clinicians. In this cross-sectional study, the utility and limitations of CPGs for EM post-stroke as decision support tools for expert stroke clinicians was explored using a two-pronged approach: review of EM CPGs and interviews with expert stroke clinicians. This cross-sectional study in presented in accordance with Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.[209] Ethics approval was obtained from The University of Melbourne Human Research Ethics Committee (1851680.1) and can be found in Appendix C. The chapter is presented in the *original submitted version* of the manuscript. This work is followed by Section 3.6 outlining the contribution of the chapter to the thesis objective.

This chapter has been accepted for publication to Frontiers in Neurology.

3.2. Introduction

Clinical decision-making is thought to be a contextual, continuous, and evolving process, where data are gathered, evaluated, and interpreted to select an *evidence-based* choice of action.[4] The definition emphasises the importance of both *evidence generation* and *evaluation*, and the *interpretation* that is relevant for individual patient decision-making. Clinical decisions are embedded in evidence-based medicine, which integrates clinically relevant (patient-oriented) research, clinical expertise and patient values and preferences.[2,3] However, evidence generated by RCTs, often expressed as an average treatment effect at a population level, cannot directly support clinical decision-making at an individual patient-level.[5,6] CPGs are decision support tools that serve to fill this gap by providing recommendations, generally with consideration of clinical applicability to the patient and setting, and by explicitly stating the options for, and implication of different care options, i.e. trade-off between harm vs. benefits, to support clinical judgements.[7,67,88]

In the domain of stroke rehabilitation and recovery, early mobilisation (sitting out of bed, standing or walking early after stroke) has long been considered an important part of stroke unit care.[162] CPGs have historically reflected this, and contained recommendations to mobilise patients as early as possible post-stroke.[8,10] However, a publication in 2015 of the largest RCT, A Very Early Rehabilitation Trial (AVERT; n=2104), demonstrated poorer outcomes in the early mobilisation group compared to usual care.[9] This resulted in many CPG recommending against [intensive] out-of-bed activity starting within 24 hours post-stroke,[10] with significant uncertainties about best practice care remaining. Furthermore, the recent publication of a Cochrane Review[14] and supplementary individual participant data meta-analysis[210] of early

mobilisation RCTs demonstrated that although early mobilisation is not recommended within 24 hours post-stroke, there is still a need for more detailed research to understand the optimal timing, frequency, and intensity of the intervention. This has an obvious implication for developing decision support tools.

The need for decision support stems from the significant challenges faced by stroke clinicians when translating population-level evidence from recovery and rehabilitation trials to individualised clinical decision-making. The inherent heterogeneity of stroke patients and their recovery patterns are not explicitly covered by CPG recommendations. This contributes to the complexity and ambiguity of decision-making in early mobilisation practices and potential inconsistencies in the delivery of care. It is evident that clear and clinically applicable CPGs for early mobilisation are required to effectively support decision-making at an individual patient level, standardise care, and ultimately optimize patient outcomes. The existing gaps in knowledge include the utility of current early mobilisation CPGs to support clinical decision-making, the areas of uncertainty in decision-making, and the key recommendations to improve future early mobilisation CPGs.

Our overall aim was to investigate how well CPGs for early mobilisation after stroke support individual patient-level decision-making for expert stroke clinicians. To meet this aim we:

(1) Systematically evaluated how well specific normative decision-support criteria from the Appraisal of Guidelines Research and Evaluation–Recommendations Excellence (AGREE-REX) tool are met by early mobilisation CPGs.

(2) Empirically investigated the influence of early mobilisation CPGs on decision-making of Australian stroke clinicians', the areas of uncertainty in their decision-making, and suggestions to improve CPG recommendations.

3.3. Methods

3.3.1. Study Design

This cross-sectional study was designed using a conceptual framework developed in accordance with the definition of clinical decision-making[4] to emphasise the importance of both *evidence generation and evaluation*, as well as the *interpretation* component relevant for *individual patient decision-making* (Figure 3.1 top panel). For that reason, this study includes two components:

1. Normative component – a review of stroke CPGs containing recommendations for early mobilisation practices. Recommendations were appraised using purposely selected AGREE-REX items that are relevant to decision-making for clinicians.
2. Empirical component – semi-structured interviews with Australian expert stroke clinicians to investigate the influence of early mobilisation CPGs on decision-making of stroke clinicians', the areas of uncertainty in their decision-making, and suggestions to improve CPG recommendations.

We report this cross-sectional study in accordance with Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.[209]

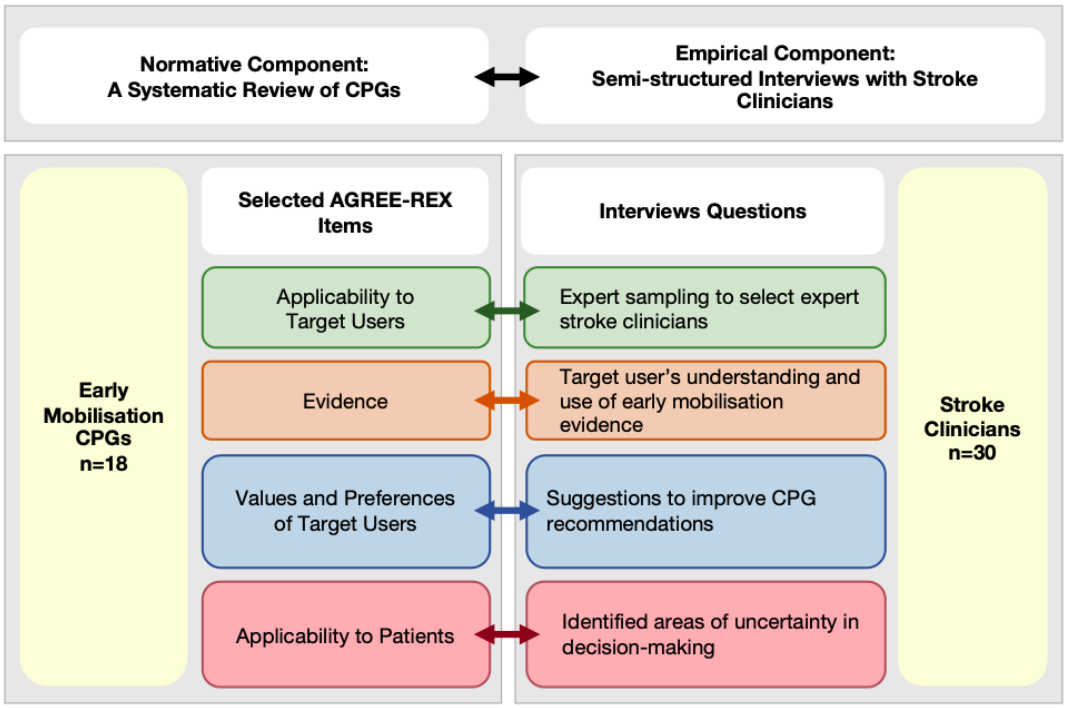


Figure 3.1. Conceptual framework integrating normative and empirical components. AGREE-REX, Appraisal of Guidelines Research and Evaluation–Recommendations Excellence; CPG, clinical practice guidelines; EM, early mobilisation.

Normative Component

3.3.2. Search Strategy and Data Extraction

In November 2019, the following resources were used to search for CPGs:

- Stroke CPGs identified by the World Stroke Organisation[211]
- 30 early mobilisation CPGs identified by Bernhardt et al.[162]
- LMICs stroke CPGs identified by the Lancet Neurology Commission[212]
- US National Guidelines Clearinghouse
- Guidelines International Network
- Scientific Electronic Library Online (SciELO) Network
- PubMed and Excerpta Medica Database (EMBASE)
- Websites of known CPG development bodies

Search terms included Topic = “country name” AND Topic = “guideline” OR “consensus” OR “standards” OR “recommendations” AND Topic = “stroke OR cerebrovascular disorder/disease OR intracranial haemorrhage OR cerebrovascular accident OR “early mobilisation”. CPGs were selected if they were publicly available; currently active; produced with the support of a health professional association or society, public or private organisation, health care organisation or plan, or government agency; published after 2015 to ensure data from AVERT[9] was considered; and contained recommendations about early mobilisation practices. CPGs for the management of childhood stroke were excluded. CPGs in Dutch, German, Portuguese, Russian and Ukrainian were also independently reviewed by an individual fluent in those languages. Other non-English CPGs were examined using Google Translate, which has been shown to be a viable and accurate tool for translating non–English-language trials to conduct systematic reviews.[213,214] Each CPG was reviewed and information regarding the title, year of publication, name of organisation/society/government agency, specific early mobilisation recommendations, evidence informing recommendations, and methodological approach was extracted for appraisal.

3.3.3. Appraisal of Clinical Practice Guidelines

The AGREE-REX (Appraisal of Guideline Research and Evaluation-Recommendations Excellence) tool was used to evaluate the quality of CPG recommendations for early mobilisation post-stroke.[89] The tool consists of nine items structured within three theoretical domains: clinical applicability, values and preferences, and implementability. All nine AGREE-REX items contain an

operational definition of the item and several item-specific criteria (ranging between 2 to 10) for CPGs to be evaluated on.

Since the overall aim of this study was to understand the level of decision support provided in CPGs for clinicians, the four most relevant AGREE-REX items were selected for detailed review: *Evidence* (8 criteria), *Applicability to Target Users* (5 criteria), *Applicability to Patients* (4 criteria), and *Values and Preferences of Target Users* (3 criteria). The *Evidence* and *Applicability to Patient* items of the AGREE-REX relate to the tailoring of clinical decisions to individual and/or patient subgroups. The *Applicability to Target Users* and *Values and Preferences of the Target Users* items relate to providing decision support to the stroke clinicians.

The excluded five AGREE-REX items included: *Values and Preferences of Policy/Decisionmakers*, *Values and Preferences of Guideline Developers*, *Purpose*, *Local Application and Adoption*, and *Values and Preferences of Patients/Population*. These items were excluded because we focused on investigating *how well CPGs support decision-making for stroke clinicians (target user).* Although the *Values and Preferences of Policy/Decision-makers* and *Guideline Developers* items are certainly important in developing clinically credible and implementable guidelines, these are not directly relevant to individual patient-level decision-making for stroke clinicians about early mobilisation. The *Purpose* item was excluded because it was related to "the implementation goals of the guideline (e.g., for advocacy, policy change, etc.)".[89] The *Local Application and Adoption* item was excluded not to diminish the role of systems of care factors and social structures, but rather to emphasise that the *four chosen items* will be present and play a crucial role in supporting individual patient-level decision-making across different healthcare systems and social structures. The *Values and Preferences of Patients* item was excluded because we specifically

focused on clinical decision-making by clinicians (target user of CPGs) rather than shared decision-making. Certainly, a future study following a similar integrated research design (i.e., collecting both normative and empirical evidence *from patients*) presents a viable future direction and could provide valuable insights into shared decision-making.

Every CPG was independently assessed against the criteria (yes/no) by two reviewers (VR, KH). The AGREE-REX items were given a score (percentage of total achievable score) depending on the number of criteria addressed by the CPG i.e., if a CPG addressed 6 out of 8 criteria for the *Evidence* item, the CPG would receive a score of 75% for that item. The following outcomes were reported: (1) a summary of AGREE-REX item scores (%) achieved by each CPG and (2) the percentage of CPGs addressing each criterion.

Empirical Component

3.3.4. Study Participants and Recruitment

In this study, expert sampling[215] was used to recruit a representative sample of CPG target users and post-stroke early mobilisation content experts (i.e., participated or conducted early mobilisation research) who can provide a level of expertise necessary to understand the evidence and applicability to patient components of this study. This sample typically forms part of a guideline development group in Australia.[57] As a result, we identified a sample of stroke physicians, physiotherapists and nurses/nurse practitioners with a high level of expertise in delivering acute stroke care and early mobilisation practices (senior position, >6 years stroke experience). According to this sampling strategy, clinicians identified as appropriate for participation were invited to participate via email and were recruited between June 2019 to March 2020. Ethics

approval was obtained from The University of Melbourne Human Research Ethics Committee (1851680.1). Consent was obtained from all participants involved in the study.

3.3.5. Semi-Structured Interview

We conducted the interviews in two distinct parts. Each addressed an individual aim: (1) to investigate the utility and limitations of current early mobilisation CPGs and (2) identify the decision-making factors and specific high-interest patient subgroups using a data visualisation tool. The analysis and results for the second part are presented elsewhere.

In this study, the design of the CPG-related interview questions was informed by the conceptual framework that links the normative and empirical components of the study (Figure 3.1 bottom panel). Therefore, the topics of interest were pre-determined using selected AGREE-REX items to investigate both the evidence generation and evaluation component and the interpretation component relevant for individual patient decision-making. The semi-structured interview questions were designed to elicit the desired information (Appendix C). During the interview, the following topics of interest were examined: the influence of early mobilisation evidence and CPGs on decision-making, areas of uncertainty in decision-making, and additional information and parameters required in early mobilisation CPG recommendations to support decision-making. Participants were encouraged to provide more detailed responses using probing techniques and prompts. All participants were interviewed either in-person or via videoconference. Demographic data on participants were collected using a brief questionnaire.

3.3.6. Coding and analysis

The interviews were audio-recorded and transcribed verbatim using a paid transcription program (Amberscript) to facilitate the analysis of data. A *deductive* and *directed thematic content approach*[216] was used based on the main outcomes for the study: clinicians' use of and view on CPGs, specific areas of uncertainty in early mobilisation decision-making, and recommendations for improvement of early mobilisation CPGs. VR read all transcripts to make initial analytical observations about the data. Transcripts were imported into QSR International NVivo 9 to organize and analyse data using the set of a priori and predefined interview questions. As per the directed thematic content approach, predetermined codes were formed based on the three desired outcomes.[217] The strategy for coding involved reading through each transcript and conducting line-by-line coding using the predetermined codes. Subcodes were determined during this process with subsequent analysis. The data that could not be initially coded are identified and analysed later to determine if they represent a new category or a subcategory of an existing code.

Cross-checking occurred in two ways: firstly, KH double-coded 25% of all transcript across all three coding themes. Differences which occurred were resolved by consensus. Ongoing discussions established trustworthiness and credibility to clarify the interpretation of the data. All subcodes were discussed between VR and KH to determine overlap or divergence in sub-themes within the three broad outcomes. Secondly, KH and VR also randomly selected codes in NVivo to ensure the quotes from the transcripts accurately reflected the theme it was coded to. Descriptive

statistics with frequencies and proportions were produced and reported using Stata (version 14.2; StataCorp LP, College Station, TX, USA).

3.4. Results

Normative Component

3.4.1. CPG Selection and Data Extraction

The initial search yielded 57 CPGs. We excluded 39; 31 were published predating the completion of the largest trial on this topic (AVERT); 8 did not include EM recommendations (Supplementary Figure C1 & Table C1, Appendix C). The remaining 18 CPGs were from the following geographical regions: Argentina, Australia, Canada, China, Europe, Finland, India, Italy, Korea, Netherlands, Norway, Peru, United Kingdom and United States of America (Supplementary Table C2, Appendix C).[218-220]

3.4.2. AGREE-REX Appraisal

In most CPGs, *Applicability to the Target User* and *Values and Preferences of the Target User items* were met (Table 3.1). However, the *Applicability to Target Users* criterion "the guideline differentiates between recommended actions for which clinical flexibility and individual patient tailoring is more appropriate in the decision-making process and those for which it is less appropriate" was only addressed in 61% of CPGs (Figure 3.2). *Applicability to Patients* was the item with the least number of criteria met by CPGs (Table 3.1). Within the *Evidence* item, more than 78% of CPGs addressed the magnitude of benefits vs. harms, risk of bias of included studies, and the consistency of results (Figure 3.2). However, less than 61% sufficiently addressed the directness of the evidence, confounding factors, publication bias, and dose-

response gradient in their recommendations. The *Applicability to Patients* criterion relating to "the tailoring of recommendations to individual patients" was only addressed by 61% of CPGs, while the criteria relating to patient-centred outcomes were addressed by only 17% of CPGs (Figure 3.2).

Empirical Component

3.4.3. Study Participants

Thirty expert stroke clinicians from Australia (Victoria and South Australia participated in this study; 11 physicians, 11 physiotherapists and 8 nurses (Table 3.2). No clinicians approached declined to participate. On average, clinicians practised in a stroke context for 14 years (IQR 10-25), 87% are currently practising, and 74% work in a metropolitan acute stroke care unit.

Table 3.1. Summary of AGREE-REX item scores

	Evidence	Applicability to Target Users	Applicability to Patients	Values and Preferences of Target Users
Argentina, 2019[221]	25%	100%	0%	75%
Australia, 2019[220]	100%	100%	100%	100%
Canada, 2018[222]	88%	100%	33%	100%
Canada, 2019[223]	88%	100%	33%	100%
China, 2019[219]	50%	100%	33%	100%
Europe, 2018[224]	50%	100%	0%	75%
Finland, 2016[225]	50%	100%	0%	75%
India, 2019[226]	13%	80%	33%	75%
Italy, 2016[227]	88%	100%	33%	100%
Korea, 2017[228]	50%	100%	0%	75%
Netherlands, 2019[229]	100%	100%	100%	100%
Norway, 2017[230]	100%	100%	100%	100%
Peru, 2020[231]	75%	100%	33%	100%
UK, 2016[205]	63%	100%	33%	100%
UK, 2019[232]	88%	100%	33%	100%
USA, 2019[233]	63%	100%	0%	75%
USA, 2016[234]	63%	100%	0%	75%
USA, 2019[235]	63%	100%	0%	75%
All CPGs Median (IQR)	**63% (50%-88%)**	**100% (100%-100%)**	**33% (0%-33%)**	**100% (75%-100%)**

The proportion of criteria (%) achieved by each CPG and median (IQR) summary scores across all CPGs. Shading emphasises poor (dark) to good (light) meeting of criteria for each item. CPG, clinical practice guidelines; IQR, interquartile range; UK, United Kingdom; USA, United States of America.

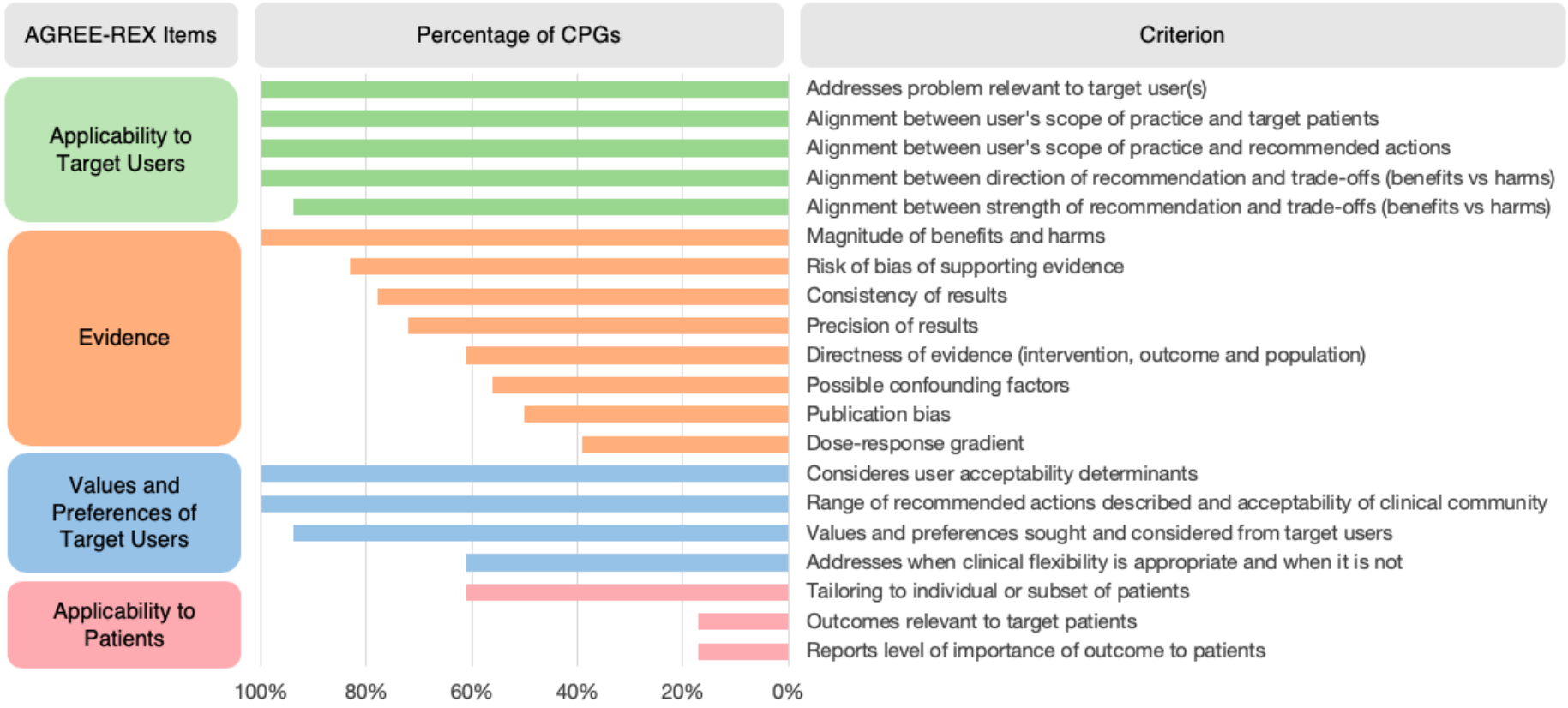


Figure 3.2. Proportion of CPGs addressing AGREE-REX criteria. Summary of percentage of CPGs addressing each AGREE-REX criterion for the four items.

Table 3.2. Demographic characteristics of expert stroke clinicians

Variables	No. (%)
Occupation	
Physician	11 (37%)
Physiotherapist	11 (37%)
Nurse	8 (26%)
Highest level of education	
Ph.D.	14 (47%)
Clinical Doctorate	1 (3%)
Masters (clinical)	5 (17%)
Bachelor's Degree	7 (23%)
Graduate Diploma	3 (10%)
Currently practising	
Yes	26 (87%)
No	4 (13%)
Primary stroke environment	
Acute stroke unit	22 (74%)
Inpatient rehabilitation	3 (10%)
Outpatient rehabilitation	1 (3%)
Research institute	4 (13%)
Level of knowledge on early mobilisation	
High, well informed about evidence	15 (50%)
Average, up to date with evidence	15 (50%)
Low, not up to date with evidence	0 (0%)
Number of years practising in a stroke context, median (IQR)	14 (10-25)

IQR, interquartile range.

3.4.4. Use of CPGs as Decision-Support Tools

Out of the total sample, 50% of clinicians perceived their current level of knowledge on early mobilisation evidence to be high (Table 3.2). Only 43% of clinicians indicated that CPGs provide decision-support, 40% expressed their preference to use clinical reasoning with an understanding of current evidence over the use of CPGs. Forty-seven percent of clinicians found current CPGs "too broad or vague", while 10% thought it was only "useful for junior or non-specialised staff".

3.4.5. Areas of Uncertainty in Decision-Making

The areas of uncertainty were generally related to the process of clinical reasoning, the intervention dose, tailoring to specific patient and stroke characteristics and the impact of early mobilisation practices on recovery trajectories of patient's post-stroke (Figure 3.3). Specifically, the three most frequently mentioned areas of uncertainty were related to optimal intervention dose (27%), tailoring early mobilisation practices for people with moderate and severe stroke (27%), and the lack of knowledge of potential responders and non-responders to early mobilisation practices (23%).

3.4.6. Areas of Improvement for Early Mobilisation CPGs

In response to the areas of uncertainty, the clinicians expressed the need for CPGs to report more granular descriptions of patient and stroke characteristics to allow for appropriate tailoring of decision-making to an individual or subsets of patients. Other recommendations included clearer statements about when clinical flexibility is appropriate and when it is not, physical assessment criteria to aid decision-making, more detailed descriptions about safety parameters to consider, and a detailed description of the intervention dose. The clinicians also provided specific suggestions for parameters they would like to be reported in early mobilisation recommendations (Figure 3.3)

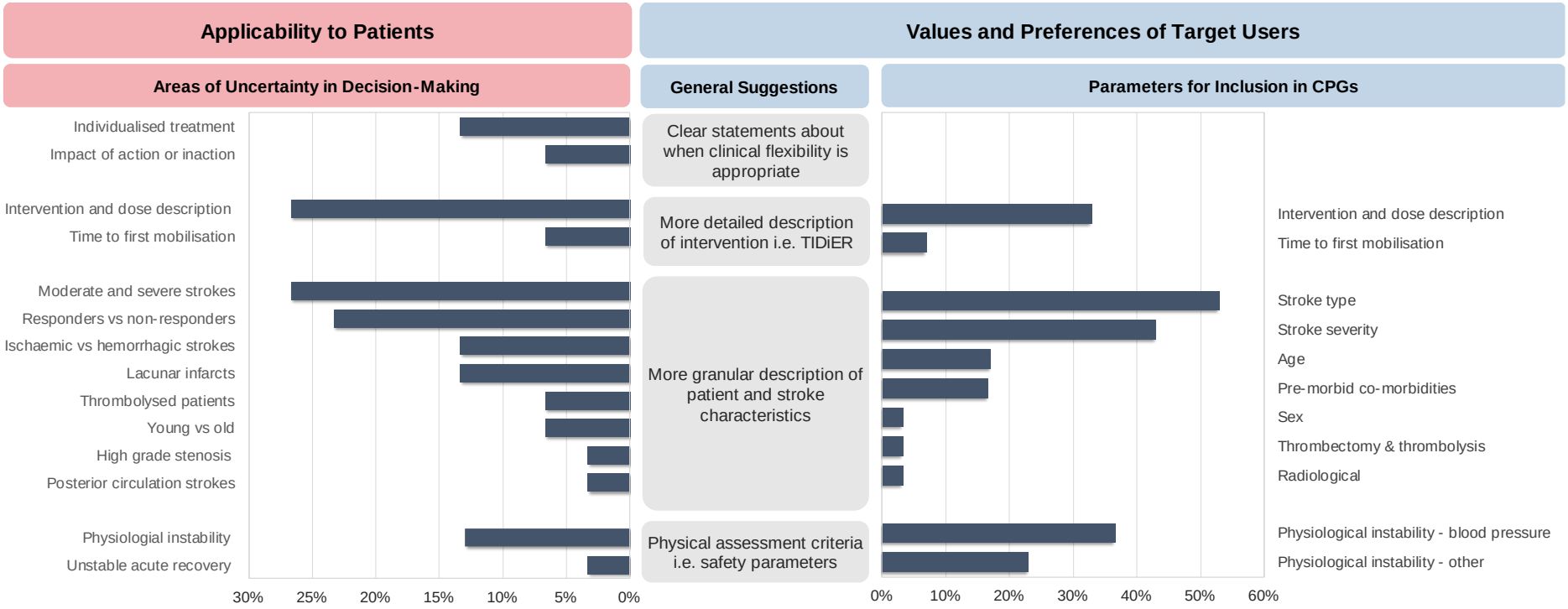


Figure 3.3. The areas of uncertainty and key recommendations for future CPGs. TIDieR, Template for Intervention Description and Replication.

3.5. Discussion

In this study, we used a novel two-pronged approach to investigate how well current CPGs to support early mobilisation decision-making post-stroke for Australian stroke clinicians. The review of early mobilisation CPGs demonstrated that almost all CPGs addressed the *Applicability to the Target User* and *Values and Preferences of the Target User* items however, the *Evidence* and *Applicability to Patients* items were not met to the same degree. Therefore, it is not surprising that only 43% of interviewed expert stroke clinicians indicated that CPGs provide decision support, with many often relying on clinical reasoning and individual interpretation of the evidence to select an evidence-based choice of action. Many clinicians (47%) found current CPGs "too broad or vague". This is an important finding given that specific guidelines can change a physician's decision for the better, while non-specific guidelines can change it for the worse.[103] Another important point is that although many Australian stroke experts found Australian CPGs to be "too broad or vague", the included AGREE-REX item scores for the Australian CPG reached 100%. The discrepancy between our normative and emprical findings may be due to the limited evidence base to effectively support clinical decision-making. There is still a need for more detailed early mobilisation research to understand the optimal timing, frequency, and intensity of the intervention. The need for a comprehensive evidence base has obvious implications on how well CPGs can truly support early mobilisation decision-making for a clinician, which may be reflected in our empirical findings. Several areas of uncertainty in decision-making and related suggestions for the improvement of early mobilisation CPG recommendations were identified and discussed in more detail below.

The normative component of this study revealed several factors mediating and impeding the use of early mobilisation CPGs as decision-support tools. Concerning the *Evidence* item, most CPGs described the magnitude of benefits and harms of early mobilisation (mainly informed by AVERT), supporting clinical decision-making by explicitly stating the options for, and implication of, different options. However, many CPGs did not address the possibility of confounding factors or a dose-response gradient despite all CPGs utilising findings from AVERT, which included safety and intervention parameters. Interestingly, in many CPGs, the exact intervention, population, and outcomes of interest to the clinical problem were not addressed. This was reflected in the interviews with clinicians expressing uncertainty about delivering the optimal dose of early mobilisation (exact intervention) to specific subgroups of stroke patients (population) and the impact of early mobilisation on a patient's recovery trajectory (outcomes). As such, the findings may highlight potential gaps in the early mobilisation evidence base that require future exploration of existing data and the development of new clinical trials to better support evidence-based clinical decision-making.

The *Applicability to Patients* item and the identified areas of uncertainty in clinicians' decision-making exposed several insufficiencies of early mobilisation CPGs in providing decision-support. Specifically, many CPGs did not describe how decisions should be tailored to specific patient and stroke characteristics or include patient-centred outcomes. This is supported by the empirical findings from clinicians, which demonstrated uncertainty when making patient-centred decisions. In particular, the main areas of uncertainty included tailoring decisions to potential responders and non-responders, those with moderate and severe strokes, and different stroke types (ischaemic vs. haemorrhagic). By using a novel approach to integrate the newly

developed AGREE-REX tool [87,89] with empirical investigations, this study provides an in-depth understanding about why previous assessment of the methodological quality of CPGs for rehabilitation post-stroke using the AGREE II do not sufficiently address the "applicability" domain.[236] It has been suggested that barriers in the uptake of CPGs to support individualised decision-making may be due to the perceived rigidity of CPGs, insufficient clinical flexibility or loss of clinical autonomy.[36,37,101,104] However, as our normative and empirical data suggests, the absence of granular recommendations is perhaps the motivating reason early mobilisation CPGs are underutilised. This is an important barrier that needs to be acknowledged and addressed for the future development of CPGs to support clinicians effectively.

The identified areas of uncertainty directly led to four broad sets of recommendations by stroke clinicians. The recommendations included clearer statements about when clinical flexibility is appropriate, more granular descriptions of patient and stroke characteristics to allow for appropriate tailoring of decision-making, a more detailed description of the intervention dose, and the need for physical assessment criteria including safety parameters. The expert stroke clinicians also suggested specific parameters to be reported in future early mobilisation CPGs they deemed necessary when making clinical decisions at an individual patient level. These included consideration of safety parameters and intervention parameters that should be modified to people with different stroke types, stroke severity, age, and comorbidities.

Overall, the apparent need for clear and clinically applicable recommendations directly corresponds to the *Values and Preferences of Target Users* item. While many CPGs sought and considered the values and preferences of clinicians, the criterion requiring specification about when clinical flexibility is appropriate, and when it is

less appropriate, was not sufficiently addressed. Clinical flexibility allows “leeway for clinical judgment, patient preferences, and clinically relevant conditions of the delivery system (including necessary equipment and skilled personnel)”.[48] Indeed, this is a critical aspect of evidence-based medicine that involves balancing clinical autonomy and reasoning with best available evidence with patients' values and preferences. A decision-support tool needs to specify circumstances when clinical flexibility is necessary to ensure decisions are optimally tailored for specific patients, and when evidence-based standardised care is required. Clinical decisions about early mobilisation post-stroke are complex, multifactorial, and not sufficiently addressed in current CPGs.

This study has some limitations. The use of Google Translate without consultation of clinicians in the country may have been insufficient in contextualising those recommendations to the local context. While we established the study's scope to enable a comprehensive investigation of the clinical credibility and applicability of current EM CPG using AGREE-REX, we did not assess the methodological rigour and transparency of CPGs. It could be useful to investigate the internal validity of EM CPGs to understand how to improve the overall quality of current CPGs. Another limitation is that we did not include all AGREE-REX items. The unexplored items were related to the values and preferences of patients/populations, values and preferences of policy/decision-makers, purpose, and local application and adoption.[89] Consideration of patients and funders' values and perspectives, and alignment across different viewpoints, may enhance the utilisation of CPGs.[87,88,237] Future investigations could adopt our novel two-pronged approach to ensure these normative standards of AGREE-REX are linked with empirical investigation of the different stakeholders (patients, policy or decision makers etc.). All expert stroke clinicians

were from Australia, and therefore, our findings are naturally within the scope of an Australian healthcare system and may limit the generalisability of our findings. Nonetheless, the use of expert sampling[215] to obtain a multidisciplinary sample of clinical (physicians, physiotherapists, and nurses/nurse practitioners) and academic experts in acute stroke and early mobilisation practices allowed the identification of important insufficiencies in decision-support of current CPGs.

The identified lack of specificity, clinical applicability, and adaptability of current CPGs to effectively respond to the heterogeneous clinical stroke context has provided a clear direction for improvement. The four key recommendations for future early mobilisation CPGs include more granular descriptions of patient and stroke characteristics for appropriate tailoring of decisions to individual or subgroups of patients; clear statements about when clinical flexibility is appropriate; detailed description of the intervention dose; and physical assessment criteria including safety parameters.

3.6. Chapter contribution to the thesis objective

In this chapter, several limitations in current EM CPGs to support clinical decision-making were identified. Given the current uncertainties in the EM field, these gaps in the translation of evidence to decision-making provide a clear direction for improvement in the development of CPGs to better support evidence-based clinical decisions.

An interesting finding from this study was the variability in CPGs addressing the Evidence item of AGREE-REX. This may highlight potential gaps in the EM evidence base that require future exploration of existing data to better support evidence-based clinical decision-making. In the next chapter, research that aims to enrich the EM evidence base by generating patient-specific evidence is presented.

Chapter 4

Study 2: Early Mobilisation Post-Stroke: A Systematic Review and Meta-Analysis of Individual Participant Data

4.1. Chapter overview

From the literature review in Chapter 2, we know that findings from EM RCTs have been largely inconclusive and inconsistent, with some trials showing benefit, while the single, largest RCT (AVERT) showed harm. It is also known that the recent Cochrane review found no difference between EM and usual care despite the inclusion of AVERT and could not shed light on these uncertainties. Therefore, in this chapter, we unpacked the granularity of RCT data at an individual patient-level to investigate subgroup-treatment and dose effects of EM. In doing so, it was hoped that patient-specific evidence could be used to better inform individual patient-level decisions for EM post-stroke.

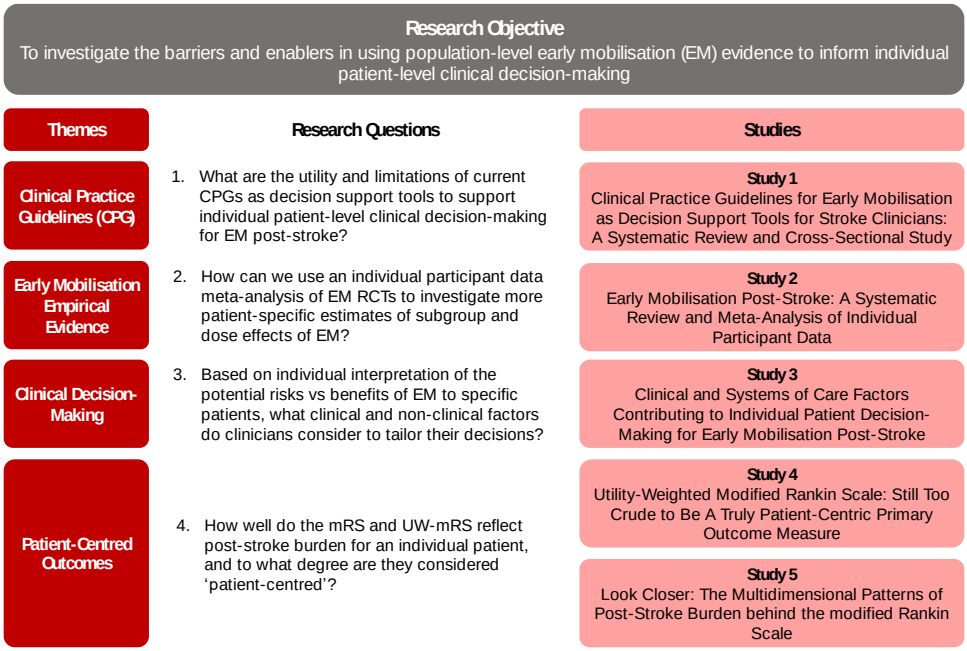


Figure 1.1. Thesis Framework.

This chapter contains a published study (Study 2), which aimed to use an individual participant data (IPD) meta-analysis of EM RCTs to investigate more precise estimates of subgroup and dose effects of EM (Research question 2, Figure 1.1). The chapter contains the *accepted version* of the manuscript (Sections 4.2 – 4.6) followed by Section 4.7 outlining the contribution of the chapter to the thesis objective.

Ethics approval was not sought for this IPD meta-analysis, on the premise that it addressed the same research question as the original studies for which participants gave informed consent. However, a Memorandum of Understanding (Appendix D) was signed by all trialists to verify that the use of de-identified data for the purposes of this study was approved by the respective institutional review boards. Peter Langhorne, Fanny Herisson, Simone Poletto, Yanna Tong and Julie Bernhardt were lead authors from the included trials. They were not involved in the trial selection and rating processes. Due to copyright restrictions, the Version of Record (PDF) of the manuscript was not included in the Appendices.

This chapter contains the Author's Original Manuscript of an article published by Taylor & Francis Group in Disability and Rehabilitation available online here.

4.2. Introduction

Early mobilisation (EM) post-stroke has been the subject of considerable research and included as a recommendation in several CPGs globally.[211,220,222,230,232,233] Yet, numerous trials have produced surprising or inconclusive findings that justify more in-depth exploration of the effects of EM post-stroke.[9,238-240] The most notable of all EM trials, A Very Early Rehabilitation Trial (AVERT), demonstrated an association between higher doses of EM (<24 hours post-stroke) and a reduction in the odds of a favourable functional outcome at 3 months.[9] As a consequence, many CPGs recommend EM within 48 hours, but no earlier than 24 hours.[211,220,222,230,232,233] However, it has been argued that these findings should not prolong immobility or inactivity of patients post-stroke.[11] While AVERT challenged conventional beliefs, there are still gaps in knowledge concerning the optimal dose of mobilisation with regards to duration, frequency and intensity, and how to identify responders and non-responders.

A recently published Cochrane Review found EM did not increase the proportion of patients who survived or had a good outcome post-stroke compared to usual care.[14] However, due to the aggregated nature of the meta-analysis, there was limited ability to include predictive factors i.e., stroke severity and age, leading to uncertainty in the estimations of effect, and an inability to explore subgroup–treatment effect interactions. Furthermore, elements of the intervention (i.e., time to first mobilisation, duration, frequency, intensity), patient characteristics, and outcomes are often inconsistently reported. Traditional methods of data aggregation do not effectively consider this complex and multifaceted nature of stroke rehabilitation trials and the heterogeneity of intervention reporting.

The incremental value of using individual patient data (IPD) when meta-analysing large trials includes a reduction in publication bias, accounting for missing data at an individual level, standardising statistical analyses across studies and adjusting for baseline factors consistently across studies.[106] The aims of this IPD meta-analysis were to determine the safety and efficacy of EM (<48 hours from stroke onset) compared to usual care, to investigate dose effects of EM and to assess the heterogeneity in trial design and reporting.

4.3. Methods

4.3.1. Search strategy and selection criteria

This meta-analysis was performed according to the Preferred Reporting Items for Systematic review and Meta-Analyses of IPD (PRISMA-IPD) guideline.[241] IPD was sought from RCTs comparing an intervention group that started out-of-bed mobilisation within 48 hours of stroke onset and aimed to reduce time-to-first mobilisation with or without an increase in the amount or frequency, to usual care (normal practice or no routine intervention) for acute stroke patients. All eligible trials included participants with a clinical diagnosis of stroke (focal neurological deficit of cerebrovascular origin) with or without confirmation on brain imaging, and no contra-indication to mobilisation within 48 hours of stroke onset. The trials were identified and sourced from the recently published Cochrane Systematic Review investigating very early vs. delayed mobilisation after stroke.[14] Full details of the inclusion and exclusion criteria, search strategy and quality appraisal have been published and extensively described.[14] The initial Cochrane Review searched the Cochrane Stroke Group Trials Register, 19 online databases and relevant ongoing trials and research registers, the Chinese medical database, Wanfangdata and

reference lists, and contacted researchers in the field (dated July 2017) yielding 21,395 studies. Nine eligible trials were identified (n=2958) and four trials were additionally identified as "ongoing" at the time of the initial search. Since the initial Cochrane Review search was completed in 2017, we additionally searched relevant completed trials using Cochrane Stroke Group Trials Register, US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov and World Health Organization (WHO) International Clinical Trials Registry Platform (searched June 2019) for the purposes of the IPD meta-analysis.

4.3.2. Data collection

The protocol for the meta-analysis was pre-specified and registered with PROSPERO (ID CRD42019132724 submitted 24/Apr/2019) before analysing the combined dataset. The corresponding authors from nine eligible trials were first contacted via email in July 2018 and invited to share deidentified, IPD for pre-specified variables (Table 4.1). In March 2019, the authors from the four "ongoing trials" identified in the Cochrane review were contacted to determine if their findings were published and invited to share data. A copy of the study protocol and a Memorandum of Understanding stipulating the terms of the data sharing agreement were provided. All corresponding authors were followed-up if no contact was made. Out of nine eligible trials, five agreed to share IPD for the purposes of this meta-analysis: AVERT II [238], AVERT III,[9] VERITAS,[239] VERIS-BRAZIL,[240] and SEVEL.[242] Four "ongoing" trials were identified and included in the Cochrane review. One trial (Tong et al) [243] had recently completed recruitment and published their findings when the selection process took place for the current IPD meta-analysis resulting in the inclusion of an additional dataset.

Table 4.1. Prespecified variables

Variables	
Treatment allocation	Time to first mobilisation
Age	Presence and type of adverse events
Sex	Death at 14 days & 3 months
Stroke type	Dependency at 3 months
Stroke risk factors	ADL score at 3 months
Baseline NIHSS	Quality of life at 3 months
Premorbid mRS	Number of out-of-bed sessions per day
Co-intervention	Number of minutes per session per day
Time to randomisation	Total intervention period (days)

IPD was extracted by direct access to the study databases. Two reviewers (VR and LC) checked for discrepancies in sample size and prespecified variables between the provided databases and the original published articles, noting missing data. If discrepancies were identified, clarification was sought from the corresponding author. Once all trial databases were cross-referenced and coding of each prespecified was standardised across the trials, they were merged into a common database. The master database was cross-referenced once more for data checking and correct coding with each original trial database by two reviewers (VR, LC). Risk of bias was assessed for all trials using the Cochrane Risk of Bias Tool by two reviewers (VR, KH). Quality of intervention development, monitoring and reporting was assessed by two reviewers (VR, KH) using a checklist comprising of the Stroke Recovery and Rehabilitation Roundtable (SRRR) intervention development recommendations[12] and Template for Intervention Description and Replication (TIDieR) checklist.[244]

4.3.3. Specification of outcome and dose measures

The modified Rankin Scale (mRS) is a frequently used functional outcome measure in acute stroke trials. The ordinal scale quantifies the degree of disability or dependence post-stroke ranging from 0 (no symptoms) to 6 (death).[245] Favourable outcome is commonly denoted as mRS 0-2 and poor outcome as mRS 3-6. Therefore,

the prespecified primary efficacy outcome of this meta-analysis was the proportion of patients with a favourable outcome (mRS 0-2) at 3 months post-stroke.

The primary safety outcome was the proportion of patients that died at 3 months post-stroke. Secondary outcome measures included an ordinal shift of the mRS score across the entire range of the scale, Barthel Index, death at 14 days, length of acute hospital stay at 3 months, and number and type of adverse events. The prespecified dose measures included time from stroke onset to first mobilisation (hours), number of out-of-bed sessions per day (frequency), minutes of out-of-bed activity per day (amount), and/or total intervention period (days).

4.3.4. Statistical analyses

All statistical analyses were performed using Stata (version 14.2; StataCorp LP, College Station, TX, USA). A one-stage approach to IPD meta-analysis using mixed-effect logistic regression models, adjusting for baseline age and NIHSS was performed for the primary efficacy, safety and secondary outcomes.[246,247] Mixed-effects modelling with random intercepts for individual trials and random slope for treatment effect across trials was utilised to account for between-trial variance. Adjusted odds ratios (aORs, for binary outcomes) and adjusted common odds ratio (acOR, for full ordinal scale analysis) with respective 95% confidence intervals (95%CI) were reported. Pre-planned subgroup analyses were performed to explore the effects of age (<65; 65-79; >80), baseline severity (NIHSS <8; NIHSS 8-16; NIHSS>16), stroke type (ischaemic vs. haemorrhagic), treatment with thrombolysis, premorbid mRS, geographic region (Australia, New Zealand, France, Malaysia, Singapore, China and Brazil), and time to first mobilisation (<24 hours; 24-48 hours; >48 hours). For all subgroup analyses, treatment-by-subgroup interactions

were tested by including multiplicative interaction terms in respective regression models. A complete case analysis was performed for all the outcomes.

4.4. Results

Datasets from six out of 10 eligible trials were obtained representing 82% of all eligible participants. Figure 4.1 displays the study inclusion flowchart. The data from one trial are openly available in Dryad at https://doi.org/10.5061/dryad.jh15q.[242] The remaining data that support the findings of this study are available upon reasonable request. Individual trial-level characteristics such as geographical region, sample size, and patient and treatment characteristics are presented in Table 4.2.

Table 4.2. Summary of individual trial-level characteristics

	Sample size	Region(s)	Aim for mobilisation activity	Participant median age (% male)	Median stroke severity (IQR)	Median time to first mobilisation for EM (median; IQR)	Median time to first mobilisation for usual care (median; IQR)
AVERT II[238]	2104	Australia, New Zealand, Singapore, Malaysia, UK	Earlier and more	75 (54%)	8 (5-15)	19 (14-22	31 (23-40)
AVERT III[9]	71	Australia	Earlier and more	71 (61%)	7 (4-12)	19 (13-22)	23 (17-29)
SEVEL[242]	138	France	Earlier	72 (64%)	6 (4-10)	26 (22-30)	71 (67-74)
Tong et al[243]	248	China	Earlier and more	61 (77%)	5 (3-8)	36 (21-42)	N/A
VERIS-Brazil[240]	37	Brazil	Earlier and more	65 (35%)	8 (5-15)	43 (29-48)	72 (61-108)
VERITAS[239]	32	Scotland	Earlier and more	65 (50%)	4.5 (2.5-8)	27 (26-29)	32 (23-47)

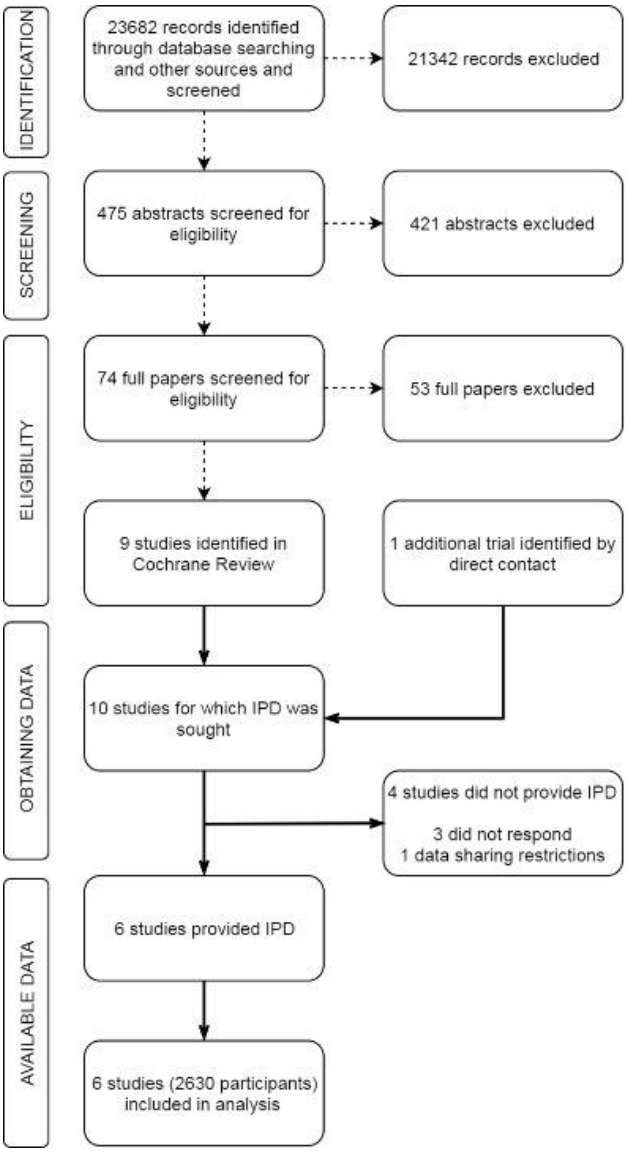


Figure 4.1. Modified PRISMA flowchart. Original 2019 Cochrane Review (dashed arrows and current IPD meta-analysis (solid arrows).

Data were available in all trials for the following pre-specified variables: age, sex, hypertension, ischaemic heart disease, diabetes, smoking status, atrial fibrillation, premorbid mRS, stroke type, NIHSS, thrombolysis, time to first mobilisation, mRS, Barthel Index and death status at 3 months. The proportion of missing data for the primary outcome was less than 1%. The following variables were not available from the indicated trials: dyslipidaemia (VERITAS), death at 14 days (SEVEL), length of acute hospital stay (Tong et al) and Oxford Classification (VERIS-BRAZIL, SEVEL,

Tong et al). Intervention data (number of out-of-bed sessions per day, minutes of out-of-bed activity per day and total intervention period) was sought, however the variables were usually not available or were not comparable (Table 4.3).

Table 4.3. Early mobilisation dose characteristics

Trial	Available Dose Data	Intervention Description
AVERT II [238]	The total minutes per day over a 14-day period or until discharge.	Mobilisation in the usual care group was provided by physiotherapists (30–60 min per day) and nurses. Patients in the intervention group were upright and out of bed (sitting or standing as able) at least twice per day; in addition to their usual care, 6 days per week.
AVERT III [9]	Session duration (mins) and number of sessions per day over a 14-day period or until hospital discharge.	Result in at least three additional out-of-bed sessions to usual care. Functional ability dictated intervention dose, with four levels specified, and dose was adjusted in line with recovery (titrated). The intervention period lasted 14 days or until discharge from stroke-unit care
SEVEL [242]	Total number of minutes during the first out-of-bed sitting session.	For both protocols, minimal duration of the first sitting was 15 minutes and maximum of 60 minutes depending on patient fatigue and tolerance. Sitting was repeated daily according to initial tolerance of the procedure, as approved by the physician in charge.
Tong et al [243]	Average daily session duration (mins) over 10-14-day period.	(1) ERM - besides standard medical care, patients in this group started <1.5 h/d (lower dose) of out-of-bed mobilisation within 24–48 hours after stroke onset. (2) EIM - besides standard medical care, patients in EIM started the ≥3 h/d out-of-bed mobilisation within 24–48 hours after stroke onset. (3) VEIM - besides standard medical care, patients in VEIM started the ≥3 h/d of out-of-bed mobilisation within 24 hours of stroke onset. Out-of-bed mobilisation included sitting, standing, and walking which were performed with or without assistance as described by the AVERT III protocol. The frequency, dose, and content of mobilisation varied according to physical ability and were recorded in detail by therapists or nurses. Mobilisation continued for 10–14 days including the weekend.
VERIS-Brazil [240]	Session duration (mins) and number of sessions over a 14-day period or until discharge.	The duration of standard-care therapy sessions was approximately 15 min. Intervention group patients were mobilized 5 times a week, once a day, for approximately 30 min per session, in addition to sitting out of bed for at least 30 min whenever possible.
VERITAS [239]	Daily minutes of standing or walking, recorded using activity monitors.	Standard care group received early mobilisation by a physiotherapist (30-60 mins per day) and nurses. The intervention group received standard care plus a protocol of early mobilisation based on AVERT II. The aim was to get patients to sit, stand and walk within 24 hours of stroke onset and continue this at least 4 times per day. An accelerometer was used to measure time (min) spent sitting/lying, standing, and stepping for patients. This was measured on days 3, 4, and 5, with recordings on the first day considered most reliable.

Overall, data from 2630 individual patients was obtained, 1437 (54.6%) assigned to EM and 1193 (45.4%) assigned to usual care. There was an anticipated difference in the proportion of patients between the EM and usual care groups due to the design of the trial by Tong et al,[243] which randomised participants into three intervention groups that differed by dose, but all starting EM <48 hours post-stroke. The mean age was 69.8±13.0 years, 62% were male and median baseline NIHSS was 7 (IQR, 4-11). The baseline characteristics were similar between both groups (Table 4.4). Participants were recruited from Australia and New Zealand (50%), Europe (30%), Asia (19%) and South America (1%). Out of the six studies, three included patients with ischaemic stroke only and three included both ischaemic and haemorrhagic strokes. The proportions of stroke type were the same for both EM and usual care groups (90% ischaemic, 10% haemorrhagic). The median (IQR) delay to starting mobilisation post-stroke onset was 20hrs (14.5-23.8) in the EM group and 23hrs (16.7-34.3) in the usual care group (p<0.001).

4.4.1. Primary and Secondary Outcomes

Fewer EM participants had a favourable outcome at 3 months post-stroke compared to the usual care group (678 [48%] vs. 611 [52%]; aOR=0.75, 95%CI:0.62-0.92, p=0.005). No difference in death at 3 months post-stroke between EM and usual care was observed (102 [7%] vs. 84 [7%]; aOR=1.46, 95%CI: 0.92-2.31, p=0.108). No significant difference in Barthel Index scores between both groups at 3 months post-stroke was observed (MD 0.16; aOR=1.21, 95%CI: 0.71-2.06, p=0.489). The ordinal analysis did not show a significant difference between groups across the entire modified Rankin Scale (Table 4.5). The pre-specified subgroup analyses for the primary outcomes are shown in Figure 4.2. Within each individual subgroup analysis,

no significant interactions were recorded. The median length of acute care hospital stay was 7 days (IQR 4-13) for both EM and usual care groups.

Table 4.4. Baseline characteristics of all patients

Variables	UC n=1193	EM n=1437
Age, years	71.05 (SD 12.70)	68.73 (SD 13.20)
Sex		
Female	476 (40%)	522 (36%)
Male	717 (60%)	915 (64%)
Hypertension	811 (68%)	965 (67%)
Ischaemic Heart Disease	283 (24%)	292 (20%)
Dyslipidaemia	470 (40%)	634 (45%)
Type 2 Diabetes Mellitus	257 (22%)	345 (24%)
Atrial Fibrillation	272 (23%)	280 (20%)
Smoking		
Never Smoked	596 (50%)	686 (48%)
Smoker	198 (17%)	315 (22%)
Past Smoker	385 (32%)	407 (28%)
Unknown	14 (1%)	29 (2%)
Geographical Region		
Australia/New Zealand	659 (55%)	655 (46%)
Asia	125 (10%)	374 (26%)
Europe	390 (33%)	390 (27%)
South America	19 (2%)	18 (1%)
Premorbid modified Rankin Scale		
0	893 (75)	1140 (79)
1	178 (15)	167 (12)
2	114 (10)	120 (8)
3	6 (1)	9 (1)
Stroke Type		
Ischaemic	1070 (90%)	1292 (90%)
Intracerebral Haemorrhage	123 (10%)	145 (10%)
NIHSS	7 (4;12)	6 (4;11)
Mild	656 (55%)	842 (59%)
Moderate	347 (29%)	416 (29%)
Severe	190 (16%)	179 (12%)
Thrombolysis	267 (22%)	309 (22%)

Data are mean (SD), n (%), median (IQR). NIHSS, National Institutes of Health Stroke Scale.

Table 4.5. Primary and secondary outcomes

	UC	VEM	Adjusted analysis*		Unadjusted analysis	
	(n=1185)	**(n=1420)**	OR (95% CI)	p	OR (95% CI)	p
			Primary			
Favourable outcome (mRS 0-2)	611 (52%)	678 (48%)	0.75 (0.62-0.92)	0.005	0.88 (0.74-1.03)	0.112
Death at 3m	84 (7%)	102 (7%)	1.28 (0.91-1.80)	0.154	1.22 (0.90-1.65)	0.195
			Secondary			
Barthel Index	81.13 (SD 30.04)	81.95 (SD 28.52)	1.07 (0.60-1.92)	0.813	1.21 (0.71-2.06)	0.489
Barthel Index > 75	811 (74%)	787 (74%)	0.92 (0.73-1.15)	0.453	1.01 (0.83-1.23)	0.916
mRS category			0.89 (0.77-1.03)	0.128	0.98 (0.65-1.46)	0.914
0	119 (10%)	125 (9%)				
1	234 (20%)	278 (20%)				
2	257 (22%)	273 (19%)				
3	242 (20%)	328 (23%)				
4	135 (12%)	214 (15%)				
5	114 (10%)	100 (7%)				
6	84 (7%)	102 (7%)				

Data are n (%) or mean (SD). mRS, modified Rankin Scale; SD, standard deviation. *Adjusted for age and stroke severity (NIHSS).

4.4.2. Trial Quality and Conduct

The risk of bias assessments for each study are shown in Table 4.6. Overall, only two trials were identified as low risk of bias.[9,238] One trial demonstrated unclear risk in selection bias, [243] while all other trials displayed low risk of selection bias (random sequence generation and allocation sequence) and reporting bias. Two trials demonstrated high risk of performance bias;[240,243] two trials demonstrated unclear risk of performance bias.[239,242] High risk of detection bias was identified in one

trial.[242] A high risk of attrition bias was identified in two trials; one trial demonstrated unclear risk of attrition bias.

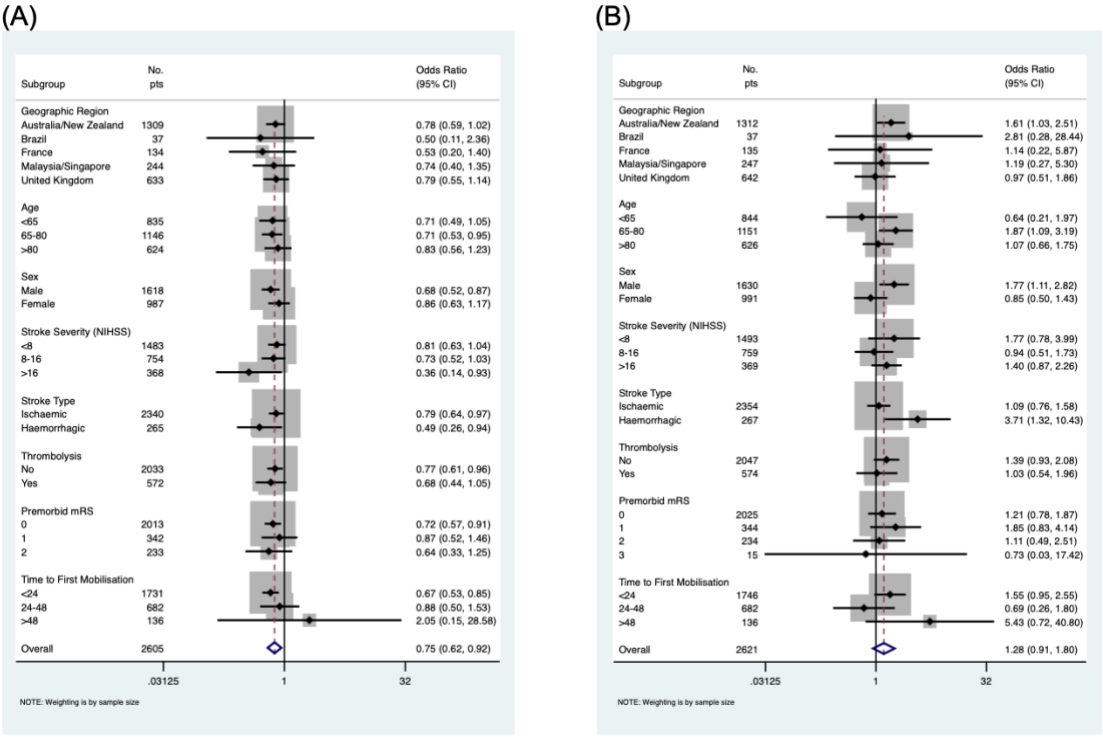


Figure 4.2. Subgroup analyses. Forest plots for the primary outcome of (A) favourable functional outcome (modified Rankin Scale score 0-2), and (B) death at 3 months in all patients. None of the individual subgroup analyses had significant treatment-by-subgroup interactions ($p>0.05$).

Table 4.6. Risk of bias assessment

Trial	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding (performance bias and detection bias)	Blinding of outcome assess-ment (detection bias)	Incomp-lete outcome data (attrition bias)	Selective reporting (report-ing bias)
AVERT II[238]	Low	Low	Low	Low	Low	Low
AVERT III[9]	Low	Low	Low	Low	Low	Low
SEVEL[242]	Low	Low	Unclear	High	High	Low
Tong et al[243]	Low	Unclear	High	Low	Unclear	Low
VERIS-Brazil[240]	Low	Low	High	Low	High	Low
VERITAS[239]	Low	Low	Unclear	Low	Low	Low

Intervention protocols, protocol development and monitoring reporting varied considerably between trials. In one study, EM was defined as out-of-bed activity starting within 48 hours of stroke onset,[240] while in the remaining five studies EM was defined as out-of-bed activity starting within 24 hours of stroke onset, [9,238,239,242,243] with further description of these activities and their titration over time described in only a few trial reports (Table 4.3). We observed differences in description of intended dose and type of intervention between trials and in the reporting of actual (delivered) intervention in the trials. While the trials in this IPD were completed before the development of the SRRR intervention development recommendations[12] and the TIDieR checklist[244] helped us expose the heterogeneity in intervention development, monitoring and reporting in the trials in this review (Table 4.7). We found that most SRRR recommendations and TIDieR Checklist items were partially described in the included trial reports. In three trials the usual care comparator was described in the same level of detail as the intervention.[9,238,239]

Intervention training and monitoring was not described or partially described in trial reports. In all trial reports, the timing of first administration of the intervention and the dose administered was provided. While further detail related to the intervention may have been provided in the trial protocols, few publications included the full protocol in supplementary materials.

Table 4.7. SRRR recommendations and TIDieR checklist items

		AVERT II	AVERT III	SEVEL	TONG	VERIS-Brazil	VERITAS
Intervention Development	Intervention development followed a structured framework	ND	ND	ND	ND	ND	ND
	Co-design research (where appropriate)	ND	FD	ND	ND	ND	ND
	Theoretical rationale underpins the development	PD	FD	FD	PD	PD	FD
	Usual care is understood	PD	PD	PD	PD	PD	FD
	Clear rationale for each intervention component	PD	PD	PD	PD	PD	PD
	Clear description of core components	PD	FD	FD	PD	PD	FD
	Clear description of the potential variable contexts in which the intervention can be delivered	NA	NA	NA	NA	NA	NA
	Considered and addressed issues impacting intervention implementation into clinical practice	FD	PD	PD	PD	FD	PD
	Iterative modelling prior to definitive evaluation	PD	PD	ND	ND	ND	PD
Intervention Monitoring	Assessment of fidelity in protocol (also Item 11 TIDieR)	PD	PD	ND	PD	ND	ND
	Method for measuring that is trial and intervention-specific	PD	PD	ND	PD	ND	PD
	Routine reporting of training in delivering the intervention	PD	PD	ND	ND	PD	ND
Intervention Reporting	Used appropriate reporting guideline for study design	ND	ND	FD	ND	ND	FD
	Provide trial registration number	FD	FD	FD	FD	FD	FD
	#1 Brief Name of Intervention	FD	FD	FD	FD	PD	FD
	#2 Why this intervention?	PD	FD	FD	PD	PD	FD
	#3 What materials are used?	PD	PD	FD	FD	PD	PD
	#4 What procedure is followed?	PD	PD	FD	PD	PD	PD
	#5 Who provides each aspect of the intervention?	PD	PD	PD	PD	PD	PD
	#6 How is the intervention delivered	FD	FD	FD	PD	FD	FD
	#7 Where is the intervention delivered?	PD	PD	PD	PD	FD	PD
	#8 When and how much the intervention is delivered?	FD	FD	FD	FD	FD	FD
	#9 Was the intervention tailored?	PD	PD	PD	PD	PD	PD
	#10 Modifications?	PD	PD	PD	PD	PD	PD
	#11 Adherence or fidelity assessed?	PD	PD	ND	ND	PD	PD
	#12 Was the intervention delivered as planned?	PD	PD	FD	PD	PD	PD

ND, Not described; PD, Partly described; FD, Full described; NA, Not applicable. Note: unpublished, full protocols, including intervention development, training and monitoring information from authors at the time of other data acquisition were not used when assessing these items.

4.5. Discussion

In this IPD meta-analysis we found that EM significantly reduced the odds of a favourable functional outcome (mRS 0-2) at 3 months post-stroke compared to usual care. This was observed despite a 3-hour difference in time to first mobilisation favouring EM; possibly suggesting that the often-larger dose of therapy received in the EM group had a greater influence on outcome at 3 months. No significant difference in death at 3 months between groups was observed, which was similar to the findings presented in the recently published Cochrane review of VEM.[14] An interesting discrepancy, however, is the aggregated Cochrane analysis found no significant impact of EM on favourable outcome at 3 months. Similarly, the Cochrane analysis found a higher mean Barthel Index score at 3 months in the EM group compared to usual care. In contrast, our IPD meta-analysis, which we acknowledge did not include all studies in the Cochrane review, found no significant difference between groups. The observed difference could also indicate an artefact of differential survival. The absence of all available trials resulted in AVERT III contributing to 80% of the total sample size in our IPD meta-analysis. If all eligible trials provided IPD, AVERT III would have still accounted for 66% of the total sample. The discrepancy in findings may also be due to the inability to adjust for prognostic variable such as age and stroke severity and meta-analysing IPD data rather than aggregated data particularly when the authors of the Cochrane identified statistical heterogeneity and missing data across the included trials, which are somewhat overcome when using IPD.

The availability of IPD allowed pre-defined subgroup analyses, but no statistically significant interactions were found. The potential signals of harm may be clinically important and justify further research. For example, similar to the findings from

AVERT III, EM appears more harmful for patients with severe stroke and haemorrhagic stroke, although this was also not significant in this IPD. When time-to-first mobilisation subgroups were compared (i.e. <24 hours vs. 24-48 hours vs. >48 hours), no significant differences in favourable outcome and death at 3 months were found. Several CPGs recommend starting intensive out-of-bed activities within 48 hours, but no earlier than 24 hours post-stroke.[205,220,222] However, further research is required to confirm the optimal timing, frequency and duration of early mobilisation within the 48-hour post-stroke time period.

Aggregating IPD is helpful to detect clinically meaningful signals since many rehabilitation trials are often underpowered to detect treatment effects in subgroups. We have identified several limiting factors that can prevent aggregation and comparison of IPD in this research domain. Only two out of six trials were identified as having low risk of bias. We also found a lack of consistency in intervention development, monitoring and reporting, which is concerning since many factors can influence how an intervention is delivered and ultimately the success of that intervention. Our results should be viewed in this light. Appropriate and uniform intervention development, outcome selection and publication reporting are required to accurately investigate optimal mobilisation strategies. A number of general (TIDieR) and more specific (SRRR recommendations) tools and checklist are available to help support trial development and reporting and may, in future, make aggregation of data more effective.[12,244,248] In future IPD comparisons, it may be advantageous to acquire full protocols, including intervention development, training and monitoring information from authors at the time of other data acquisition to more accurately assess these issues. Likewise, the relatively recent momentum towards data sharing has obvious implications on the future trial design, conduct and reporting.

There is an ethical obligation to responsibly share deidentified IPD generated by interventional clinical trials as proposed by the International Committee of Medical Journal Editors.[249] However, inability to share data including data sharing restrictions and participation non-responders were reasons that inhibited the inclusion and analysis of all eligible trials in this study.

This study has limitations. Authors of four out of the 10 eligible trials did not respond or were not able to share data. Data from only one trial was submitted to a public repository.[242] Secondly, the original search was conducted in 2017, but all ongoing trialists at the time of the initial search were contacted resulting in the inclusion of an additional trial, and systematic searches were also conducted in clinical trial registry websites. Thirdly, studies recorded actual (not intended) time to first mobilisation so delayed time to first mobilisation could be due to a) trial allocation and/or b) a clinical decision to delay mobilisation because of ill-health for someone randomised to EM. Lastly, although the timing and dose of EM were generally well defined in all trials, the frequency, duration and intensity of the EM interventions were variably pre-specified inhibiting exploration of dose effects of EM.

The strengths of this study relate to the strengths of conducting an IPD meta-analysis to supplement the Cochrane review.[14] By using IPD, we accounted for missing data, standardised statistical analyses across studies, adjusted for prognostic factors and explored subgroup–treatment effect interactions consistently across studies increasing the validity of our findings. This study also highlighted aspects of intervention development and monitoring that inhibit IPD meta-analyses in this research domain that need to be addressed before future meta-analyses of EM can be explored.

Aggregation of IPD offers the possibility of improved precision in treatment effects. The constraints we experienced in this study mean that the optimal dose (frequency, duration and intensity) of mobilisation, as well as identification of responders and non-responders to treatment remains unclear. Optimising trial design and improving reporting standards must progress in parallel with data sharing expectations if the opportunity IPD can offer is to be realised.

4.6. Chapter contribution to the thesis objective

In this chapter, the findings from Study 2 suggest that caution with mobilisation commencing within 24 hours is reasonable. This aligns with several current CPGs recommendations identified in the previous chapter. However, insufficient detail about mobilisation interventions and usual care in many studies limited any further interpretation. The identified inconsistencies in intervention development and reporting is a barrier in using population-level evidence to inform individual patient-level decision-making.

If the optimal dose (frequency, duration and intensity) of mobilisation, as well as identification of responders and non-responders to EM remain unclear, then how do clinicians tailor their decisions to individual patients (as indicated in Chapter 2) particularly when CPGs do not adequately support decisions? This is explored in the next chapter.

Chapter 5

Study 3: Clinical and Systems of Care Factors Contributing to Individual Patient Decision-Making for Early Mobilisation Post-Stroke

5.1. Chapter overview

Since the publication of AVERT, the clinical and non-clinical factors that influence the way clinicians tailor their decisions based on their own interpretation of potential risks vs. benefits of EM to individual patients has not been investigated.

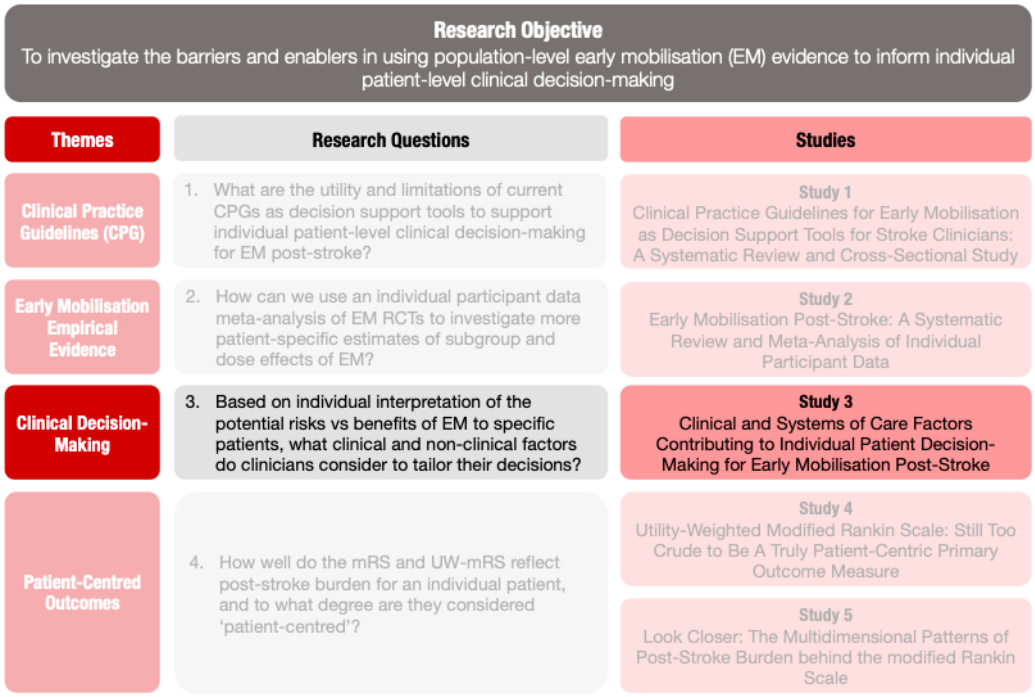


Figure 1.1. Thesis Framework.

In this chapter, the work from Study 3 is presented which aimed to identify: (1) important patient and systems of care factors influencing EM decision-making; and (2) high-interest patient subgroups, i.e., suspected responders (patients who might

benefit from EM), non-responders (patients who might be harmed from EM), and patient subgroups that create uncertainty in EM decision-making. The research addressed Research question 3 (Figure 1.1).

The chapter is presented in the *original submitted version* of the manuscript (Sections 5.2 – 5.5) followed by Section 5.6 outlining the contribution of the chapter to the thesis objective. Ethics approval was obtained from The University of Melbourne Human Research Ethics Committee (1851680.1) and can be found in Appendix C.

This chapter has been submitted for publication to *Disability and Rehabilitation.*

5.2. Introduction

It is well-established that the delivery of care in organised acute stroke units is effective in reducing death and long-term dependency.[157] Early mobilisation (EM), defined as sitting out of bed, standing or walking early after stroke, is a component of this care model.[162] The susceptibility of patient inactivity during hospitalisation[163,164] and the increased risk of immobility-related complications such as infections, falls, and thromboembolism leading to poor functional recovery, forms the central theoretical rationale for EM.[9,165-167] Equally, some concerns exist about EM including increased risk of falls,[165] impairment of cerebral blood flow and perfusion due to hemodynamic changes,[173,174] and risk of further bleeding in haemorrhagic strokes[175] or after thrombolysis.[176]

Significant changes to CPGs after the publication of the single largest EM RCT, A Very Early Rehabilitation Trial (AVERT; n=2104),[9] which demonstrated poorer outcomes in the EM group compared to usual care. Many guidelines now recommend *against* starting [intensive] out-of-bed activity within 24 hours post-stroke.[10] Few guidelines address care after the first 24-48 hours. Of those guidelines that recommend EM within 48 hours, they contain little information about how this should be tailored to individual patients.[220,222,230,232,233]

We currently do not know how clinicians make individualised clinical decisions about which patients should be mobilised, particularly in the absence of clinical trials that are powered to detect subgroup effects. Differences in decision-making have been shown to be influenced by the *interpretation* and *understanding* of the evidence, the *heterogeneity of stroke factors*, and individual clinical judgements of the *benefits and harms of EM* to different patient subgroups.[8,175,189] Investigating the specific

clinical and non-clinical factors that influence clinicians to tailor EM decision-making is important to understand.

We conducted an interactive session with expert clinicians that included an introductory semi-structured interview followed by an assisted data exploration using an EM data visualisation tool. The aim of the session was to: (1) investigate important patient and systems of care factors influencing EM decision-making; and (2) identify high-interest patient subgroups, i.e., suspected responders (patients who might benefit from EM), non-responders (patients who might be harmed from EM), and patient subgroups that create uncertainty in EM decision-making.

5.3. Methods

We report this cross-sectional study in accordance with Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.[209] Ethics approval was obtained from The University of Melbourne Human Research Ethics Committee (1851680.1). Consent was obtained from all participants involved in the study.

5.3.1. Study Participants and Recruitment

Expert sampling[215] was used to recruit a representative sample of experienced stroke clinicians. The role of these individuals was to provide an in-depth understanding of the factors that influence early mobilisation decision-making. Our sample targeted senior stroke physicians, physiotherapists and nurses/nurse practitioners with a high level of expertise in the delivery of acute stroke care and early mobilisation practices (senior position, consistent with >6 years stroke-specific experience). This sample represents content experts in early mobilisation (i.e., participated in early mobilisation trials or conducted early mobilisation research, but might not currently work in an acute stroke unit). Clinicians identified as appropriate for participation

according to this sampling strategy were invited to participate via email and were recruited between June 2019 to March 2020.

5.3.2. Introductory Semi-Structured Interview

The interviews were semi-structured, and questions were designed to elicit the desired topics of interest: important clinical and non-clinical factors for effective decision-making and high-interest patient subgroups (Appendix E). Participants were encouraged to provide more detailed responses using probing techniques and prompts. All participants were interviewed either in person or via videoconference. Demographic data on participants were collected using a brief questionnaire. A section of the introductory interview had an independent aim to investigate the utility and limitations of current early mobilisation CPGs. The results of this investigation are beyond the scope of this study and reported elsewhere.

5.3.3. Assisted Data Exploration

The AVERT Atlas is an interactive data visualisation tool (Figure E1-E3, Appendix E) for understanding and investigating complex clinical trial data from the 2104 patients included in AVERT. It was developed to be run using the statistical software R,[250] and was built using the Shiny[251] and ggplot2[252] packages. The AVERT Atlas allows the user to select a group-level combination of different patient, stroke, and dose variables as well as the outcome on the modified Rankin Scale (mRS) collected in AVERT. Based on the subgroup selection, the participant could visualise the within-subgroup distribution of the mRS, represented by a segmented bar chart. The investigator (VR) assisted the navigation of the tool if this was desired by the clinical experts. To achieve the aim of this study, the tool was used to investigate influential clinical factors that represent high-interest patient subgroups that represent perceived

characteristics of patients who might benefit from EM (responders), patients who might be harmed from EM (non-responders), and patient subgroups that create uncertainty in decision-making. Participants were given a list of all available variables and subsequently asked three questions:

1. Which variables define the patient group of highest interest to you?

Prompt: This could be a group that you believe to be a responders or non-responders or a group that you are uncertain about.

2. Why are you interested in this group of patients?
3. What response to EM do you intuitively anticipate for this group, and why?

5.3.4. Coding and analysis

The interviews were audio-recorded and transcribed verbatim using a paid transcription program (Amberscript) by VR (PhD Candidate) to facilitate the analysis of data. A *deductive* and *directed thematic content approach*[216] was used based on the main outcomes for the study: (1) the frequency distribution of individual decision-making factors being nominated by stroke clinicians and (2) the frequency distribution of patient and stroke variables used to define patient subgroups of interest.

VR read all transcripts to make initial analytical observations about the data. Transcripts were imported into QSR International NVivo 9 to code and analyse data using a deductive approach. As per the directed thematic content approach, predetermined codes were formed based on the set of a priori and predefined interview questions.[217] The strategy for coding involved reading through each transcript and conducting line-by-line coding using the predetermined codes. Subcodes were determined during this process with subsequent analysis. The data

that could not be initially coded are identified and analysed later to determine if they represent a new category or a subcategory of an existing code.

KH double-coded 25% of all transcript across all coding themes. Differences which occurred were resolved by consensus. The rationale for only double-coding 25% of transcripts was due to the *deductive* and *directed thematic content approach* that was utilised. In this approach, a top-down process was used where many of the codes were predetermined and aligned directly with the interview questions. Ongoing discussions between the two reviewers established trustworthiness and credibility to clarify the interpretation of the data. All subcodes were discussed between VR and KH to determine overlap or divergence in sub-themes within the broad predetermined codes. KH and VR also randomly selected codes in NVivo to ensure the quotes from the transcripts accurately reflected the theme it was coded to. Descriptive statistics with frequencies and proportions were produced and reported using Stata (version 14.2; StataCorp LP, College Station, TX, USA). Representative verbatim quotes were selected for presentation.

5.4. Results

Thirty clinical stroke experts participated in this study: 11 physicians, 11 physiotherapists and 8 nurses (Table 5.1). All participants approached provided consent to participate. On average, clinicians had been practising in a stroke context for 14 years (IQR 10-25), and 50% highly rated their current level of knowledge on early mobilisation evidence.

Table 5.1. Demographic characteristics of expert stroke clinicians

Variables	No. (%)
Occupation	
Physician	11 (37%)
Physiotherapist	11 (37%)
Nurse	8 (26%)
Highest level of education	
Ph.D.	14 (47%)
Clinical Doctorate	1 (3%)
Masters (clinical)	5 (17%)
Bachelor's Degree	7 (23%)
Graduate Diploma	3 (10%)
Currently practising	
Yes	26 (87%)
No	4 (13%)
Primary stroke environment	
Acute stroke unit	22 (74%)
Inpatient rehabilitation	3 (10%)
Outpatient rehabilitation	1 (3%)
Research institute	4 (13%)
Level of knowledge on early mobilisation	
High, well informed about evidence	15 (50%)
Average, up to date with evidence	15 (50%)
Low, not up to date with evidence	0 (0%)
Number of years practising in a stroke context, median (IQR)	14 (10-25)

IQR, interquartile range.

5.4.1. Introductory Interview: Clinical Factors

Table 5.2 displays the frequency distribution of clinical factors being nominated by stroke clinicians as important for decision-making. More than 80% of the sample considered stroke severity, medical stability (e.g., haemodynamic stability, cardiovascular stability, risk of complication, respiratory and febrile) and type of stroke to be the most important factors contributing to decision-making about early mobilisation practices. Around 50% of the sample considered age, delivery of thrombolysis and/or endovascular clot retrieval, and pre-morbid comorbidities to be important clinical factors.

Table 5.2. List of clinical factors considered important to early mobilisation decision-making

Clinical Factors	n (%)	Example Quotation
Stroke severity	27 (90%)	PA24 (Physio): “I think the first things that you see as a clinician are their stroke factors. So, what was their stroke severity? So, what was their NIHSS?... I actually think that's probably a big driver...”
Medical stability **Haemodynamic** **Cardiovascular** **Risk of complications** **Respiratory** **Febrile**	25 (83%) 24 (80%) 13 (43%) 10 (33%) 6 (20%) 4 (13%)	PA2 (Doctor): Also, to consider how sick they are if they have an infection and all these things come into play. Do they have complications? That probably contributes to this mix of "How stable are they?" PA17 (Physio): “But the things I'll be looking at would be medical stability. So, you know, appropriate blood pressure, heart rate, all those sorts of things.”
Type of stroke	24 (80%)	PA5 (Nurse): “In terms of stroke classification, I think that is probably important. So if I had on my list a significant penumbra or presence of large vessel occlusion and that's within the first, you know, I’m going to say eight hours after stroke onset, I'd be very cautious about moving those patients around too much including sitting in bed.” PA6 (Doctor): “With haemorrhages, intuitively, you have to wait until they settle down. Mobilizing them, I think I'd be quite gentle”
Functional ability	18 (60%)	PA25 (Doctor): They [densely hemiplegic patients] won't be able to mobilise early because they need a lot of assistance and then you judge based on the functional outcome of the patient.” PA28 (Doctor): “If I walk in and they are walking around the stroke unit saying, ‘Can I go home?’. That's a determinant.”
Age	16 (53%)	PA18 (Physio): “Potentially elderly patients that didn't have a good baseline. So then on the flipside of that, there are those patients that had a really good baseline that show good potential.”
Thrombolysis and/or endovascular clot retrieval	16 (53%)	PA30 (Doctor): “Post procedural ECR or thrombolysis, I would not be pushing it at all. I'd just wait and see to make sure the groin is okay.” PA14 (Nurse): “I guess if someone's had thrombolysis, part of our protocol would be to keep them resting in bed for about 24 hours.”
Premorbid comorbidities **Premorbid mRS** **Fitness level**	15 (50%) 6 (20%) 5 (17%)	PA26 (Physio): “If they've got loads of comorbidities, you know they're a bit more fragile usually. Premorbid Rankin - that's certainly useful.” PA2 (Doctor): “Their premorbid capabilities in terms of moving and ambulating and age probably play a little bit of a factor in terms of general fitness...”
Consciousness	11 (37%)	PA13 (Nurse): “The patients...that I feel confident in leaving in bed would be the really altered conscious state patients”

Alertness /Delirium	11 (37%)	PA3 (Physio): “I give therapy, but tiredness or alertness or comfortability, neglect, engagement etc, are the factors that dictate how the intervention is given.” PA23 (Doctor): “I guess also other patient factors unrelated to stroke such as if they were very delirious, I'd sort of say I want to make sure that they're not a high falls risk”
Imaging **Persistent occlusion** **Stenosis** **Penumbra**	8 (27%) 7 (23%) 4 (13%) 3 (10%)	PA2 (Doctor): “Now we also have a little bit more information about the anatomy and pathophysiology, and at the same time there has been a secular trend where more and more people have advanced imaging. So, we also have more information about what vessels are open or which vessels are still blocked. I say be careful when they're still blocked vs. when they are open. Although, I must say it has not been really formalized as a ‘no, don't do it’.”
Frailty	7 (23%)	PA19 (Doctor): “In terms of spending a couple of days in bed, the damage that that might do is probably greater for people who are frail and elderly with borderline mobility to start with.”
Falls risk	6 (20%)	PA23 (Doctor): “I'd sort of say I want to make sure that they're not a high falls risk or other medical issues that might necessitate them staying in bed”
Patients’ feeling	6 (20%)	PA29 (Nurse): “The other thing, it's what the patient wants often trumps everything else at times.”
Cognition	5 (17%)	PA25 (Doctor): “Some patients have cognitive impairment or dysphasia so they can't follow commands. Hence why they are not great for mobilization.”
Communication	5 (17%)	PA27 (Nurse): “Whether they had language or communication impairment.
Neglect	3 (10%)	PA11 (Nurse): “It's that moderate [NIHSS] section that I think we [early mobilisation] have the biggest impact on because...neglect is a big factor...”
Fatigue	2 (7%)	PA17 (Physio): “I would start intervening, but I would very much grade it depending on... their level of fatigue”

5.4.2. Influential Non-Clinical Factors

Table 5.3 displays important non-clinical factors that influence EM decision-making. The importance of resources such as adequate staffing, staff experience and the availability of equipment in enabling early mobilisation practices was also apparent. The stroke experts expressed that a lack of resources are barriers to implementing EM practices.

Table 5.3. List of non-clinical factors influencing early mobilisation decision-making

Non-Clinical Factors	n (%)	Example quotation
Staffing	16 (53%)	PA26 (Physio): "Generally... you'd need two people. A minimum of two people. So, if you can't get that extra pair of hands and it's not going to be safe, you don't do it..."
Equipment	9 (30%)	PA19 (Physio): "This is problematic. I'm getting better at this as you get more senior or working longer in an area, but if you don't have access to equipment or interested staff to support you, then it takes a lot of time and a lot of energy."
Level of Education/Expertise of Staff	4 (13%)	PA7 (Physio): "I think it's this there's a discrepancy between that decision-making and whether you get somebody up based on how experienced you are, how comfortable are you moving the actual physical stroke patient"
Hospital Appointments	3 (10%)	PA22 (Physio): "So is there a lot of activity happening with this person, in which case, you wouldn't even be thinking about it. You'd just be waiting to see what's going on."
Organisation of Care	1 (3%)	PA20 (Doctor): "…the most important thing is where are they. Are they at a stroke unit or not stroke unit? Do they have a team around them which is geared towards thinking about stroke care in an organized way?"

5.4.3. High Interest Patient Subgroups

Figure 5.1 demonstrates the distribution of selected variables by expert clinicians that represent high interest patient subgroups (responders or non-responders to early mobilisation). During the interactive session with stroke experts, the following variables were found to be characteristics of responders to early mobilisation practices: patients with moderate stroke severity (NIHSS 8-16), who are younger (less than 65 years old), had an ischaemic stroke (particularly small vessel occlusion and subcortical stroke), few comorbidities (0-1) and an alert state of responsiveness.

The variables that were characteristic of non-responders to EM practices included: severe stroke severity (NIHSS>16), older age, haemorrhagic stroke, lacunar strokes, posterior circulation ischaemic strokes, persistent occlusion or high grade stenosis,

hemodynamic instability, many comorbidities (atrial fibrillation, ischaemic heart disease and type 2 diabetes mellitus) and altered state of consciousness.

There were several variables that were contradictory in response, i.e., some clinicians identified as responder characteristics and others as non-responder characteristics, highlighting some uncertainty in these subgroups. These variables included people with moderate stroke severity, elderly patients, and those with lacunar infarcts.

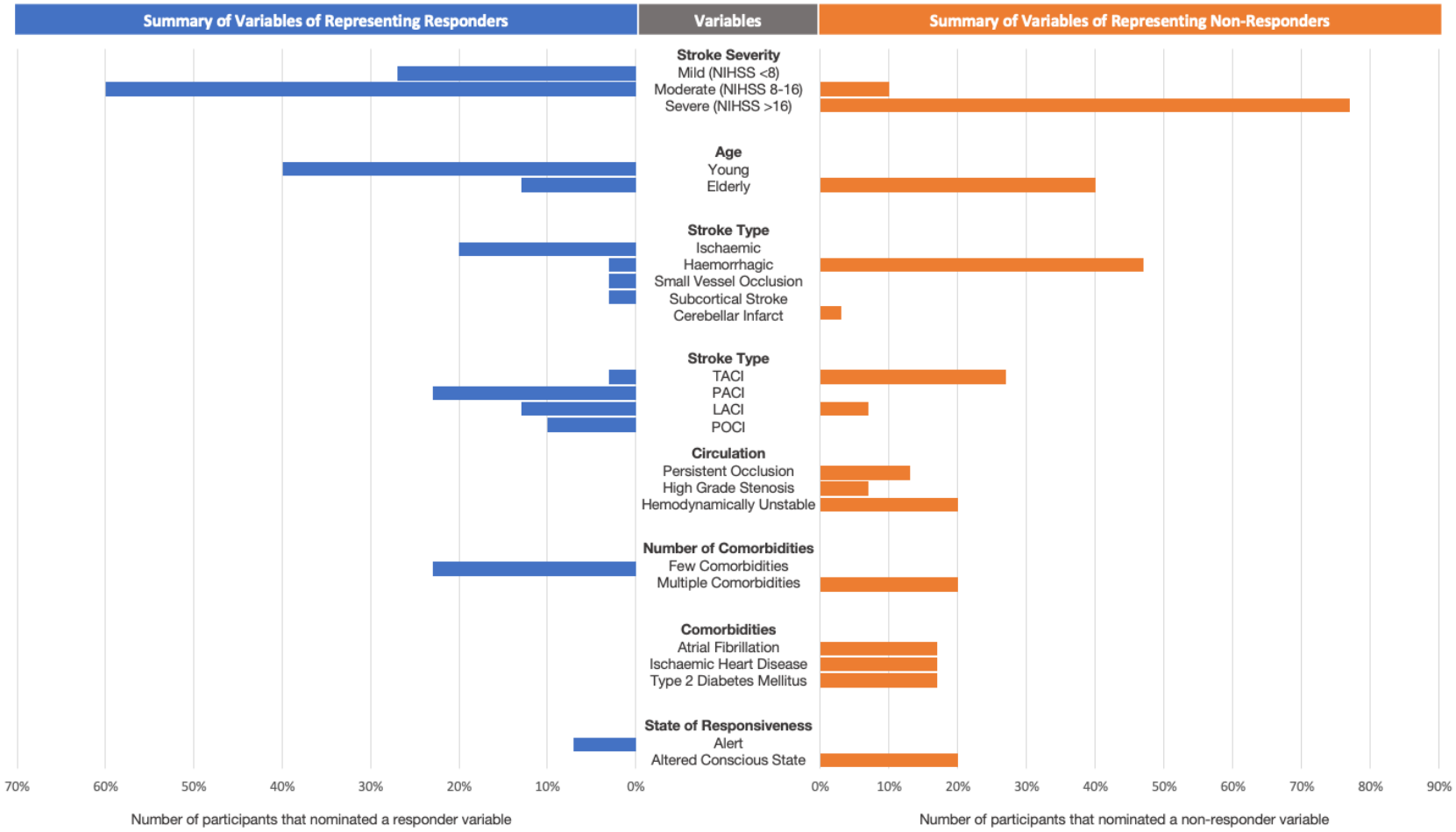


Figure 5.1. Summary of characteristics representing responders and non-responders to EM decision-making. NIHSS, National Institute of Stroke Scale; TACI, total anterior circulation ischaemic stroke; LACI, lacunar ischaemic stroke; POCI, posterior circulation stroke.

5.5. Discussion

The findings from this cross-sectional study provide in-depth insights of key patient and systems of care factors, which contribute to individualised expert decision-making for EM post-stroke. The novel interactive data visualisation session demonstrated the perceived characteristics of suspected responders, non-responders, and patient subgroups that create uncertainty in decision-making.

The top three clinical factors that influence EM decision-making were stroke severity, stroke type and medical stability. The importance of stroke severity was reiterated in the interactive data visualization session where many clinicians perceived patients with mild or moderate strokes to be responders, and patients with severe strokes to be non-responders to EM. This finding may reflect the evidence from AVERT, which demonstrated an association (albeit non-significant) between EM and poorer functional outcome in patients with severe stroke.[9] In contrast, a recent retrospective observational study in patient with severe stroke found EM was significantly associated with a lower risk of total complication rate of immobility, incidence of pneumonia, and incidence of pressure sore without increasing falls.[165] The inconsistent and inconclusive evidence base was highlighted in a recent systematic review of rehabilitation in patients with severe stroke, which found no differences in physical function and immobility-related complications between EM and the usual care group. People with severe stroke are more likely to develop acute medical complications (particularly immobility-related complications) in the initial hospitalisation phase post-stroke resulting in poor functional recovery and death.[253-255] However, the question regarding the benefits vs. harms of EM in patients with severe stroke remains. This highlights an important area for decision support and

guidance for clinicians. Future empirical investigations need to explore this area of uncertainty to help current decision-making processes.

Our findings also demonstrated that stroke type influenced decision-making i.e. patients with ischaemic vs haemorrhagic strokes. Clinicians perceived patients with haemorrhagic stroke to be a subgroup that might respond unfavourably to EM. However, it was interesting that a small sample of clinicians saw some benefit of EM in this subgroup, which highlights the heterogeneity of evidence and opinions. Prior to the publication of AVERT, and during a time with limited decision-support, many stroke clinicians believed patients with haemorrhagic strokes should stay in bed longer than those with ischaemic strokes.[175] There is some evidence that suggest EM within 24 hours might be detrimental,[9] while other studies suggest that complacency in treatment.[256,257] A recent RCT investigating EM within 24-72 hours post-stroke in mild to moderate haemorrhagic patients found a significant improvement in early functional independence and shorter length of stay compared to standard early rehabilitation.[30] Similarly, findings from an ongoing observational cohort study demonstrated that each additional day between admission and initiation of acute rehabilitation therapy (median 3 days) was significantly associated with increased odds of poor outcome (mRS 4-6) at 30 days and at 90 days post haemorrhagic stroke.[256] Perhaps the combination of stroke severity and stroke type factors may result in different responses to EM. Since there is no guidance or specific recommendations after the 24-hour period post-stroke for patients with haemorrhagic strokes,[10] variability in the delivery of care due to differences in the interpretation of harms vs benefits is to be expected, and this was demonstrated in our findings.

Finally, medical stability (e.g. haemodynamic stability, cardiovascular stability and risk of complications) was also considered an important factor contributing to

decision-making about EM practices. Establishing physiological safety criteria including systolic blood pressure, heart rate, and consciousness to guide the initiation and progress of EM has been recognised.[8] Our findings confirm the need for such criteria.

Age and premorbid comorbidities were, to a lesser extent, considered important factors in EM decision-making. Clinicians perceived patients who are elderly and with many comorbidities to be non-responders and those younger than 65 years with few comorbidities to be responders. This may reflect an understanding of the complex interplay of both factors such that an elderly patient with few morbidities is perceived to have the capacity to respond more favourably to EM than an elderly person with many comorbidities. This was similar to findings in an audit of 300 acute stroke patient records which found patients < 65 years were significantly more likely to receive EM than older patients.[258] The audit also found quality of care provided by allied health professionals was significantly associated with premorbid independence and comorbidities. A recent study also found that comorbidity burden is independently associated with patient-rated perceived impact (Stroke Impact Scale index) within the first-year post-stroke.[259] In summary, the level of comorbid burden of a patient may be an important factor that needs to be considered when assessing the potential benefit or risk of EM for a patient post-stroke.

Around half of the sample considered the delivery of thrombolysis or endovascular clot retrieval to be important factors that impact EM decision-making. The reasons for these included unresolved groin puncture post endovascular clot retrieval, and increased risk of falls and symptomatic intracerebral haemorrhage post thrombolysis. A study of 54 clinicians using hypothetical vignettes revealed that neurological decline, neurological decline with symptomatic intracerebral haemorrhage, infection

of uncertain cause, severe chest infection, severe stroke, drowsiness, and confusion significantly influenced decisions to mobilise thrombolysed patients early post-stroke.[190] In the absence of decision support or guidance, a tendency to take a conservative approach exist[260] in the context of thrombolysis or thrombectomy however, there is no evidence to support this approach.[191-193]

Some important systems of care barriers were identified that warrant consideration. This included adequate staffing and the availability of equipment to enable early mobilisation practices. This supports a previously published study which identified poor teamwork, inadequate staffing, various organisational barriers, staff attitudes and beliefs, and patient-related barriers to implementing the AVERT EM intervention.[261] These are important bottlenecks in the delivery of best care and need to be considered when developing decision-support tools such as CPGs so that EM is implementable and applicable to the local setting. It may also justify allocation of healthcare resources to ensure best care is delivered and optimal patient outcomes are achieved.

The strength of this study of this study relates to the use of an interactive data visualisation tool to observe the expert stroke clinicians' interaction with a large dataset. The tool provided a structured method to explore complex stroke factors. This was advantageous in understanding specific clinical factors, and to elicit high-interest subgroups that require a greater level of clinical judgment in their decision-making process. It must be noted that all participants were repeatedly reminded that that dataset was not powered to look at subgroup effects and therefore, any effects of EM in the subgroups they selected is not should not influence their current decision-making. Nonetheless, the use of the tool allowed the transition from aggregate data

to individual patient data and providing rich information about important decision-making factors.

There are also limitations in this study. Firstly, the expert stroke clinicians in this study were from Australian metropolitan hospitals. As a result, our findings are naturally within the scope of this demographic and may not represent the opinions of clinicians and stroke experts globally, or in regional centres in Australia. However, the multidisciplinary sample of clinical (physicians, physiotherapists and nurses/nurse practitioners) and academic experts in acute stroke provided rich detail about EM decision-making. Secondly, other allied health professionals who may play an active role in EM were not included. Early work in developing EM trial interventions identified physiotherapists and nurses as the main drivers of EM practices.[163] Physicians also play an important role in determining whether patients are medically fit to commence EM in cases of uncertainty. We therefore deliberately selected physicians, physiotherapists and nurses for these studies. Future exploration of important EM decision-making factors, systems of care bottlenecks, and perceived responders and non-responders to EM in more geographically and clinically diverse samples may validate the findings from this study.

As expected, many clinical and systems of care factors play a fundamental role in decision-making at an individual patient-level. There are some patient and stroke characteristics that determine whether a patient is mobilised early or not, which are not necessarily informed by evidence. We need to be careful that inconclusive and inconsistent EM findings do not instil a sense of complacency or confidence. Nonetheless, these factors have identified important future research areas, and potentially worthwhile patient subgroups that need to be substantiated by empirical investigations to better support clinical decision-making.

5.6. Chapter contribution to the thesis objective

In this chapter, several clinical and non-clinical factors, and characteristics of different patient subgroups that influence EM decision-making post-stroke were identified. The work from this study has enabled an in-depth exploration into how individualised decisions are tailored in a sample of Australian expert stroke clinicians. By understanding these decision-making factors, it provides a feedback loop to both EM evidence generation and CPG development. Population-level evidence is needed to substantiate these processes while understanding the decision-making process is needed to inform the iterative development of tools that support this process i.e., CPGs.

Thus far in the thesis, the utility of CPGs to support decision-making, patient-specific EM evidence, and the factors important for decision-making were explored to collectively understand some of the barriers and enablers in using evidence to inform clinical decisions. However, there is growing awareness to move towards a patient-centric approach to the delivery of healthcare. With this in mind, the next chapter will focus on investigating how new and old measures of post-stroke burden might be considered ‘patient-centred’.

Chapter 6

Study 4: Utility-Weighted Modified Rankin Scale: Still Too Crude to Be A Truly Patient-Centric Primary Outcome Measure?

6.1. Chapter overview

The mRS is the most widely employed outcome measure in acute stroke trials however, is not without its limitations. The UW-mRS is a newly proposed primary outcome measure that aims to incorporate both clinically and socially defined measures to improve the interpretability and statistical efficiency of the mRS.

In the next two chapters, the work conducted was to address Research question 4 (Figure 1.1) which aims to investigate how well the mRS and UW-mRS represent post-stroke burden for an individual patient.

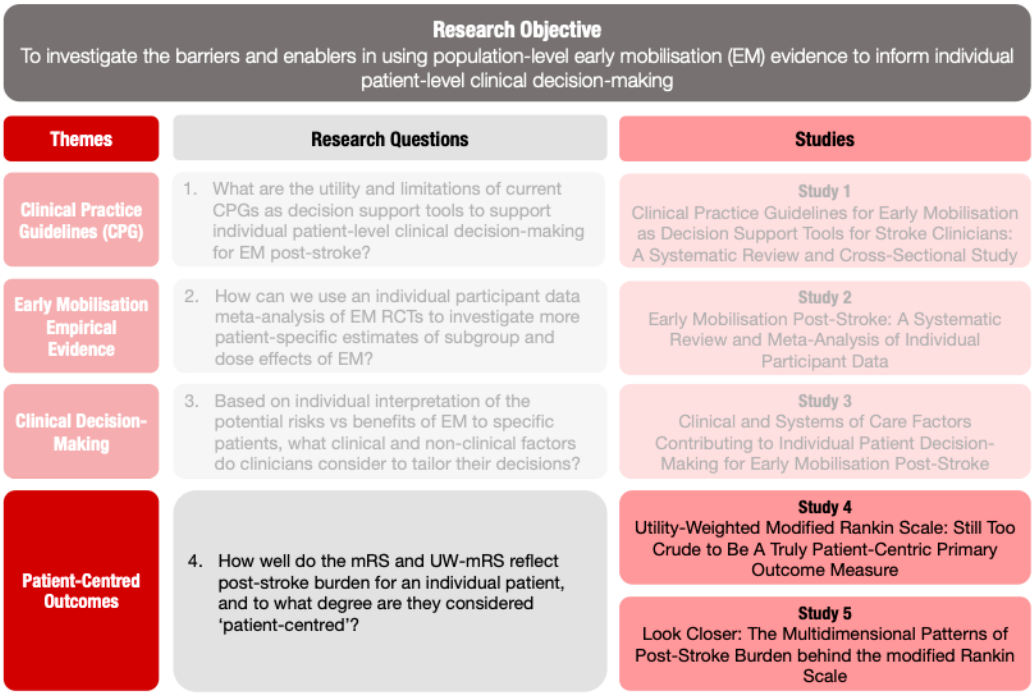


Figure 1.1. Thesis Framework.

In this chapter, the *published* work from Study 4 investigated the construct validity of the UW-mRS since the emergence and utilisation of this outcome measure will inevitably affect the interpretation of clinical trials findings and the communication to patients about the relevant benefits, risks and costs associated with different clinical decisions.

This chapter contains the *accepted version* of the manuscript (Sections 6.2 – 6.6) followed by Section 4.7 outlining the contribution of the chapter to the thesis objective. The Version of Record (PDF) of the manuscript can be found in Appendix A.

This chapter contains the Author's Original Manuscript of an article published by International Journal of Stroke that is available here.

6.2. Introduction

The recently proposed UW-mRS is derived by translating ordinal mRS categories into a utility-weighted scale.[16] It is an emerging outcome measure which aims to reflect patient perception of quality of life to improve interpretability of the mRS.[17,154] The UW-mRS also claims to improve statistical efficiency by providing a continuous alternative to either an ordinal or dichotomised analysis of the mRS.[17]

The development of a UW-mRS has been advocated by several associations including the Stroke Therapy Academic Industry Roundtable.[262] Utility, centred on health-related quality of life, is a widely accepted measure of a patient's preference for a specific health outcome.[151,152] UW-mRS values are obtained by assigning a mean utility weight to each mRS category based on patient, clinician or societal valuation of various disability health states.[155]

Since its introduction, research efforts have focused on the statistical efficiency of the UW-mRS to detect treatment effects in stroke intervention trials.[17,154] These investigations showed similar efficiency compared to the ordinal mRS, and greater efficiency compared to the dichotomised mRS. While prior investigations have focused on statistical properties, construct validity of the UW-mRS has yet to be established.

It is unclear whether a mean utility weight per mRS category sufficiently reflects the variability of utility values within mRS categories, and whether the difference in utility values adequately discriminates between mRS categories. Similarly, it is unclear whether utility weights reflect changes in patient perspectives over time post-stroke, and consequently, the long-term outcomes of an intervention to a stroke survivor. Lastly, uncertainty concerning the most appropriate utility derivation

method for a stroke population persists challenging the use of the UW-mRS as a primary outcome measure in stroke trials. [263,264]

The overall objective was to investigate the construct validity of the UW-mRS as a primary outcome measure for stroke trials. The aims of this study were:

1. Derive AQoL-4D predicted UW-mRS values per mRS category from the AVERT population at 3- and 12-months post-stroke;
2. Investigate between- and within-patient variability of individual utility values;
3. Compare the derived UW-mRS values and utility derivation method with published UW-mRS studies.

6.3. Methods

6.3.1. Conceptual Model for Assessing Threats to UW-mRS Validity

We conceptualised the potential validity threats as belonging to the following two set of circumstances. Firstly, there is the potential inability of UW-mRS values to reflect the observed variability in utility values between and within individual mRS categories. Under this scenario, considerable mismatch between the UW-mRS value and the true utility value may exist. Secondly, UW-mRS values can be derived using various elicitation methods, country-specific value sets, multi-attribute utility instruments and stroke population datasets.[265,266] Obtaining a utility involves defining and valuing health states of interest to develop a *value set*. Utility values can be elicited using two methods: direct (e.g. time trade-off method, standard gambling, visual analogue scale) or indirect (e.g. patient's response to a multi-attribute utility instrument: AQoL-4D, European Quality of Life Scale [EQ-5D-3L] and Health Utilities Index).[265]

As many stroke trials use multi-attribute utility instruments to measure quality of life in parallel to the mRS, a UW-mRS can be derived using a stroke population dataset containing responses to both measures.[16,17,154,155,266] To produce UW-mRS values, utility values are stratified by mRS category and then averaged. We, therefore, will refer to individual patient AQoL-4D utility values as *utility values* and the mean of AQoL-4D utility values per mRS category as *UW-mRS values* in this study.

6.3.2. Study Population

We used individual patient data mRS and AQoL-4D data from AVERT.[9] AVERT was a parallel-group, single-blind, RCT conducted at 56 acute stroke units in five jurisdictions: Australia, New Zealand, Malaysia, Singapore, and the UK. The aim was to compare the effectiveness of frequent, higher dose, very early mobilisation (sitting, standing or walking <24 hours post-stroke onset) with usual care after stroke. mRS data was obtained face-to-face by a trained and accredited blinded assessor. AQoL-4D was completed by the patient (88%) or a proxy (12%) when communication or cognitive impairments precluded direct questioning of the patient. The primary and secondary outcomes of the trial have been previously reported.[9] Ethics approval was obtained from institutional review boards at all sites.

6.3.3. AQoL-4D Derived Utilities

The UW-mRS values per mRS category were determined using AQoL-4D utility values and mRS scores from AVERT at 3 months and 12 months. The descriptive system for AQoL-4D was developed within the framework of the World Health Organisation's classification of impairment, disability and handicaps.[267] AQoL-4D consists of 4 dimensions and 3 levels: independent living – self-care, household tasks and mobility; relationships – friendships, isolation and family role; mental health –

sleeping, worrying and pain; senses – seeing, hearing and communication. The utility scoring algorithm is based on a stratified sample of the Australian population using the time trade-off method. This algorithm weights the items and applies a multiplicative function to achieve an index producing a life-death utility scale. The scale is bounded by the following points: 1.00 (best possible health state), 0.00 (death equivalent) and -0.04 (worse-than-death states).[268] AVERT participants who died before their 3 or 12 month follow-up were assigned a utility value of 0.

6.3.4. Statistical Analyses

To analyse the variability in utility values between individual mRS categories, we conducted the following tests. A one-way ANOVA was performed to estimate differences between mean utility values across mRS categories at 3- and 12-months post-stroke, respectively. ROC curves were generated to differentiate between sequential mRS categories based on utility values. A ROC curve area of 0.5 indicates no discrimination, whilst 1.0 indicates perfect discrimination.[269] Finally, a misclassification matrix was used to classify each patient into a "matched" mRS category based on the closest UW-mRS value to their observed individual AQoL-4D utility value. The formal agreement between the "actual" and "matched" mRS were estimated using weighted Cohen's Kappa with quadratic weights.

To explore the variability in utility values within individual patients over time post-stroke, intraclass correlations (ICC) were calculated using a random effects linear regression model using patients (random effect), time i.e. 3 vs. 12 months (input), and utility values (output). The ICC obtained from this model indicates the degree of variability that can be explained by within- and between-patient differences. The ICC is 0 when there are no between-patient differences and increases up to 1 when the

between-patient variance increases relative to within-patient variance.[270] Variability in utility values over time post-stroke in patients who did not change in mRS score from 3 months to 12 months were calculated using the previously described random effects linear regression model. If the UW-mRS is a valid representation of the mRS, particularly in patients with the same mRS score at 3 and 12 months and covariates such as age, stroke severity and co-morbidities are accounted for, similar utility values would be expected, and any observed difference should be attributed to time.

To explore the variability in AVERT UW-mRS values and previously published studies, all UW-mRS values from nine UW-mRS sets at 3 months, and two UW-mRS sets at 12 months were applied to individual patients from the AVERT dataset.[16,17,154,155,266] ICCs were calculated using a variance component model implemented as random effects linear regression using age and baseline stroke severity (inputs), utility values (output) and patients (random effects). Statistical analyses were performed with Stata (version 14.2; StataCorp LP, College Station, TX, USA).

6.4. Results

6.4.1. Study Population

All 2,104 participants from AVERT were included in our analysis. The utility values and mRS scores at 3 months were available in 2,030 participants (96.5%) including 160 participants who died before follow-up. Utility values and mRS scores at 12 months were available in 2,017 participants (95.7%), including 257 participants who died before follow-up. The study population had a mean (SD) age of 70.7 (12.7) years and 61% of participants were male (Table 6.1).

Table 6.1. Characteristics of 2030 AVERT stroke patients

Variables	Total, n=2030
Recruitment region	
Australia and New Zealand	1216 (60%)
Asia	240 (12%)
UK	574 (28%)
Age (years)	
<60	593 (29%)
60-80	906 (45%)
>80	531 (26%)
Sex	
Male	1238 (61%)
Female	792 (39%)
Risk Factors	
Hypertension	1410 (68%)
Ischaemic Heart Disease	483 (23%)
Hypocholesterolaemia	832 (40%)
Diabetes Mellitus	459 (22%)
Atrial Fibrillation	465 (22%)
Premorbid modified Rankin Scale	
0	1528 (75%)
1	292 (15%)
2	210 (10%)
Stroke History	
NIHSS	7 (4-12)
Ischaemic Stroke	1784 (88%)
Intracerebral Haemorrhage	246 (12%)
rtPA Treatment	503 (24%)

NIHSS: National Institutes of Health Stroke Scale; rtPA: recombinant tissue plasminogen. Data are n (%) or median (IQR).

6.4.2. AVERT UW-mRS Values

The UW-mRS values i.e. mean AQoL-4D utility values (SD), derived for mRS categories 0-6 at 3 months include: 0.80 (0.22), 0.78 (0.21), 0.63 (0.23), 0.37 (0.21), 0.11 (0.12), 0.03 (0.06) and 0 (0), respectively (one-way ANOVA $p<0.0001$). The UW-mRS values for mRS categories 0-6 at 12 months include 0.85 (0.19), 0.77 (0.21), 0.63 (0.24), 0.34 (0.21), 0.11 (0.12), 0.03 (0.07) and 0 (0), respectively (one-way ANOVA $p<0.0001$).

6.4.3. Utility Variability Within and Between mRS

We observed considerable variation in utility values within each mRS category at both 3 months and 12 months (Figure 6.1). The misclassification matrix showed excellent statistical agreement between the "actual" and "matched" mRS categories at 3 months (kappa, 0.81; CI, 0.79–0.84) and 12 months (kappa, 0.86; CI, 0.85–0.87) (Figure 6.2). However, based on the AQoL-4D utility values, good discrimination between sequential pairs of mRS categories was only observed for mRS 3 vs. 4 at 3 months (area under the curve [AUC], 0.87; 95% CI, 0.84–0.90) and 12 months (AUC, 0.83; 95% CI, 0.80–0.87) (Table 6.2). In contrast, utility was less discriminant for mRS 1 vs. 2 and 4 vs. 5 which reflect the matching inconsistencies in the misclassification matrix. For example, an individual with an mRS of 2 at 3 months could be classified into mRS 0 (29%), mRS 1 (13%), mRS 2 (28%) and mRS 3 (24%) based on the closest UW-mRS value. Similarly, an "actual" mRS of 4 and 5 displayed a mirrored inability to discriminate between "matched" mRS 4, 5 and 6.

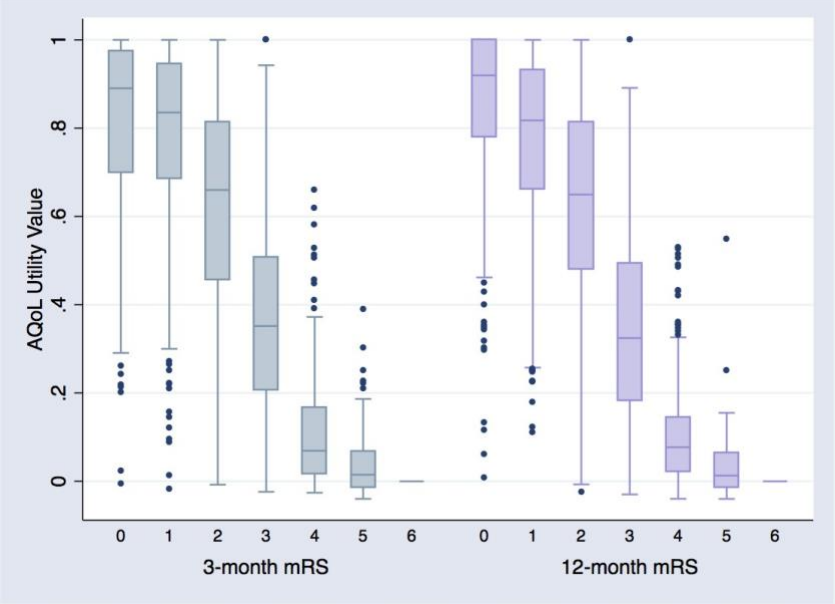


Figure 6.1. Distribution of AQoL utility values across mRS categories at 3 and 12 months.

Table 6.2. ROC curve analyses between sequential mRS categories based on AQoL utility values at 3 and 12 months

mRS Comparisons	3-month AUC (95% CI)	12-month AUC (95% CI)
0 vs. 1	0.54 (0.49–0.60)	0.64 (0.60–0.68)
1 vs. 2	0.70 (0.66–0.74)	0.68 (0.64–0.72)
2 vs. 3	0.80 (0.77–0.83)	0.81 (0.78–0.84)
3 vs. 4	0.87 (0.84–0.90)	0.83 (0.80–0.87)
4 vs. 5	0.72 (0.67–0.76)	0.75 (0.70–0.80)
5 vs. 6	0.63 (0.56–0.70)	0.61 (0.52–0.69)

mRS: modified Rankin Scale; AUC: area under curve; CI; confidence interval.

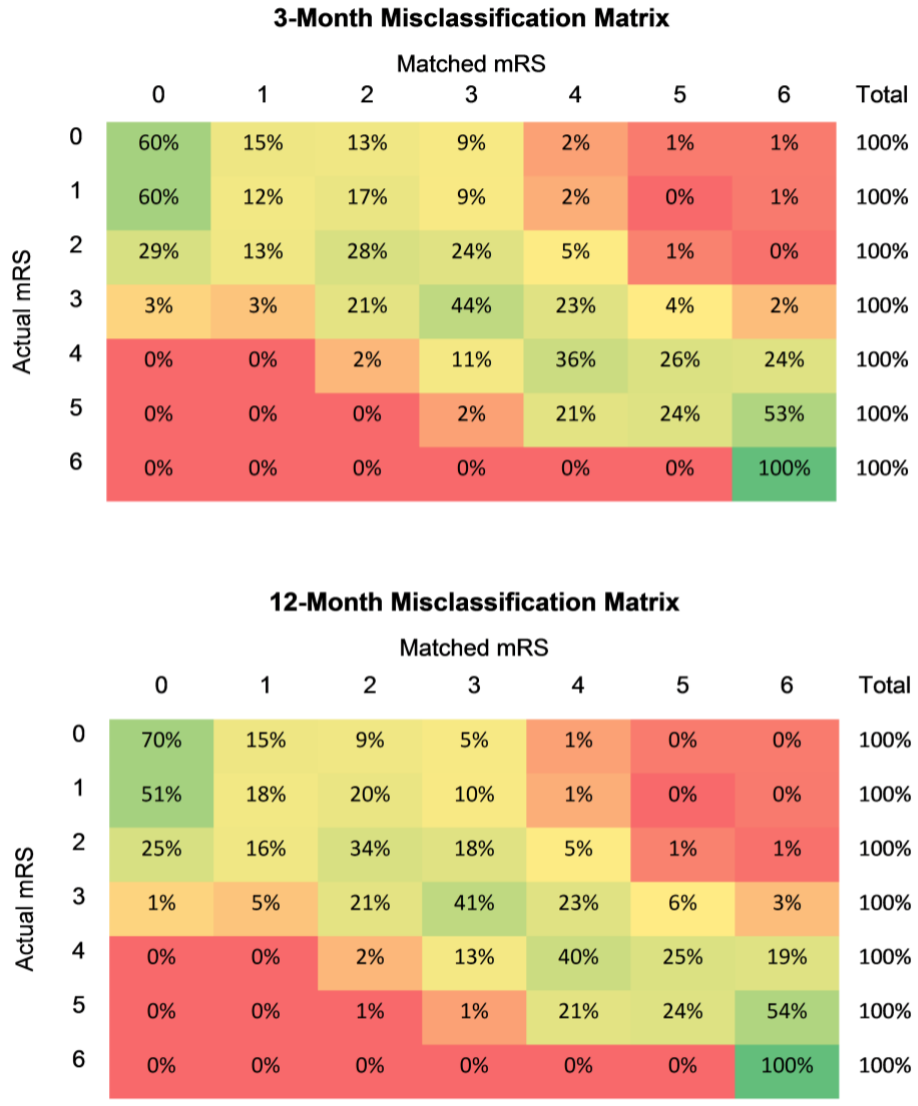


Figure 6.2. Misclassification matrices at 3- and 12-months post-stroke. The proportion of patients classified into a matched mRS based on the closest UW-mRS value to their observed AQoL-4D utility value.

6.4.4. Utility Variability Over Time Post-Stroke

We observed substantial within-patient variability in utility values for patients with no change in mRS between 3- and 12-months post-stroke (n=1189) (Figure 6.3). The degree of variability that could be explained by within-patient differences (time post-stroke) was 47% (ICC=0.53). The proportion of variability in utility values attributable to time post-stroke for the "no change" groups were: 45% (mRS 0), 42% (mRS 1), 44% (mRS 2), 54% (mRS 3), 56% (mRS 4) and 53% (mRS 5).

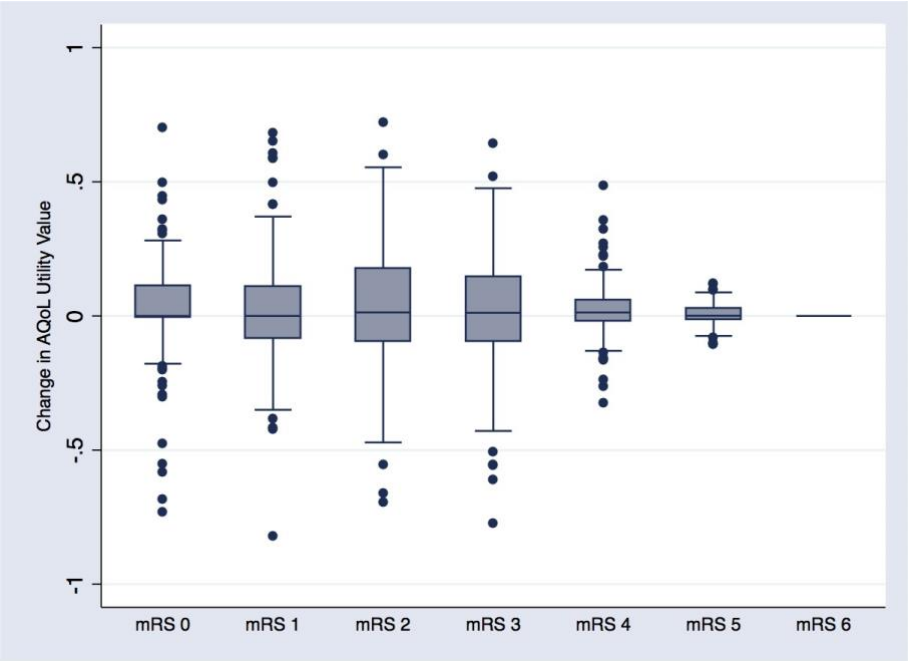


Figure 6.3. Distribution of change in AQoL-4D utility values over time. Change in utility values between 3 and 12 months for patients whose mRS did not change (n=1189).

6.4.5. Utility Variability by Derivation Method

The UW-mRS values displayed a similar sigmoidal relationship across the mRS categories consistent with previously published utility weights (Table 6.3). UW-mRS values, however, were generally lower for all categories except mRS 5 and the UW-

mRS set based on an applied Singaporean value set by Ali et al 2017. We found variability in UW-mRS values due to the choice of valuing method, the country value set applied, and the stroke population dataset analysed (Table 6.3). Taking into consideration age and stroke severity, the choice of the derivation method contributed to 25% and 22% of the variability in UW-mRS values from nine UW-mRS sets at 3 months and two UW-mRS sets at 12 months, respectively (Figure 6.4).

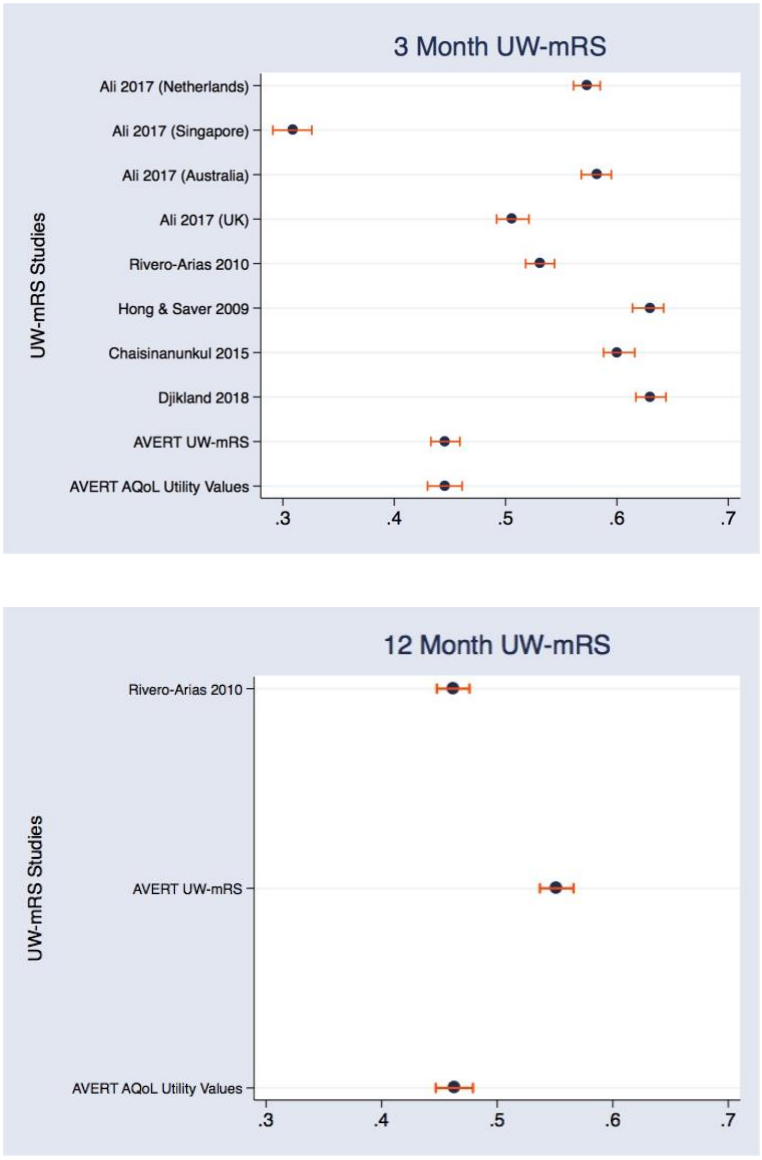


Figure 6.4. UW-mRS values from different studies applied to the AVERT dataset at 3 and 12 months (mean, 95% CI).

Table 6.3. Comparison of UW-mRS values and derivation methods between AVERT and previously published studies

Studies	Valuing Method	Country Value Set Applied	Country of Population Analysed	n	mRS Levels					
					0	1	2	3	4	5
			3 Month UW-mRS [a]							
AVERT	AQoL TTO	Australia	International	2083	0.80	0.78	0.63	0.37	0.11	0.03
			UK	601	0.78	0.77	0.60	0.35	0.09	0.04
			Australia New Zealand	1238	0.80	0.79	0.64	0.38	0.12	0.03
			Singapore Malaysia	244	0.81	0.76	0.66	0.37	0.09	-0.00
Dijkland, 2018	EQ-5D TTO	Netherlands	Netherlands	500	0.95	0.93	0.83	0.62	0.42	0.11
Ali, 2017	ED-5D TTO	UK	UK	70	0.90	0.82	0.70	0.53	0.20	-0.15
		Australia	Australia New Zealand	80	0.93	0.86	0.76	0.61	0.35	0.02
		Singapore	Singapore Malaysia Philippines	160	0.88	0.74	0.51	0.23	-0.16	-0.48
		Netherlands	Netherlands France Belgium	215	0.91	0.83	0.73	0.59	0.35	0.12
Chaisinanunkul, 2015 [b]	EQ-5D TTO & WHO-GBDP PTO	N/A	International	N/A	1.00	0.91	0.76	0.65	0.33	0.00
Hong & Saver, 2009 [c]	WHO-GBDP PTO	N/A	International	N/A	1.00	0.95	0.79	0.67	0.35	0.06
Rivero-Arias, 2010 [d]	EQ5D TTO	UK	UK	365/ 478	0.92	0.81	0.69	0.57	0.24	-0.01
			12 Month UW-mRS [e]							
AVERT	AQoL TTO	Australia	International	2052	0.85	0.77	0.63	0.34	0.11	0.03
			UK	585	0.85	0.78	0.54	0.35	0.10	0.02
			Australia New Zealand	1226	0.84	0.76	0.65	0.34	0.11	0.03
			Singapore Malaysia	241	0.91	0.80	0.66	0.34	0.13	0.02
Rivero-Arias, 2010	EQ5D TTO	UK	UK	665	0.96	0.83	0.68	0.55	0.32	-0.07

UW-mRS: utility weighted modified Rankin Scale; AQoL: Assessment of Quality of Life; TTO: time trade-off; EQ-5D: European Quality of Life Scale; WHO-GBDP: World Health Organisation Burden of Disease Project; PTO: person trade-off.

[a] Derived from data collected at 3 months post-stroke.

[b] Country value set and sample number not specified as UW-mRS was derived by averaging values from TTO and PTO studies.

[c] Country value set and sample number cannot be specified due to PTO valuing method.

[d] UW-mRS derived by averaging 1-month and 6-month UW-mRS.

[e] Derived from data collected at 12 months post-stroke.

6.5. Discussion

This study is the first to derive AQoL-4D predicted UW-mRS values which were consistent with prior studies using different derivation methods confirming the unequally ordered intervals of the mRS.[16,17,154,155,266] Our AQoL-4D UW-mRS values however, indicated poorer patient perception of their health states compared to previously derived UW-mRS values. This concurs with previous investigations of discrepancies in predicted utilities by different instruments which evaluate related but distinct constructs.[271,272] For example, nearly 80% of the difference in utility weights predicted by AQoL-8D and EQ-5D-3L can be explained by differences in their descriptive systems i.e. the set of questions and response categories which seek to describe a person's health.[271] The remaining observable difference is due to scale effects and scale-adjusted differences attributable to the utility formula. This is problematic as it compromises the comparability of utility values and, by implication, the UW-mRS.

Extensive research has established the psychometric properties of EQ-5D-3L in a stroke population.[273-275] Accordingly, EQ-5D-3L is an appealing instrument for mapping utility weights to mRS in previously published studies.[154,155,266] However, a criticism of the EQ-5D-3L, and consequently these studies, is its lack of sensitivity to health states.[265,276] The instrument was not designed to assess the needs and deficits of stroke patients. AQoL-4D yields the same generic weakness, but accounts for more stroke- and health-related quality of life domains than the EQ-5D-3L.[265] Given that AQoL-4D has shown predictive validity in a stroke population, it is possible the lower UW-mRS values observed in this study indicate a more accurate representation of stroke outcomes based on the perspective of stroke patients.[276]

Differences in UW-mRS values can also be attributed to the country-specific value set applied and the population sample from which quality of life and mRS data are collected. The application of different country-specific value sets has previously been shown to result in significantly different utility distributions.[266] Guidance on the most appropriate value set to correctly depict post-stroke outcomes while allowing aggregation of international datasets is still required.

Another threat to construct validity was the inability of UW-mRS values to reflect variability in AQoL-4D utility values between and within individual mRS categories. This may be explained by individual patient differences in the affected dimensions of the AQoL-4D, which has been previously postulated.[266] Neurological deficits vary with the region of sustained damage (e.g. motor vs. cognition), as such, the affected dimensions of AQoL-4D can also vary (e.g. selfcare vs. relationships). These dimensional differences, however, are often classified as the same level of functional impairment by the mRS.

Our findings demonstrate an inability to sufficiently discriminate levels of functional disability based on a patient's perspective of their health. This highlights an obvious issue concerning the simple yet inadequate choice to average utility values per mRS category. Weighting the mRS in this way does not reflect the variability of individual utility values within each category, reduces the sensitivity of the UW-mRS, and disregards crucial information about quality of life. If the UW-mRS only captures a marginally greater understanding of patient perspectives, the justification of its superiority over the mRS as a primary outcome measure is somewhat questionable, particularly when statistical superiority has not been established.[154]

Based on the derived utility values, mRS does not effectively capture patient perceptions of their stroke outcome or final state health outcome. The substantial variability in utility values over time post-stroke, particularly for patients whose mRS score did not change between 3 and 12 months is indicative of the mRSs oversimplification as an outcome measure and the potential disregard of changes in patient perspectives of their health over time. Similarly, we should question whether the UW-mRS is too crude to reflect long-term effects of stroke. Perhaps, utility values should exclusively supplement the mRS rather than be integrated to produce what our data suggest may be a primary outcome measure with limited validity.

The are several strengths of this study; it is the first to derive AQoL-4D UW-mRS values, AVERT provides a large, broad sample of patients early post-stroke, and the longitudinal data facilitated the exploration of patient perspectives over time post-stroke. An outcome measure should consider whether an intervention has long-term benefits for a stroke survivor and reflect changes in quality of life over time. The first limitation of this study is the selection criteria of AVERT. However, compared to other acute stroke trial datasets, it is reasonably non-restrictive, and the large, multinational sample promotes the generalisability of the conclusions. The second limitation is the application of an Australian value set to an international population which, ignores country specific differences in quality of life.[266] However, the application of customised value sets is problematic when pooling data for analyses. Standardisation of UW-mRS derivation methods is required to preserve the validity of data and the ability for data aggregation.

By introducing a meaningful concept of the difference between mRS categories for an otherwise ordinal scale, the UW-mRS provides an imperfect solution to an important problem. Findings from this study should not discourage the quest to

overcome the shortcomings of the mRS. The need to employ outcome metrics devoid of crudity while reflecting patient perspectives of quality of life is still critical.[132] Further investigation of the UW-mRSs added value, including validation on longitudinal data, is time-critical given its recent approval as a co-primary outcome measure by the Food and Drug Administration in a thrombectomy trial.[156] Nonetheless, parallel collection of longitudinal quality of life and mRS data is still warranted. However, the sources of variability shown in this study need to be considered to ensure the use of UW-mRS as a primary outcome measure is justifiable, valid, standardised, statistically efficient, and accurately represents final health states of stroke patients.

6.6. Chapter contribution to the thesis objective

In this study, we found that patients with a given mRS may exhibit a wide of range of utility values, and patients with different mRS scores can have similar utility values, making the UW-mRS not sufficiently granular to represent post-stroke burden. Therefore, the identified threats to construct validity warrants caution when using the UW-mRS as a primary outcome measure and may be too crude to be truly patient-centric.

The high level of variability in utility values in the UW-mRS scores formed the rationale for the next chapter. We *looked closer* at the *multidimensional patterns* for each mRS and UW-mRS score using individual domains of quality of life (AQoL-4D) to obtain a granular view of 3-month stroke burden.

Chapter 7

Study 5: Look Closer: The Multidimensional Patterns of Post-Stroke Burden behind the modified Rankin Scale

7.1 Chapter overview

This chapter systematically investigates both objective and patient perceived measures of how different aspects of function and disability post-stroke are impacted at each level of the mRS. We hypothesised that different post-stroke burden domains would display different patterns of decline across mRS categories. In other words, the variability see in Chapter 6 could be explained by differing patterns of decline based of the type of post-stroke impairment or burden.

Considering the mRS is currently the most utilised primary outcome measure and is not likely to be challenged in this position, adding depth of understanding that transcends the current simple description of each mRS category is essential. Furthermore, the effect of an intervention on various domains of post-stroke outcome must be first understood in order to effectively define the parameters which bound "favourable outcome". This is particularly relevant for the development of CPGs and by association clinical judgement and reasoning which are informed by this type of empirical data.

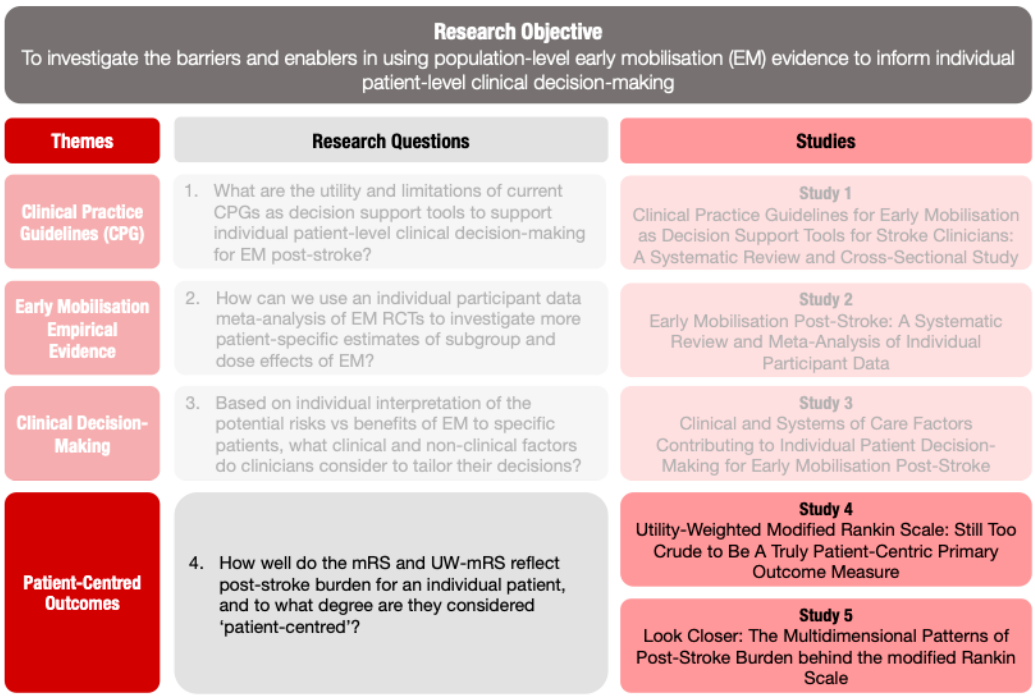


Figure 1.1: Thesis Framework.

This chapter contains the *accepted version* of the manuscript (Sections 6.2 – 6.6) followed by Section 4.7 outlining the contribution of the chapter to the thesis objective. The Version of Record (PDF) of the manuscript can be found in Appendix A.

This chapter contains the Author's Original Manuscript of an article published by International Journal of Stroke that is available here.

7.2. Introduction

The modified Rankin Scale (mRS), initially described as a measure of "post-stroke disability"[245], includes six ordered grades of functional outcome and a death state. While all mRS categories consistently and explicitly refer to a generic degree of post-stroke disability, some categories are varyingly described in terms of the presence of symptoms, ability to perform activities of daily living, walking ability and self-care. The variability in description across the scale may explain the multidimensionality of post-stroke disability. Modern stroke trials also adopt different interpretations of the mRS, inconsistently referring to the scale as a measure of neurologic functional disability[277], disability[278,279], global disability,[280-282] functional outcome[283-285], global functioning[286], and disability or dependence in daily activities[287] post-stroke. Consequently, it is challenging for a clinician or trialist to consistently interpret and communicate the multidimensional disease burden associated with different mRS categories.

The International Classification Functioning, Disability and Health (ICF) Core Set for Stroke was developed to help evaluate and classify the multidimensionality of post-stroke disability, function and health.[15] The Comprehensive and Brief ICF Core Sets for Stroke highlight the necessity for trial outcomes, such as the mRS, to describe a more representative spectrum of problems in the functioning of patients post-stroke. The core set includes domains, not limited to, neurological deficits, inability to perform activities of daily living and self-care tasks, neuropsychological dysfunction and neuropsychiatric deficits.[15,123,124] Despite this, different mRS categories do not uniformly reflect post-stroke burden, which is further amplified by the "ordered, but unequally spaced" nature of the scale.[16] The utility-weighted mRS (UW-mRS) has been proposed as a solution to this important problem.[16,17] The UW-mRS aims to

reflect patient perspectives of quality of life in a single measure that provides a continuous alternative to the mRS. Despite the advantages of the UW-mRS, patients with a given mRS may exhibit a wide of range of utility values, and patients with different mRS scores can have similar utility values, making the UW-mRS not sufficiently granular to represent post-stroke burden.[288]

This limitation may be due to the insufficient ability of the single dimensional UW-mRS to represent the multiple domains of post-stroke burden across mRS categories. In A Very Early Rehabilitation Trial (AVERT), mRS, Assessment of Quality of Life (AQoL-4D) and Barthel Index (BI) scores were simultaneously collected as primary and secondary outcomes. Collectively, the mRS, AQoL-4D and BI can measure several aspects of the Brief ICF Core for Stroke. AQoL-4D captures patient perspectives of post-stroke burden and can provide valuable information to supplement the mRS. Similarly, BI can reflect intricacies of physical function that are not represented in the mRS.[289] The aim of this study was to formulate and investigate the *multidimensional patterns* of post-stroke burden for individual mRS categories using AQoL-4D and BI.

7.3. Methods

All data that support the findings of this study are available from the corresponding author on reasonable request. Individual participant 3-month mRS, AQoL-4D and BI scores from A Very Early Rehabilitation Trial (AVERT) were obtained.[9] AVERT was a parallel-group, single-blind, RCT conducted in 54 acute stroke units across Australia, New Zealand, Malaysia, Singapore and the UK with a total sample of 2104 patients.[9] The aim of the trial was to compare the effectiveness of frequent, higher dose, very early mobilisation (sitting, standing or walking <24 hours post-stroke) with

usual care after stroke. The 3-month mRS scores were obtained from an unstructured interview (in person or by telephone) by a trained assessor who was not working in the study stroke unit and was blinded to treatment allocation. Detailed description of baseline demographics, and primary and secondary outcomes of AVERT have been previously published.[9] Ethics approval was obtained from institutional review boards at all sites.

7.3.1. Assessment of Quality of Life-4D and Barthel Index Outcomes

The descriptive system for the AQoL-4D was developed within the framework of the World Health Organization's definition of health-related quality of life.[268] AQoL-4D consists of 4 domains each measured on a scale from 0.00 (worst possible health state) to 1.00 (best possible health state) and containing 3 items: Independent Living – self-care, household tasks and mobility; Relationships – friendships, isolation and family role; Mental Health – sleeping, worrying and pain; Senses – seeing, hearing and communication.[268]

The BI measures a patient's performance in ten activities of daily living (ADL) to indicate level of functional independence, ability to mobilise in ADL and the need for assistance in care.[134,290] These items can be divided into two domains: Self-care – feeding, grooming, bathing, bowel and bladder care, and toilet use; and Mobility – ambulation, transfers and stair climbing. [127] Each of these functional categories may be scored from 0 to 1, 0 to 2, or 0 to 3, depending on the item. The summed score ranges from 0 to 20.

The use of composite outcome measures to increase trial efficiency and to capture numerous relevant patient factors have been previously investigated and shown to improve trial power compared to unadjusted mRS analysis alone.[291] In this study, the

inclusion of BI provides a clinician assessment of function and independence post-stroke that supplements the patient-reported Independent Living domains of AQoL-4D. To enable the comparison of individual domains within and between mRS categories, AQoL-4D and BI domain scores were scaled to 0% to 100% for BI and 0.00 to 1.00 for AQoL-4D. A percentage of the total achievable score within each BI domain was utilised e.g. a patient with a score of 6 out of 8 for Mobility received a score of 75%. Similarly, AQoL-4D domains are bounded by the range of the entire AQoL-4D scale, i.e. 0 (worst health state) to 1 (best health state).

7.3.2. Statistical Analyses

Kruskal-Wallis test was used to investigate differences in individual AQoL-4D and BI domains across mRS categories. Receiver Operator Characteristics (ROC) curve analyses were conducted to examine the ability of individual AQOL-4D and BI domain scores to discriminate between sequential mRS categories. A ROC curve area of 0.5 indicates no discrimination, whilst 1.0 indicates perfect discrimination.[269] The comparative speed of decline in AQoL-4D and BI domain scores across increasing mRS categories were investigated with random effects linear regression models using individual AQOL-4D and BI domains and mRS categories as independent variables, individual scores as dependent variables, domain-by-mRS category interaction terms and patients treated as random effects. Patient who died at 3 months (mRS 6) were not included in the analyses. Statistical analyses were performed with Stata (version 14.2; StataCorp LP, College Station, TX, USA).

7.4. Results

7.4.1. Study Population

From 2104 participants recruited in AVERT, 160 participants died at 3 months post-stroke. AQoL-4D, BI and mRS scores were available in 1870 stroke survivors (96.2%) at 3 months. The proportions within each category are as follows: mRS 0 (n=182), mRS 1 (n=393), mRS 2 (n=404), mRS 3 (n=448), mRS 4 (n=261) and mRS 5 (n=182). The selected sample had a mean (SD) age of 70.1 (12.7) years, median (IQR) baseline NIHSS of 7 (4-12), 61% were male and 88% were ischaemic strokes. AQoL-4D was completed by the patient (88%) or a proxy (12%) when communication or cognitive impairments precluded direct questioning of the patient. At 3 months, the median mRS was 3 (IQR 1-4).

7.4.2. Post-Stroke Burden across the mRS using AQoL-4D and BI Domains

Kruskal-Wallis test showed significant differences in scores for individual domains of AQoL-4D and BI across mRS categories ($p<0.0001$) (Figure 7.1). From the results of the ROC curve analysis, the scores for the Independent Living domain demonstrated excellent discrimination for mRS 3 vs. 4 and mRS 2 vs. 3 at 3 months post-stroke (Table 7.1). The scores for the Relationships domain could moderately discriminate mRS 2 vs. 3 and 3 vs. 4. Discrimination by other AQoL-4D domains were generally poor for the remaining comparisons. BI Self-care and Mobility domains demonstrated excellent discrimination between mRS 3 vs. 4 and 4 vs. 5 (Table 7.1).

Since the overall pattern of decline in scores of AQoL-4D Independent Living domain across the mRS was not matched by the comparable decline in scores of Mobility and

Self-care domains of BI, we undertook an additional exploratory analysis of individual items which contribute to the AQoL-4D Independent Living domain (household tasks, self-care and mobility items). A higher proportion of patients exhibiting poor performance was found in the household tasks items compared to mobility and self-care items of AQoL-4D's Independent Living domain across the mRS (Figure 7.2). A significant Independent Living individual items-by-mRS interaction effect was found ($p<0.0001$).

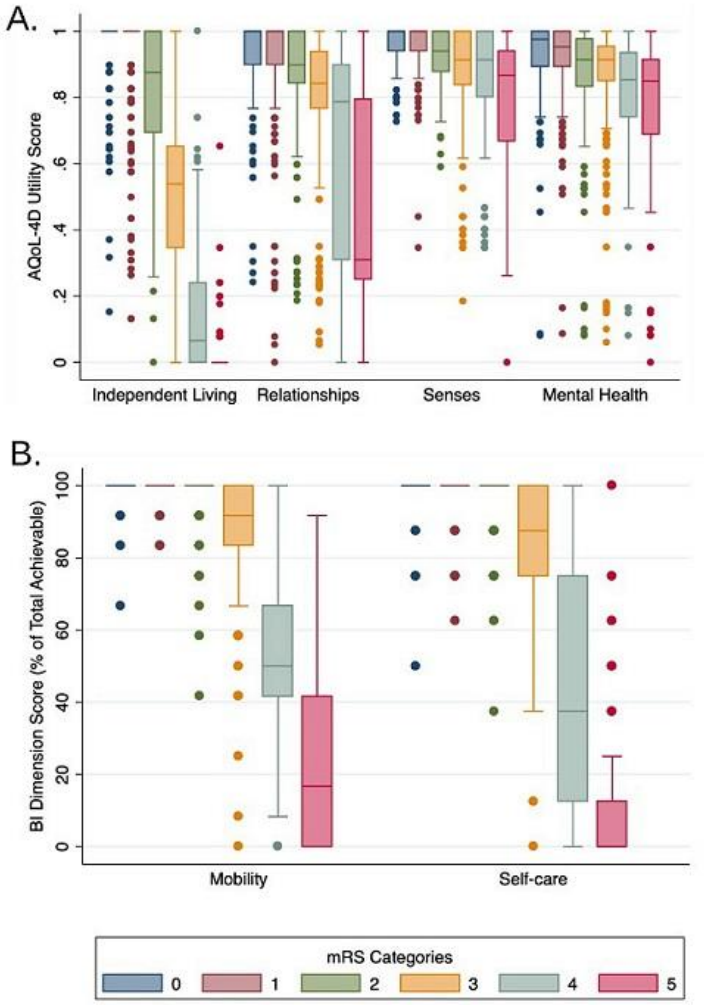


Figure 7.1. Distribution of AQoL-4D and BI domain scores stratified by mRS. BI: Barthel Index; AQoL-4D: Assessment of Quality of Life; mRS: modified Rankin Scale.

Table 7.1. AUC values (95% CI) for discrimination between sequential mRS categories for individual domains of post-stroke burden at 3 months.

	mRS 0 vs. 1	mRS 1 vs. 2	mRS 2 vs. 3	mRS 3 vs. 4	mRS 4 vs. 5
Independent Living	0.51 (0.47-0.55)	0.69 (0.66-0.73)	0.85 (0.83-0.88)	0.92 (0.89-0.94)	0.72 (0.68-0.75)
Relationships	0.54 (0.50-0.59)	0.60 (0.56-0.64)	0.65 (0.61-0.69)	0.65 (0.60-0.69)	0.63 (0.58-0.69)
Senses	0.53 (0.49-0.58)	0.61 (0.57-0.64)	0.59 (0.55-0.62)	0.53 (0.48-0.57)	0.60 (0.54-0.65)
Mental Health	0.53 (0.49-0.58)	0.62 (0.58-0.66)	0.51 (0.48-0.55)	0.59 (0.55-0.64)	0.56 (0.51-0.62)
Self-care	0.51 (0.49-0.53)	0.55 (0.53-0.57)	0.71 (0.68-0.74)	0.91 (0.89-0.93)	0.85 (0.81-0.88)
Mobility	0.50 (0.48-0.52)	0.54 (0.52-0.56)	0.71 (0.68-0.74)	0.89 (0.86-0.91)	0.83 (0.87-0.80)

AUC: area under the receiver operating characteristic curve; CI: confidence interval; mRS: modified Rankin Scale.

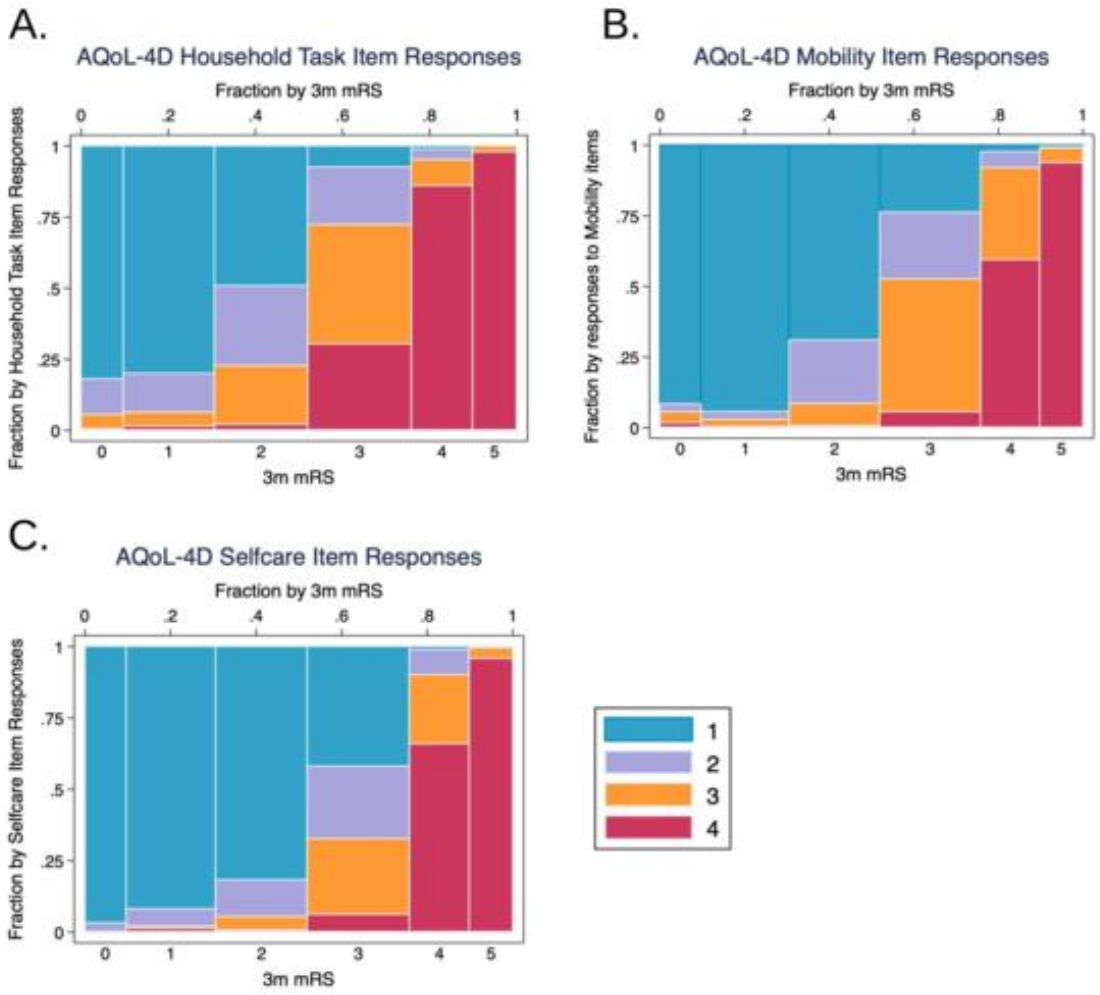


Figure 7.2: Level of performance in (A) Household Tasks, (B) Mobility and (C) Self-care items at 3 months. The performance level ranged from 1 (best) to 4 (poor).

7.4.3. Multidimensional Patterns of Post-Stroke Burden across the mRS

A significant AQoL-4D/BI domain-by-mRS interaction effect was found ($p<0.0001$), i.e. different domains displayed different patterns of change across mRS categories. The multidimensional patterns of post-stroke burden encompassing mean AQoL-4D and BI domain scores at 3 months for each mRS category are represented in a radar plot (Figure 7.3). The mean scores for all AQoL-4D and BI domains declined with increasing mRS categories. All independence-related domains (Independent Living,

Self-care, Mobility) displayed a large decline in scores across the mRS while Relationships was the most impacted non-independence-related domain (Figure 7.3). Based on the mean scores in patients clinically classified as symptomless (mRS 0) or with no significant disability (mRS 7.1), the most impacted domains post-stroke were Mental Health (sleep, worrying and pain) and Relationships (friendships, isolation and family role (Figure 7.4), For patients with an mRS 2-5, Independent Living was the most impacted domain.

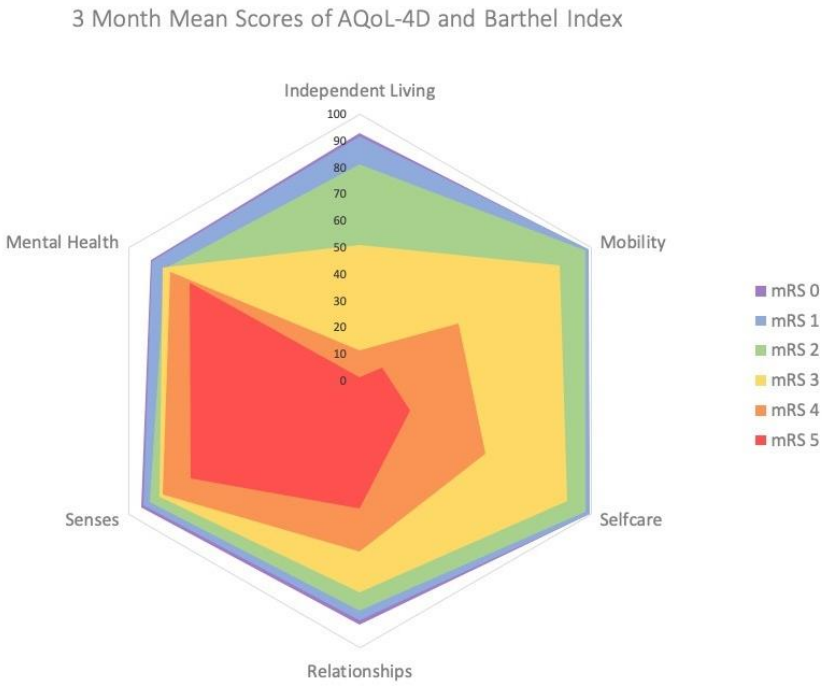


Figure 7.3. Mean AQoL-4D and BI domain scores per mRS category. The individual axes are scaled from 0 (poor performance) to 100 (best performance) and correspond to respective domains of post-stroke burden. The level of performance for all six domains is depicted using a coloured polygon whereby a smaller area of the polygon is indicative of poorer performance. mRS: modified Rankin Scale.

	Most Impacted Domain ←					→ Least Impacted Domain
mRS 0	Mental Health	Relationships	Independent Living	Senses	Mobility	Self-care
mRS 1	Relationships	Mental Health	Independent Living	Senses	Mobility	Self-care
mRS 2	Independent Living	Mental Health	Relationships	Senses	Mobility	Self-care
mRS 3	Independent Living	Relationships	Mental Health	Mobility	Senses	Self-care
mRS 4	Independent Living	Mobility	Self-care	Relationships	Mental Health	Senses
mRS 5	Independent Living	Mobility	Self-care	Relationships	Mental Health	Senses

Figure 7.4. The most impacted domain (lowest mean score) to the least impacted domain (highest mean score) per mRS category.

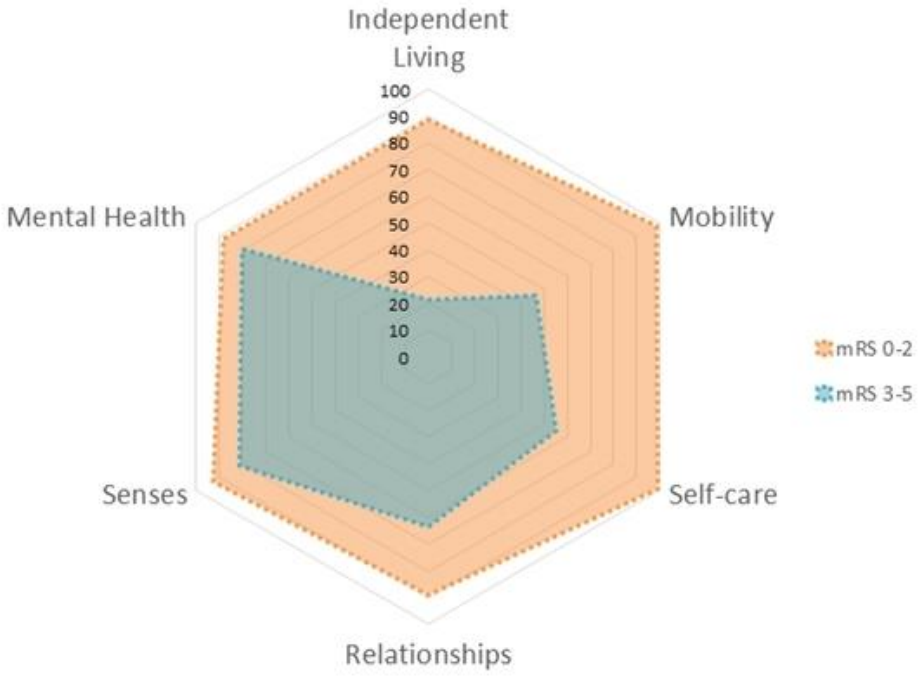


Figure 7.5. mRS 0–2 vs. 3–5 patterns of post-stroke burden.

The multidimensional *patterns of stroke burden* encompassing mean AQoL-4D and BI domain scores at 3 months for dichotomised mRS 0-2 and 3-5 outcomes are represented in Figure 7.5. Radar plots for mRS 0-1 vs. 2-5 and 0-3 vs. 4-5 are presented in Appendix F (Supplementary Figures F1 and F2).

7.5. Discussion

This study has systematically investigated both objective and patient-perceived measures of post-stroke burden at each level of the mRS. We found a significant domain-by-mRS interaction effect i.e. different post-stroke burden domains displayed different patterns of decline across mRS categories. This contextualises the outcome of an individual patient that receives a particular mRS score allowing better judgement about the way the mRS should be assessed and interpreted. Our findings support a recent observational cohort study, which examined patient-reported domains of health across levels of the mRS and found that as disability increased, symptoms in multiple disability domains increased in variable degrees.[123] The present study, however, supplements this work by systematically investigating the interactions between, and compensatory nature of, domains of stroke burden at various levels of the mRS in a sample nearly twice the size and across multiple centres internationally.

We observed a more pronounced decline in mean scores for the domains Mental Health and Relationships compared to Independent Living, Mobility and Self- in patients clinically classified as mRS 0 or 1. This is interesting given several studies have demonstrated difficulties with participation, social roles and reintegration, mental health, cognition, and feelings of isolation even after physical (basic self-care and mobility) post-stroke recovery appears complete.[123,292-296] Thus, caution should be taken when generalising patients with an outcome of mRS 0-1 as exhibiting "no symptoms at all" or "no significant disability" particularly when it is evident that nonmotor-related domains of stroke burden contribute to long-term outcomes post-stroke.

Patients with a mRS>3 displayed a deterioration in the domains Independent Living, Mobility, Self-care and Relationships. The decline in mean scores for Independent Living, however, were more marked than the decline in mean scores for Mobility and Self-care domains across most mRS categories. Two potential explanations exist: 1) since BI Mobility and Self-care are clinician-rated and AQoL-4D Independent Living is patient-rated, patients may perceive poorer outcomes compared to what is observed and scored by a clinician for the same health state; and/or 2) the more pronounced decline in the Independent Living domain is due to the deterioration in the household tasks item that is not explicitly measured by both domains of the BI. This complements a study which mapped the mRS to the ICF Core Set for stroke and found the strongest association between Self-care and Mobility, followed by Domestic life.[124] Unlike our study, they did not explicitly consider patient perspectives, therefore difficulties in domestic living, rather than mobility and self-care, may be more important to a stroke survivor. These findings ultimately highlight the multifaceted nature of post-stroke outcomes i.e. a stroke patient with hemiparesis (impairment) may have poor mobility (activity restriction), that has significant implications on social interactions, ability to return to work, driving, cooking, cleaning etc. (participation restrictions).[15,297] Further work is required to discriminate between mRS health states by capturing greater information about participation restrictions post-stroke.

Our findings displayed good discrimination between several sequential mRS comparisons in Independent Living, Mobility and Self-care domain scores, particularly mRS 3 vs. 4. This aligns with a recent suggestion that mRS 0–3, rather than 0–2, to be considered a "good" outcome category in moderate to severe ischaemic stroke.[298] The mechanisms of an intervention and type of patients enrolled

in clinical trials are important in determining the "right" statistical approach to analyse the mRS.[128,299]The recommended ordinal analysis approach still raises questions about the ability to interpret the treatment effects and its clinical importance.[137] Thus, by attempting to understand the clinical and societal impact of stroke at each categorical level of the mRS, we are able to understand the monotonic yet unequal impact of stroke burden across the mRS.

There are several strengths of this study. We included domains which do not contribute to the description of each mRS category i.e. Mental Health (sleeping, worrying and pain), Senses (seeing, hearing and communication) and Relationships (friendships, isolation and family role). These domains formed a more representative spectrum of post-stroke burden of that transcends the current description, which only measure the presence of symptoms, ability to perform activities of daily living, walking ability and self-care.[134] This is important since many trials describe the functional outcome of a patient in terms of the impairment, but it is arguably equally as important to describe domains that are more meaningful to a stroke survivor. Furthermore, the use of AQoL-4D, provided rich information about a patient's perception of their health and wellbeing post-stroke that is not explicitly measured by the mRS or BI. This is essential to progress towards a more patient-centric approach to interpreting outcome measures that accurately reflect long-term post-stroke outcomes.

There are also limitations. Clinical trial selection criteria limit the generalisability of the findings. This sample did not include patients admitted to a stroke unit >24h of stroke onset. However, the sample is broadly similar to population-based norms including a representative proportion of moderate to severe ischaemic and haemorrhagic strokes (44%).[300] The reasonably non-restrictive, large and

multinational sample may reduce external validity concerns about our conclusions. The ceiling effect of BI is a well-known limitation and may have influenced our findings. Additionally, since only two available measures from AVERT were used, our findings are naturally within the scope of AQoL-4D and BI, and only covers a fraction of the Comprehensive ICF Core Set for Stroke. Important post-stroke outcomes, e.g. cognition, were not included in this study. It is possible that other measures could provide additional domains to enrich the clinical picture of 3-month stroke outcome. The number of measures selected, the feasibility of data collection, and the interpretability of those results are important trade-offs worth considering for future stroke trials.

The findings from this study have given us a closer look at the multidimensional burden of stroke at each mRS category. There are three main implications that arise from this research. Firstly, the observed non-uniform change in different domains of post-stroke burden across the mRS needs to be considered when interpreting individual scores for future stroke trials. Secondly, the “hidden” domains of burden (i.e. mental health, relationships) in patients classified at symptomless may indicate areas of unmet need(s), which should be considered when developing targeted clinical strategies to support patient’s post-stroke. Lastly, the demonstrated importance of domestic living (e.g. driving, cooking, cleaning) to the decline in Independent Living across the mRS may suggest that it is an important measure of post-stroke burden from a patient's perspective. Domestic living is not explicitly captured by the mRS but may be a pragmatic measure to assess post-stroke burden.

Since the superiority of the UW-mRS or other measures over the mRS has not been established, the mRS is not likely to be challenged in its position. Therefore, the findings from our study may provide some important considerations about the

multidimensionality of post-stroke burden from a clinical and patient perspective, and the ability (and inability) of the mRS to measure those domains of post-stroke burden.

7.6. Chapter contribution to the thesis objective

In this chapter, a significant domain-by-mRS interaction effect was found, i.e., different post-stroke burden domains displayed different patterns of decline across mRS categories. We found that the mRS could capture changes in quality of life, albeit more for some domains (independent living and relationships) and less for others (senses and mental health).

From this study, we were able to represent patient-specific burden at each mRS category, which may facilitate a more patient-centric approach to articulating and interpretating each mRS score.

Chapter 8

General Discussion

8.1. Introduction

Early mobilisation (EM) is a valued component of acute stroke unit care. However, the *multifaceted* and *complex* nature of EM evidence and interventions, CPG development and reporting, and the characteristics of stroke patients and their outcomes challenges the use of population-level evidence to inform individualised decision-making.

The objective of this thesis was to investigate the barriers and enablers to the effective use of population-level evidence to inform individual patient-level clinical decision-making for EM post-stroke. This was achieved by exploring four key research questions centred upon the principles of evidence-based clinical decision-making, which emphasises the importance of evidence generation, evaluation, and interpretation to select an evidence-based decision that is tailored to the patient or setting. The four research questions were:

1. What are the utility and limitations of current CPGs as decision support tools that facilitate or inhibit the use of population-level evidence to support individual patient-level clinical decision-making for EM post-stroke?
2. How can we use an individual participant data meta-analysis of EM RCTs to investigate more precise estimates of subgroup and dose effects of EM?

3. Based on individual interpretation of the potential risks vs. benefits of EM to specific patients, what clinical and non-clinical factors do clinicians employ to tailor their decisions?
4. How well do the mRS and UW-mRS reflect post-stroke burden for an individual patient, and to what degree are they might be considered 'patient-centred'?

In this final chapter, the major findings and broader contributions from this thesis will be summarised. The limitations of this work will then be discussed, along with potential future directions that could further support the delivery of EBM within the context of EM post-stroke.

8.2. Summary of key findings and contributions

In this section, the key gaps that formed the aims within each study will be reiterated. A summary of the findings, along with key contributions, will be provided to contextualise the overall impact of the collective work from this thesis. Figure 8.1 incorporates a summary of the contributions and future directions into the overall research framework presented in Chapter 1.

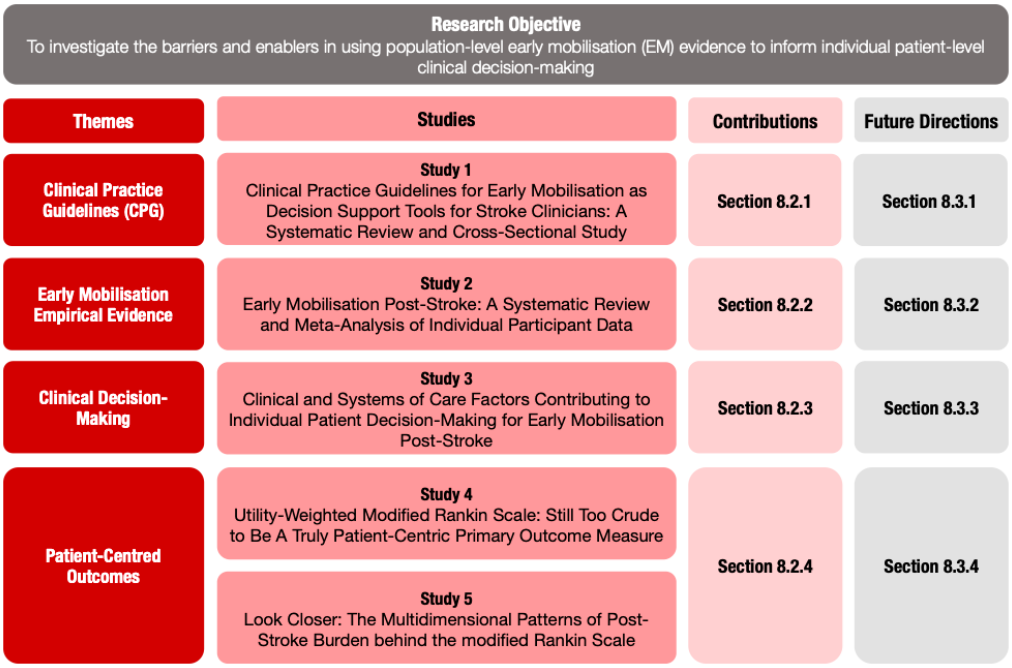


Figure 8.1. Thesis framework with contributions and future directions. This framework outlines the sections in Chapter 8 containing the associated contributions and future directions for Studies 1-5.

8.2.1 The utility of EM CPGs to support decision-making

The field of EM research is burdened by surprising and/or inconsistent findings that challenge the utility of EM evidence to inform individualised clinical decision-making. Clinically credible and applicable CPGs are needed to support and optimise clinical decisions. However, since the publication of AVERT, the utility of current EM CPGs to support individualized clinical decision-making was unknown. In Chapter 3 (Study 1), we used a two-pronged approach to investigate the utility and limitations of current CPGs in translating population-level evidence to support individualised clinical decision-making for EM post-stroke. The systematic review of EM CPGs demonstrated that almost all CPGs addressed the *Applicability to the Target User* and *Values and Preferences of the Target User* items of the AGREE-REX tool however, the *Evidence* and *Applicability to Patients* items were not met to the same degree. It is therefore not surprising that only 43% of interviewed expert

stroke clinicians indicated that CPGs provide decision support; with many often relying on clinical reasoning and individual interpretation of the evidence to select an *evidence-based* choice of action. CPGs were considered "too broad or vague" by 40% of clinicians. The investigation of areas of uncertainty in decision-making revealed four key areas of improvement for future development of EM CPGs: (1) granular descriptions of patient and stroke characteristics for appropriate tailoring of decisions, (2) clarity about when clinical flexibility is appropriate, (3) detailed description of the intervention dose, and (4) physical assessment criteria including safety parameters.

This study makes two major contributions. Firstly, by using a novel approach to integrate a newly developed appraisal tool (AGREE-REX)[87,89] with empirical investigations, this study provides an in-depth understanding about why previous appraisals using the AGREE II demonstrate that CPGs for rehabilitation post-stroke do not sufficiently address the "clinical applicability" domain.[236] Secondly, the exploration of opinions and use of current EM CPGs by expert stroke clinicians further empirically confirmed the lack of specificity, clinical applicability, and adaptability of current CPGs to effectively respond to the heterogeneous clinical stroke context, and this has provided a clear direction for improvement.

8.2.2 Patient-specific EM evidence generation

Findings from EM RCTs have been largely inconclusive and inconsistent, with the single, largest RCT (AVERT) showing harm. The recent Cochrane review found no difference between EM and usual care[14], but the analysis could not unpack the granularity of RCT EM data. In Chapter 4 (Study 2), we aimed to investigate more precise estimates of subgroup and dose effects of EM by conducting an IPD meta-analysis of EM RCTs. In this IPD meta-analysis, we found that EM significantly

reduced the odds of a favourable functional outcome (mRS 0-2) at 3 months post-stroke compared to usual care. This was observed despite a 3-hour difference in time to first mobilisation favouring EM; possibly suggesting that the often-larger dose of therapy received in the EM group had a greater influence on outcome at 3 months. The availability of IPD allowed pre-defined subgroup analyses, but no statistically significant interactions were found. The frequency, duration and intensity of the EM interventions were variably pre-specified inhibiting any exploration of dose effects of EM.

The findings from this IPD meta-analysis has two important contributions. Firstly, by using an IPD approach to unpack the granularity of RCT evidence, our findings provide additional support to the Cochrane review,[14] and several current CPG recommendations that advise against commencement of EM <24 h post-stroke.[10] Secondly, although the subgroup and dose effects of EM remains unclear, we have highlighted aspects of intervention development and monitoring that inhibit IPD meta-analyses in this research domain. These need to be addressed before future meta-analyses of EM RCTs can provide useful information based on the totality of *high-quality* evidence to support appropriate tailoring of evidence-based clinical decisions to the patient, and the development of CPGs.

8.2.3 Clinical and non-clinical factors influencing EM decision-making

The variability and ambiguity of EM evidence, and the heterogeneity of stroke patient populations requires complex decision-making. The way clinicians tailor their decisions based on their own interpretation of the potential risks vs. benefits of EM to individual patients is largely unknown. In Chapter 5 (Study 3), we aimed to identify the clinical and non-clinical factors that expert stroke clinicians employ to tailor their

decision to the clinical context, and to identify high-interest patient subgroups using an EM data visualisation tool. Based on our findings, stroke type, stroke severity, and medical stability were identified as important clinical decision-making factors by >80% of expert stroke clinicians. Inadequate staffing, equipment and low level of expertise were barriers to EM. The perceived characteristics of EM responders were those with mild or moderate stroke severity, ischaemic stroke, younger patients, and no or few co-morbidities. The perceived characteristics of EM non-responders included severe stroke severity, haemorrhagic stroke, older age, those with persistent vessel occlusion or high-grade stenosis, hemodynamic instability, numerous comorbidities and an altered state of consciousness. The subgroups that created uncertainty in EM decision-making were moderate stroke severity, older patients, and those with lacunar and posterior circulation infarcts.

This study has two noteworthy contributions. The findings have revealed unique, in-depth insights into patient and stroke characteristics involved in individual patient decision-making. This indicates that EM is not considered a benign intervention, rather it is one that may have benefits for the right group of patients yet may cause harm for the wrong group. We also identified bottlenecks in the delivery of care that might prevent the delivery of EM to the local setting. Identification of such factors can help develop structured and clinically applicable CPGs to ensure the clinicians (and patients) are well-informed to make the right decisions.

8.2.4 UW-mRS and mRS as "patient-centred" outcomes

From the words of a stroke survivor (Hannah, 2014), "healthcare providers often measure our recovery with functional status assessments and activity scales, but what these metrics cannot quantify is the tremendous loss felt"[301] In the domain of stroke,

despite the fact that the mRS in unable to capture all aspect of post-stroke disability, it remains the most widely used primary outcome measure in stroke clinical trials.[126] The UW-mRS is a newly proposed measure to improve the patient-centredness and overcome the limitations associated with the ordinal nature of the mRS. However, the construct validity of the UW-mRS as a patient-centred outcome measure had not been investigated. To explore how well the UW-mRS and mRS reflect post-stroke burden for an individual patient, and to what degree they might be considered 'patient-centred', we conducted two studies.

In Chapter 6 (Study 4), we used the large, international AVERT dataset to investigate the construct validity of the UW-mRS as a primary outcome measure by assessing variability in utility values within and between mRS categories, over time post-stroke, and by different derivation methods. This study uniquely derived AQoL-4D weighted UW-mRS values, which were consistent with prior studies using different derivation methods confirming the unequally ordered intervals of the mRS.[16,17,154,155,266] Based on AQoL-4D utility values, areas under the ROC curve varied between 0.54-0.87. Time post-stroke explained 42%-56% of variability in utility values in patients with no change in mRS between 3 and 12 months. The choice of the derivation method contributed to 25% of the variability in UW-mRS values. In summary, the high variability in individual patient-centred utility values between and within mRS categories, over time post-stroke, and using different derivation methods is not adequately reflected in the UW-mRS. This highlighted an obvious issue concerning the simple yet inadequate choice to average utility values per mRS category. Weighting the mRS in this way does not reflect the variability of individual utility values within each category, reduces the sensitivity of the UW-mRS, and disregards crucial information about quality of life.

This high level of variability in utility values present in the UW-mRS scores formed the rationale to *look closer* at individual domains of quality of life (AQoL-4D) and activities of daily living (Barthel Index) to demonstrate the *multidimensional patterns* of post-stroke burden across the mRS and UW-mRS. This was systematically investigated in Chapter 7 (Study 5). We found a significant domain-by-mRS interaction effect, i.e., different post-stroke burden domains displayed different patterns of decline across mRS categories. Mental Health and Relationship domains were most impact in patients with mRS 0-1, while patients with mRS>3 showed the greatest burden in Independent Living, Mobility and Self-care domains. However, the decline in mean scores for Independent Living were more marked than the decline in scores for Mobility and Self-care across most mRS categories. This may be due to the comparatively faster deterioration in the "household tasks" item across the mRS compared to the other items within the Independent Living domain. This work supplements the observational cohort study by Katzan et al.[123] by formally investigating the mRS-by-domain interaction effect, in a larger international sample.

The findings from both studies have several important contributions. Firstly, the threats to construct validity identified in Study 4 warrant caution when using UW-mRS as a 'patient-centred' primary outcome measure in future acute stroke trials. Secondly, the observed non-uniform change in different domains of post-stroke burden across the mRS supports findings from a recent observational cohort study that also demonstrated post-stroke symptoms in multiple domains increase to variable degrees at higher levels of disability.[123] They also found that physical function, satisfaction with social roles, and executive function were most affected domains. These non-uniform differences need to be considered when interpreting individual mRS scores. Thirdly, the "hidden" domains of burden (i.e. mental health,

relationships) in patients classified as symptomless may indicate areas of unmet need(s), which should be accounted for when developing targeted clinical strategies to support patients post-stroke. Lastly, the demonstrated importance of domestic living (e.g. driving, cooking, cleaning) to the decline in Independent Living across the mRS may suggest that it is an important measure of post-stroke burden from a patient's perspective that is not explicitly captured by the mRS. Overall, by using individual patient quality of life scores in several domains of post-stroke burden, we were able to represent patient-specific burden at each mRS category, which may facilitate a more patient-centric approach to articulating and interpretating each mRS score.

8.2.5 Contextualising of the collective contributions from this thesis

In summary, the work from this thesis has enriched the EM evidence base by generating patient-specific evidence; demonstrated important factors employed by expert stroke clinicians to tailor clinical decisions to the clinical context; investigated the patient-centredness of two potential primary outcome measures to ensure appropriate selection in empirical investigations and CPG development; and provided a clear direction for improvement in EM CPGs due to the observed lack of specificity, clinical applicability and adaptability of recommendations to effectively respond to the heterogeneous clinical stroke context.

As discussed in Chapter 2, the principles of evidence-based medicine (EBM) are fundamental to modern clinical decision-making and effective healthcare delivery.[1] We, therefore, contextualised the collective works from this thesis within the principles of evidence-based medicine (EBM) to explicitly establish the impact over and above the individual contributions of each study (Figure 8.2). Study 1 relates to

the use of *CPGs* to support evidence-based clinical decision-making; Study 2 relates to high quality *research evidence* generation; Study 3 relates to the use of *clinical expertise* to tailor decisions to the *clinical context*; Study 4 and 5 relate to the use of patient-centred outcomes to capture *patients' preferences and actions*. Together, by capturing these components of EBM as well as CPGs to support decision-making, we have *systematically* and *meaningfully* captured the barriers and enablers in using population-level evidence to inform patient-level clinical decision-making for EM post-stroke.

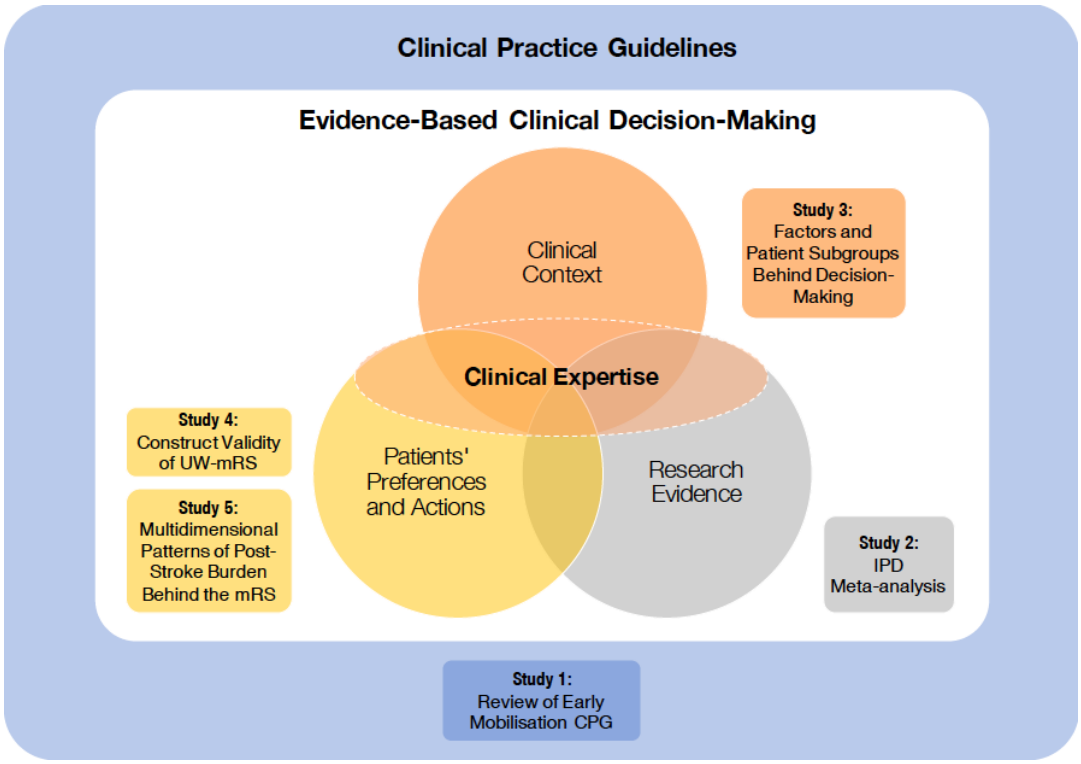


Figure 8.2. Representation of the contribution of collective works from this thesis embedded within the EBM model and CPGs. Each study from this thesis is mapped to a specific component of EBM or to CPGs to support evidence-based decision-making i.e. conducting the IPD meta-analysis contributed to the evidence component (grey) while investigating the factors important for decision-making contributed to understating how the clinical expertise of a stroke clinician is used to tailor decisions to the clinical context (orange).

8.3. Limitations and future directions

Since the limitations of each study are detailed within the relevant chapters, they will only be briefly reiterated. In this section we will focus on the limitations that are specific to the scope of this thesis, which will naturally introduce the future directions.

8.3.1 Expanding the scope in assessing the "clinical applicability" of clinical practice guideline

In Study 1, we focussed on the credibility and implementability of EM CPGs and did not assess the methodological quality potentially impacting the way in which CPGs are considered *clinically applicable* and *implementable*. A recent systematic review investigated the methodological quality appraisal of CPGs for all aspects of rehabilitation post-stroke and found only three out of 20 CPGs were rate high (>75%) across all domains of the AGREE II, and "recommended" for use.[236] While we established the scope of Study 1 to enable comprehensive investigation of the utility and limitations of current EM CPG, it could also be useful to investigate the internal validity and methodological quality of EM CPGs to understand how to improve the overall quality of current CPGs.

Similarly, as the focus of this study was to investigate the utility of CPGs as *decision-support tools for clinicians*, we purposefully excluded AGREE-REX items that were considered outside the scope of this aim. Focusing on the included items provided us with an opportunity to investigate in detail the utility and limitations of EM CPGs as decision support tools for the clinical end user. The items that were not explored in this thesis included the values and preferences of patients/populations, values and preferences of policy/decision makers, purpose, and local application and adoption.[89] Consideration of the values and perspectives of patients and funders, or alignment

across different viewpoints, may enhance the utilisation of CPGs.[87,88,237] Therefore, these domains are important to the clinical applicability of CPGs, and need to be investigated in the future. Future investigations could adopt our novel, two-pronged approach to ensure these normative standards of AGREE-REX are linked with empirical investigation of the different stakeholders (patients, policy or decision makers etc.). Together, this would establish a complete review of the quality and utility of EM CPGs.

8.3.2. Feasibility of exploring of dose and subgroup effects

In Study 2, the limitations were related to the inability to acquire IPD from all eligible trials; studies recording actual (not intended) time to first mobilisation; inconsistent development and monitoring of EM interventions; and the year of the original search.

Moreover, as the second epistemological principle of EBM suggests, the benefits and risks of an intervention should be based on a *sufficient totality of evidence* to inform clinical decisions at an individual patient-level.[2,25,302] However, in Study 2, four out of 10 eligible RCTs were unable to share data due to data sharing restrictions or because contact could not be made with the trialists. The inability to use the totality of EM evidence, denied the opportunity to potentially uncover subgroup effects of EM. The International Committee of Medical Journal Editors acknowledges an ethical obligation to responsibly share deidentified IPD generated by interventional clinical trials.[249] However, as experienced in this study, they are several challenges in conducting an IPD meta-analysis. These challenges are also well-documented in a recent scoping review.[303] Broadly, they are related to poor communication with study sponsors, frequent changes in data sharing policies of study sponsors, and the requirement for a diverse skillset related to research, administrative, statistical and

legal issues.[303] Since IPD meta-analyses are thought to be the "gold standard", [106,304-307] these challenges need to be overcome to take advantage of the more precise treatment estimates that an IPD can offer for future investigations. The benefits that an IPD meta-analysis can offer has important effects on the development of future CPGs and the design of new trials.[307-309]

Although the timing and general dose of EM were generally well defined in all RCTs, the specifics of frequency, duration and intensity of the EM interventions were variably pre-specified inhibiting exploration of dose effects of EM. This is unsurprising given Hoffman et al. showed that reporting of intervention procedures, materials, fidelity and tailoring was missing from 80% of non-pharmacological stroke interventions reviews.[310] Stroke recovery is exceedingly complex and progress towards standardised and effective treatments has been challenging. However, great advances have been made to overcome some of these challenges.[12,120,131,311-314] Most notably, the international collaborative effort by the Stroke Recovery and Rehabilitation Roundtable (SRRR) taskforce have generated consensus-based core recommendations on intervention development, monitoring and reporting,[12] which endorses the use of the Template for Intervention Description and Replication (TIDieR) checklist[244]. In addition, a stroke recovery trial development framework was more recently established.[248] With time, adoption of these recommendations, frameworks and checklists is likely to lead to high quality, comparable, and impactful research. The relatively recent momentum towards data sharing has obvious implications on the future trial design, conduct and reporting. Optimising trial design and improving reporting standards must progress in parallel with data sharing expectations if the opportunity IPD can offer is to be realised. Based on the collaborative international efforts,[12,120,131,307,311-314] trial and intervention design,

conduct, and reporting is hoped to be improve. With this in mind, and the possibility of precision in detecting treatment effects by using an IPD meta-analysis, reassessing the subgroup and dose effects of EM in the future is warranted.

8.3.3. Evidence-based clinical decision making

In Studies 1 and 3, the limitation relates to the Australian sample of expert stroke clinicians from metropolitan hospitals. Our findings are naturally within the scope of this demographic and may not be representative of the opinions of clinicians and stroke experts globally, or in regional centres in Australia. However, through the use of expert sampling[215] to obtain a multidisciplinary sample of clinical (physicians, physiotherapists and nurses/nurse practitioners) and academic experts in acute stroke and EM, we were provided with rich detail about EM decision-making. Early work in developing EM trial interventions identified physiotherapists and nurses as the main drivers of EM practices.[163] Physicians also play an important role in determining whether patients are medically fit to commence EM in cases of uncertainty. We therefore deliberately selected physicians, physiotherapists and nurses for these studies. It must be noted that other allied health professionals who may play an active role EM, such as occupational therapists, were not included. Future exploration of important EM decision-making factors, systems of care bottlenecks and perceived responders and non-responders to EM in more geographically and clinically diverse samples may validate the findings from Studies 1 and 3, and further enrich the evidence base.

In Study 3, we identified several clinical and systems of care factors that play an important role in decision-making at an individual patient-level. Individuals with haemorrhagic stroke, severe stroke, who are older, with poor function, who are

thrombolysed or receive endovascular clot retrieval were characteristics that concerned clinicians about starting EM. These decisions are not necessarily informed by, or aligned with, the evidence base.[9,191-193,256,257] However, the findings indicate that clinicians perceive EM to have benefits for the right group and harms to the wrong group. This work has identified important future research areas, and potentially important patient subgroups, that need to be investigated to substantiate and support clinical decision-making at an individual patient-level.

Some systems of care barriers were also identified that warrant consideration. This included adequate staffing and the availability of equipment to enable EM practices. This supports a previously published study within an Australian context in which poor teamwork, inadequate staffing, various organisational barriers, staff attitudes and beliefs, were identified barriers to implementing the AVERT EM intervention.[261] Our findings indicate there may be some systems of care factors that influence the administration of EM which need further investigation to ensure CPGs are implementable and applicable to the local setting. These factors will likely vary amongst countries and health providers, so future investigation in countries outside of Australia and different clinical settings (acute stroke units vs. general wards), could be of value for the future development of EM CPGs by different guideline development groups. The findings may help contextualise the evidence base to the clinical setting to ensure the recommendations are *clinically applicable* and *implementable* for the end user.[315]

Finally, in our study, the use of an interactive data visualisation tool to observe the expert stroke clinicians' interaction with a large dataset was advantageous in understanding the complex interaction of clinical factors, and to elicit high-interest subgroups that require a greater level of clinical judgment in the decision-making

process. Data collected in AVERT represents a large dataset with complex structures that are difficult to process using traditional methods. Using a data visualisation tool allowed the transition from aggregate data to subgroup or individual patient data to explore complex data in a more structured way. This may be useful in other areas that involve complex decision-making and the collection of complex clinical trial data.

8.3.4. Patient-centred outcomes and shared decision making

In Studies 4 and 5, the general limitation relates to the use of the AVERT dataset[9] to investigate the patient-centredness of the mRS and UW-mRS. Although the large, multinational sample is similar to population-based norms,[300] which reduces external validity concerns about our conclusions, the data comes from a clinical trial which naturally included restricted sampling. For example, the selection criteria of AVERT resulted in the exclusion of patients with previous disability (mRS >2) admitted to a stroke unit >24h from stroke onset. Therefore, further longitudinal exploration in less selective cohorts may be of value. A second limitation is that AQoL-4D and the mRS are both imperfect measures. If one does not agree with the other, it is difficult to know which one is 'correct', or if either are 'correct'. Nonetheless, the objective of the study was to meaningfully and empirically map these two scales.

The use of the UW-mRS is recommended by the Stroke Therapy Academic Industry Roundtable (STAIR) for "descriptive purposes" and to improve the focus on patient-centred outcomes in clinical trials.[262] However, based on the findings from Study 4, this measure may not be the best approach to incorporate patient preferences into trial results.[288] The demonstrated variability in UW-mRS values due to the choice of derivation method necessitates future consensus recommendations about the standards to derive the UW-mRS, and the appropriateness of its use as a primary

outcome measure. These recommendations are time-critical given the use of the UW-mRS as a co-primary outcome measure in an endovascular clot retrieval trial,[156] and a primary outcome measure in several ongoing trials.[316-318]

In Study 5, since only two available measures from AVERT were used, our findings are naturally limited in scope to the AQoL-4D and BI, and only cover some of the Comprehensive ICF Core Set for Stroke.[125] We advocate for the investigation of complex multidimensional patterns of post-stroke burden that may be hidden behind a single score. Composite scores have been previously proposed to provide outcomes that capture several factors of relevance to patients[291,319] however, the main limitation is in the interpretation of the score. By stratifying scores from other measures of post-stroke burden that are not extensively captured by the mRS, we were able to provide the interpretation of a given mRS score in terms of the multidimensional patterns of post-stroke burden. Post-stroke burden can include a combination domain not limited to motor, sensory, perceptual, language, cognition, and emotion.[15,123,124] It is important to note however, that the mRS is theoretically limited to stroke-related change, and BI and AQoL-4D may not be limited in the same way, which might explain some of the deviation of scores. Nontheless, future investigations of domains outside the scope of our study can enrich the clinical picture of 3-month stroke outcome, but the number of measures selected, and the interpretability of those results are important trade-offs worth considering.

We demonstrated the importance of domestic living (e.g. driving, cooking, cleaning) to the decline in Independent Living across the mRS, potentially suggesting a greater degree of relevance to the patient. Yet, domestic living is not explicitly captured by the mRS. While the mRS remains the recommended measure of global disability,[128-131] there are uncertainties concerning the best measure of "participation". Consensus

on the best measure of "participation" in development of the core measurement set by the SRRR taskforce could not be reached because potential measures generally exhibited inadequate measurement properties.[131] Based on this findings, domestic living could be an important and pragmatic domain in future post-stroke outcome measures that potentially warrants further investigation.

The need to employ pragmatic outcome metrics devoid of oversimplicity is critical.[15,125,132] It is important to emphasise that further research focusing on introducing patient-centred outcomes into clinical trials needs to appropriately and accurately reflect patient perspectives. This should facilitate better communication of relevant benefits, risks and costs associated with different clinical decisions.

Finally, the use of patient-centred outcomes is only one of several approaches to ensuring the values and preferences of patients are considered. The established scope for this thesis did not include investigations of shared decision-making in the delivery of EBM. There is often a disconnect between shared decision-making and EBM.[147] Efforts to adapt CPGs to facilitate shared decision-making are emerging.[320-322] Despite this, there are still major challenges in providing clinicians with efficient and practical tools to accommodate shared decisions.[1] Since EBM needs shared decision-making, and vice versa,[147] future investigations are necessary to facilitate and improve this connection.

8.3.5. Assessing barriers and enablers beyond early mobilisation in stroke

To conclude Section 8.3, based on the findings from five key studies, we have identified important future directions within each component of EBM and CPGs as illustrated in Figure 8.2. However, since evidence-based decision-making is a

contextual, continuous, and *evolving* process,[4] the future direction of this collective work may be the continual application of the research approach employed in this thesis to support EM post-stroke. This approach relies on conceptualising the process of decision-making that utilises population-level evidence to inform individual patient decision-making, embedding this concept into the EBM paradigm, and treating CPGs as tools to support the decision process. The distinguishing feature of this approach is the integration between the normative and empirical perspectives.

This research approach can also be applicable to other clinical areas that involve complex decision-making. As demonstrated, improving evidence-based decision-making requires several distinct, but related, components to be investigated using multiple methodological approaches. Investigations that aim to identify barriers and enablers in using evidence to inform individualised decision-making, or to develop strategies to improve clinical decision-making should systematically include all relevant EBM and CPG components. We have demonstrated how this might be achieved, and it could be a useful framework for future investigations with similar aims.

8.4 Conclusions

The work from this thesis has enriched the totality of the EM evidence base by generating patient-specific evidence; demonstrated important decision-making factors employed by expert stroke clinicians to tailor decisions to the clinical context; investigated the patient-centredness of two post-stroke outcome measures to promote appropriate selection in empirical investigations and CPG development; and provided a clear direction for improvement in EM CPGS due to the observed lack of specificity, clinical applicability and adaptability of recommendations to effectively respond to

the heterogeneous clinical stroke context. The collective work has substantially contributed to systematically identifying the barriers and enablers in using population-level evidence to inform patient-level clinical decision-making for early mobilisation post-stroke. This was achieved across all components of evidence-based medicine as well as providing an in-depth investigation of CPGs as decision support tools in this challenging clinical context.

References

1. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. The Lancet 2017;390:415-23.

2. Straus SE, Glasziou P, Richardson WS, Haynes RB. Evidence-Based Medicine E-Book: How to Practice and Teach EBM: Elsevier Health Sciences; 2018.

3. Group E-BMW. Evidence-based medicine. A new approach to teaching the practice of medicine. Jama 1992;268:2420.

4. Tiffen J, Corbridge SJ, Slimmer L. Enhancing Clinical Decision Making: Development of a Contiguous Definition and Conceptual Framework. Journal of Professional Nursing 2014;30:399-405.

5. Kent DM, Hayward RA. Limitations of Applying Summary Results of Clinical Trials to Individual PatientsThe Need for Risk Stratification. JAMA 2007;298:1209-12.

6. Permutt T. Do Covariates Change the Estimand? Statistics in Biopharmaceutical Research 2020;12:45-53.

7. Steinberg E, Greenfield S, Wolman DM, Mancher M, Graham R. Clinical practice guidelines we can trust: National Academies Press; 2011.

8. Bernhardt J, English C, Johnson L, Cumming TB. Early mobilization after stroke: early adoption but limited evidence. Stroke 2015;46:1141-6.

9. Bernhardt J, Langhorne P, Lindley RI, et al. Efficacy and safety of very early mobilisation within 24 h of stroke onset (AVERT): a randomised controlled trial. The Lancet 2015;386:46-55.

10. Bayley MT, Bowen A, English C, Teasell R, Eng JJ. Where to now? AVERT answered an important question, but raised many more. International Journal of Stroke 2017;12:683-6.

11. Luft AR, Kesselring J. Critique of A Very Early Rehabilitation Trial (AVERT). Stroke 2016;47:291-2.

12. Walker MF, Hoffmann TC, Brady MC, et al. Improving the development, monitoring and reporting of stroke rehabilitation research: Consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. International Journal of Stroke 2017;12:472-9.

13. Hoffmann TC, Walker MF. ‘TIDieR-ing up’ the reporting of interventions in stroke research: the importance of knowing what is in the ‘black box’. International Journal of Stroke 2015;10:657-8.

14. Langhorne P, Collier JM, Bate PJ, Thuy MN, Bernhardt J. Very early versus delayed mobilisation after stroke. Cochrane Database of Systematic Reviews 2018.

15. Geyh S, Cieza A, Schouten J, et al. ICF Core Sets for stroke. Journal of rehabilitation medicine 2004;36:135-41.

16. Hong K-S, Saver JL. Quantifying the value of stroke disability outcomes: WHO global burden of disease project disability weights for each level of the modified Rankin Scale. Stroke 2009;40:3828-33.

17. Chaisinanunkul N, Adeoye O, Lewis RJ, et al. Adopting a patient-centered approach to primary outcome analysis of acute stroke trials using a utility-weighted modified Rankin scale. Stroke 2015;46:2238-43.

18. Appelt KC, Milch KF, Handgraaf MJ, Weber EU. The decision making individual differences inventory and guidelines for the study of individual differences in judgment and decision-making research. Judgment and Decision making 2011.

19. Marcum JA. An integrated model of clinical reasoning: dual-process theory of cognition and metacognition. J Eval Clin Pract 2012;18:954-61.

20. Pelaccia T, Tardif J, Triby E, Charlin B. An analysis of clinical reasoning through a recent and comprehensive approach: the dual-process theory. Medical education online 2011;16:5890.

21. Arnott D. A taxonomy of decision biases. Monash University, School of Information Management and Systems, Caulfield 1998.

22. Mongtomery K. How doctors think: Clinical judgment and the practice of medicine: Oxford University Press; 2005.

23. Trimble M, Hamilton P. The thinking doctor: clinical decision making in contemporary medicine. Clin Med (Lond) 2016;16:343-6.

24. Levett-Jones T, Hoffman K, Dempsey J, et al. The ‘five rights’ of clinical reasoning: An educational model to enhance nursing students’ ability to identify and manage clinically ‘at risk’ patients. Nurse Education Today 2010;30:515-20.

25. Djulbegovic B, Guyatt GH, Ashcroft RE. Epistemologic inquiries in evidence-based medicine. Cancer Control 2009;16:158-68.

26. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. Bmj 1999;318:527-30.

27. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. How to practice and teach EBM. Edinburgh: Churchill Livingstone 2000.

28. Haynes RB, Devereaux PJ, Guyatt GH. Clinical expertise in the era of evidence-based medicine and patient choice. BMJ Evidence-Based Medicine 2002;7:36-8.

29. Woolf S, Schunemann HJ, Eccles MP, Grimshaw JM, Shekelle P. Developing clinical practice guidelines: types of evidence and outcomes; values and economics, synthesis, grading, and presentation and deriving recommendations. Implementation science : IS 2012;7:61.

30. Chalmers I. The Cochrane collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. Annals of the New York Academy of Sciences 1993;703:156-63; discussion 63-5.

31. Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. BMC Medicine 2010;8:24.

32. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383-94.

33. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ. What is "quality of evidence" and why is it important to clinicians? Bmj 2008;336:995-8.

34. Field MJ, Lohr KN. Clinical practice guidelines. Directions for a new program Washington: National Academy Pr 1990.

35. Bastian H, Glasziou P, Chalmers I. Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up? PLoS medicine 2010;7:e1000326.

36. Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? Bmj 2014;348:g3725.

37. Timmermans S, Mauck A. The promises and pitfalls of evidence-based medicine. Health Affairs 2005;24:18-28.

38. Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? Bmj 2014;348.

39. Arnott D. Cognitive biases and decision support systems development: a design science approach. Information Systems Journal 2006;16:55-78.

40. Holmström J, Ketokivi M, Hameri AP. Bridging practice and theory: A design science approach. Decision Sciences 2009;40:65-87.

41. March ST, Smith GF. Design and natural science research on information technology. Decision support systems 1995;15:251-66.

42. Huff S, McClurec R, Parkerc C, et al. Modeling guidelines for integration into clinical workflow. Medinfo 2004:174.

43. Chalmers I. The lethal consequences of failing to make full use of all relevant evidence about the effects of medical treatments: the importance of systematic reviews. Treating individuals−from randomised trials to personalised medicine Lancet 2007:37-58.

44. Makary MA, Daniel M. Medical error−the third leading cause of death in the US. Bmj 2016;353.

45. Djulbegovic B. A framework to bridge the gaps between evidence-based medicine, health outcomes, and improvement and implementation science. Journal of Oncology Practice 2014;10:200-2.

46. Pronovost PJ. Enhancing physicians ' use of clinical guidelines. Jama 2013;310:2501-2.

47. Proietti M, Nobili A, Raparelli V, Napoleone L, Mannucci PM, Lip GY. Adherence to antithrombotic therapy guidelines improves mortality among elderly patients with atrial fibrillation: insights from the REPOSI study. Clinical Research in Cardiology 2016;105:912-20.

48. Institute of Medicine Committee to Advise the Public Health Service on Clinical Practice G. In: Field MJ, Lohr KN, eds. Clinical Practice Guidelines: Directions for a New Program. Washington (DC): National Academies Press (US)

Copyright © National Academy of Sciences.; 1990.

49. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction−GRADE evidence profiles and summary of findings tables. Journal of clinical epidemiology 2011;64:383-94.

50. Djulbegovic B, Guyatt G. Evidence vs Consensus in Clinical Practice Guidelines. JAMA 2019;322:725-6.

51. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. Bmj 2008;336:924-6.

52. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. Journal of clinical epidemiology 2011;64:380-2.

53. Alonso-Coello P, Schünemann HJ, Moberg J, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. bmj 2016;353:i2016.

54. NHMRC. A guide to the development, implementation and evaluation of clinical practice guidelines. Australia: National Health and Medical Research Council; 1999.

55. Fretheim A, Schünemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 3. Group composition and consultation process. Health Research Policy and Systems 2006;4:15.

56. National Institute for H, Care E. NICE Process and Methods Guides. Developing NICE Guidelines: The Manual. London: National Institute for Health and Care Excellence (NICE)

unless otherwise stated. All rights reserved.; 2015.

57. NHMRC. Guidelines for Guidelines: Guideline Development Group. 2019.

58. Armstrong MJ, Rueda JD, Gronseth GS, Mullins CD. Framework for enhancing clinical practice guidelines through continuous patient engagement. Health Expect 2017;20:3-10.

59. Grilli R, Magrini N, Penna A, Mura G, Liberati A. Practice guidelines developed by specialty societies: the need for a critical appraisal. Lancet (London, England) 2000;355:103-6.

60. International Guideline Library. 2020, at https://g-i-n.net/library/international-guidelines-library/international-guidelines-library.)

61. Turner T, Misso M, Harris C, Green S. Development of evidence-based clinical practice guidelines (CPGs): comparing approaches. Implementation science 2008;3:45.

62. Ansari S, Rashidian A. Guidelines for guidelines: are they up to the task? A comparative assessment of clinical practice guideline development handbooks. PloS one 2012;7:e49864-e.

63. NZGG. Handbook for the Preparation of Explicit Evidence-Based Clinical Practice Guidelines. New Zealand Guidelines Group; 2001.

64. SIGN. SIGN 50: A guideline developers' handbook. Scottish Intercollegiate Guideline Network; 2004.

65. WHO. Guidelines for WHO Guidelines. Global Program on Evidence for Health Policy. World Health Organization; 2003.

66. Vernooij RWM, Sanabria AJ, Solà I, Alonso-Coello P, Martínez García L. Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. Implementation Science 2014;9:3.

67. Gagliardi AR, Brouwers MC, Palda VA, Lemieux-Charles L, Grimshaw JM. How can we improve guideline use? A conceptual framework of implementability. Implementation Science 2011;6:26.

68. Siering U, Eikermann M, Hausner E, Hoffmann-Eßer W, Neugebauer EA. Appraisal Tools for Clinical Practice Guidelines: A Systematic Review. PLOS ONE 2013;8:e82915.

69. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. Canadian Medical Association Journal 2010;182:E839-E42.

70. Consortium ANS. The AGREE II Instrument [My AGREE Plus]. 2013.

71. Ako-Arrey D, Brouwers M, Lavis J, Florez I, Spithoff K, Vukmirovic M. AGREE-HS: A Knowledge Translation Tool to Advise the Development, Reporting and Appraisal of Health Systems Guidance. 9th Annual Conference on the Science of Dissemination and Implementation; 2016: AcademyHealth.

72. Vernooij RW, Alonso-Coello P, Brouwers M, Martinez Garcia L, Panel C. Reporting items for updated clinical guidelines: checklist for the reporting of updated guidelines (CheckUp). PLoS medicine 2017;14:e1002207.

73. Brouwers MC, Kho ME, Browman GP, et al. Development of the AGREE II, part 1: performance, usefulness and areas for improvement. Cmaj 2010;182:1045-52.

74. Brouwers MC, Kho ME, Browman GP, et al. Development of the AGREE II, part 2: assessment of validity of items and tools to support application. Cmaj 2010;182:E472-8.

75. Armstrong JJ, Goldfarb AM, Instrum RS, MacDermid JC. Improvement evident but still necessary in clinical practice guideline quality: a systematic review. Journal of Clinical Epidemiology 2017;81:13-21.

76. Alonso-Coello P, Irfan A, Solà I, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. Quality and Safety in Health Care 2010;19:e58-e.

77. Kung J, Miller RR, Mackowiak PA. Failure of Clinical Practice Guidelines to Meet Institute of Medicine Standards: Two More Decades of Little, If Any, Progress. Archives of Internal Medicine 2012;172:1628-33.

78. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines?: The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. Jama 1999;281:1900-5.

79. Sabharwal S, Patel NK, Gauher S, Holloway I, Athanasiou T. High methodologic quality but poor applicability: assessment of the AAOS guidelines using the AGREE II instrument. Clin Orthop Relat Res 2014;472:1982-8.

80. Gagliardi AR, Brouwers MC. Do guidelines offer implementation advice to target users? A systematic review of guideline applicability. BMJ Open 2015;5:e007047.

81. Brosseau L, Rahman P, Toupin-April K, et al. A Systematic Critical Appraisal for Non-Pharmacological Management of Osteoarthritis Using the Appraisal of Guidelines Research and Evaluation II Instrument. PLOS ONE 2014;9:e82986.

82. Zupon A, Rothenberg C, Couturier K, et al. An appraisal of emergency medicine clinical practice guidelines: Do we agree? Int J Clin Pract 2019;73:e13289-e.

83. Knai C, Brusamento S, Legido-Quigley H, et al. Systematic review of the methodological quality of clinical guideline development for the management of chronic disease in Europe. Health Policy 2012;107:157-67.

84. Pagliari C, Grimshaw J, Eccles M. The potential influence of small group processes on guideline development. Journal of evaluation in clinical practice 2001;7:165-73.

85. Haynes B, Haines A. 10 Barriers and bridges to evidence-based clinical practice. Getting Research Findings into Practice 2008:115.

86. Oxman AD, Lavis JN, Fretheim A. Use of evidence in WHO recommendations. Lancet 2007;369:1883-9.

87. Brouwers MC, Spithoff K, Kerkvliet K, et al. Development and Validation of a Tool to Assess the Quality of Clinical Practice Guideline Recommendations. JAMA Network Open 2020;3:e205535-e.

88. Brouwers MC, Makarski J, Kastner M, Hayden L, Bhattacharyya O, the G-MRT. The Guideline Implementability Decision Excellence Model (GUIDE-M): a mixed methods approach to create an international resource to advance the practice guideline field. Implementation Science 2015;10:36.

89. Team A-RR. The Appraisal of Guidelines Research & Evaluation − Recommendation EXcellence (AGREE-REX). 2019. 2019.

90. McGlynn EA, Asch SM, Adams J, et al. The Quality of Health Care Delivered to Adults in the United States. New England Journal of Medicine 2003;348:2635-45.

91. Francke AL, Smit MC, de Veer AJE, Mistiaen P. Factors influencing the implementation of clinical guidelines for health care professionals: A systematic meta-review. BMC Medical Informatics and Decision Making 2008;8:38.

92. Kastner M, Bhattacharyya O, Hayden L, et al. Guideline uptake is influenced by six implementability domains for creating and communicating guidelines: a realist review. Journal of Clinical Epidemiology 2015;68:498-509.

93. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. The Milbank quarterly 2004;82:581-629.

94. Lugtenberg M, Zegers-van Schaick JM, Westert GP, Burgers JS. Why don't physicians adhere to guideline recommendations in practice? An analysis of barriers among Dutch general practitioners. Implementation Science 2009;4:54.

95. Green J, Britten N. Qualitative research and evidence based medicine. Bmj 1998;316:1230-2.

96. Pope C. Resisting Evidence: The Study of Evidence-Based Medicine as a Contemporary Social Movement. Health 2003;7:267-82.

97. KRAVITZ RL, DUAN N, BRASLOW J. Evidence-Based Medicine, Heterogeneity of Treatment Effects, and the Trouble with Averages. The Milbank Quarterly 2004;82:661-87.

98. Wyatt KD, Stuart LM, Brito JP, et al. Out of context: clinical practice guidelines and patients with multiple chronic conditions: a systematic review. Med Care 2014;52 Suppl 3:S92-s100.

99. Lugtenberg M, Burgers JS, Clancy C, Westert GP, Schneider EC. Current Guidelines Have Limited Applicability to Patients with Comorbid Conditions: A Systematic Analysis of Evidence-Based Guidelines. PLOS ONE 2011;6:e25987.

100. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines?: A framework for improvement. Jama 1999;282:1458-65.

101. McCartney M, Treadwell J, Maskrey N, Lehman R. Making evidence based medicine work for individual patients. Bmj 2016;353:i2452.

102. Michie S, Lester K. Words matter: increasing the implementation of clinical guidelines. BMJ Quality & Safety 2005;14:367-70.

103. Shekelle PG, Kravitz RL, Beart J, Marger M, Wang M, Lee M. Are nonspecific practice guidelines potentially harmful? A randomized comparison of the effect of nonspecific versus specific guidelines on physician decision making. Health Serv Res 2000;34:1429-48.

104. Weisz G, Cambrosio A, Keating P, Knaapen L, Schlich T, Tournay VJ. The emergence of clinical practice guidelines. The Milbank Quarterly 2007;85:691-727.

105. Murad MH. Clinical practice guidelines: a primer on development and dissemination. Mayo Clinic Proceedings; 2017: Elsevier. p. 423-33.

106. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. Bmj 2010;340:c221.

107. Trivedi M, Kern J, Marcee A, et al. Development and implementation of computerized clinical guidelines: barriers and solutions. Methods of information in medicine 2002;41:435-42.

108. Shekelle P, Woolf S, Grimshaw JM, Schünemann HJ, Eccles MP. Developing clinical practice guidelines: reviewing, reporting, and publishing guidelines; updating guidelines; and the emerging issues of enhancing guideline implementability and accounting for comorbid conditions in guideline development. Implementation Science 2012;7:62.

109. Morgano G, Amato L, Parmelli E, Moja L, Davoli M, Schünemann H. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. Recenti progressi in medicina 2017;108:402.

110. Feigin VL, Nichols E, Alam T, et al. Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Neurology 2019;18:459-80.

111. Aho K, Harmsen P, Hatano S, Marquardsen J, Smirnov V, Strasser T. Cerebrovascular disease in the community: results of a WHO collaborative study. Bulletin of the World Health Organization 1980;58:113.

112. Bonita R, Beaglehole R. Recovery of motor function after stroke. Stroke 1988;19:1497-500.

113. van der Worp HB, van Gijn J. Acute Ischemic Stroke. New England Journal of Medicine 2007;357:572-9.

114. Dabus G, Linfante I. The natural history of acute ischemic stroke due to intracranial large-vessel occlusion: what do we know? Techniques in vascular and interventional radiology 2012;15:2-4.

115. Gresham GE, Fitzpatrick TE, Wolf PA, McNamara PM, Kannel WB, Dawber TR. Residual disability in survivors of stroke−the Framingham study. New England Journal of Medicine 1975;293:954-6.

116. Phillips L, Whisnant J, Reagan T. Sudden death from stroke. Stroke 1977;8:392-5.

117. Economics DA. No postcode untouched: Stroke in Australia 2017. Australia: Stroke Foundation 2017.

118. Economics DA. The economic impact of stroke in Australia. National Stroke Foundation. 2013.

119. Gasparotti R, Grassi M, Mardighian D, et al. Perfusion CT in patients with acute ischemic stroke treated with intra-arterial thrombolysis: predictive value of infarct core size on clinical outcome. American journal of Neuroradiology 2009;30:722-7.

120. Bernhardt J, Hayward KS, Kwakkel G, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: the stroke recovery and rehabilitation roundtable taskforce. International Journal of Stroke 2017;12:444-50.

121. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. The Lancet 2011;377:1693-702.

122. Saver JL, Altman H. Relationship between neurologic deficit severity and final functional outcome shifts and strengthens during first hours after onset. Stroke 2012;43:1537-41.

123. Katzan IL, Thompson NR, Uchino K, Lapin B. The most affected health domains after ischemic stroke. Neurology 2018;90:e1364-e71.

124. Berzina G, Sveen U, Paanalahti M, Sunnerhagen KS. Analyzing the modified Rankin scale using concepts of the international classification of functioning, disability and health. Eur J Phys Rehabil Med 2016;52:203-13.

125. Organization WH. How to use the ICF: A practical manual for using the International Classification of Functioning, Disability and Health (ICF). Exposure draft for comment Geneva: WHO 2013;26.

126. Quinn T, Dawson J, Walters M, Lees K. Functional outcome measures in contemporary stroke trials. International Journal of Stroke 2009;4:200-5.

127. Sulter G, Steen C, Keyser JD. Use of the Barthel Index and Modified Rankin Scale in Acute Stroke Trials. Stroke 1999;30:1538-41.

128. Saver JL. Optimal End Points for Acute Stroke Therapy Trials. Stroke 2011;42:2356-62.

129. Salinas J, Sprinkhuizen SM, Ackerson T, et al. An International Standard Set of Patient-Centered Outcome Measures After Stroke. Stroke 2016;47:180-6.

130. Wallace SJ, Worrall L, Rose T, Le Dorze G. Measuring outcomes in aphasia research: A review of current practice and an agenda for standardisation. Aphasiology 2014;28:1364-84.

131. Kwakkel G, Lannin NA, Borschmann K, et al. Standardized measurement of sensorimotor recovery in stroke trials: Consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. Int J Stroke 2017;12:451-61.

132. Broderick JP, Adeoye O, Elm J. Evolution of the Modified Rankin Scale and Its Use in Future Stroke Trials. Stroke 2017;48:2007-12.

133. Lees KR, Bath PMW, Schellinger PD, et al. Contemporary Outcome Measures in Acute Stroke Research. Stroke 2012;43:1163-70.

134. Harrison JK, McArthur KS, Quinn TJ. Assessment scales in stroke: clinimetric and clinical considerations. Clinical interventions in aging 2013;8:201.

135. Dromerick AW, Edwards DF, Diringer MN. Sensitivity to changes in disability after stroke: A comparison of. Journal of rehabilitation research and development 2003;40:1-8.

136. Can We Improve the Statistical Analysis of Stroke Trials? Stroke 2007;38:1911-5.

137. Nunn A, Bath PM, Gray LJ. Analysis of the Modified Rankin Scale in Randomised Controlled Trials of Acute Ischaemic Stroke: A Systematic Review. Stroke Research and Treatment 2016;2016:7.

138. McArthur K, Fan Y, Pei Z, Quinn T. Optimising outcome assessment to improve quality and efficiency of stroke trials. Expert Review of Pharmacoeconomics & Outcomes Research 2014;14:101-11.

139. Mühlbacher AC, Juhnke C. Patient preferences versus physicians' judgement: does it make a difference in healthcare decision making? Appl Health Econ Health Policy 2013;11:163-80.

140. Devereaux P, Fahey T, Anderson DR, et al. Differences between perspectives of physicians and patients on anticoagulation in patients with atrial fibrillation: observational studyCommentary: Varied preferences reflect the reality of clinical practice. Bmj 2001;323:1218.

141. Redfern J, McKevitt C, Wolfe CDA. Risk management after stroke: The limits of a patient-centred approach. Health, Risk & Society 2006;8:123-41.

142. Frank L, Forsythe L, Ellis L, et al. Conceptual and practical foundations of patient engagement in research at the patient-centered outcomes research institute. Quality of Life Research 2015;24:1033-41.

143. Santana MJ, Manalili K, Jolley RJ, Zelinsky S, Quan H, Lu M. How to practice person-centred care: A conceptual framework. Health Expect 2018;21:429-40.

144. Guyatt G, Montori V, Devereaux PJ, Schünemann H, Bhandari M. Patients at the centre: in our practice, and in our use of language. Evidence Based Medicine 2004;9:6-7.

145. Press N. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.

146. Trujols J, Portella MJ, Iraurgi I, Campins MJ, Siñol N, de Los Cobos JP. Patient-reported outcome measures: are they patient-generated, patient-centred or patient-valued? J Ment Health 2013;22:555-62.

147. Hoffmann TC, Montori VM, Del Mar C. The Connection Between Evidence-Based Medicine and Shared Decision Making. JAMA 2014;312:1295-6.

148. Lawrence M, Kinn S. Defining and measuring patient-centred care: an example from a mixed-methods systematic review of the stroke literature. Health Expectations 2012;15:295-326.

149. Reeves M, Lisabeth L, Williams L, et al. Patient-Reported Outcome Measures (PROMs) for Acute Stroke: Rationale, Methods and Future Directions. Stroke 2018;49:1549-56.

150. Ali M, Fulton R, Quinn T, et al. How Well Do Standard Stroke Outcome Measures Reflect Quality of Life? Stroke 2013;44:3161-5.

151. Torrance GW. Utility approach to measuring health-related quality of life. Journal of Chronic Diseases 1987;40:593-600.

152. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. British Medical Bulletin 2010;96:5-21.

153. Albers GW, Goldstein LB, Hess DC, et al. Stroke Treatment Academic Industry Roundtable (STAIR) recommendations for maximizing the use of intravenous thrombolytics and expanding treatment options with intra-arterial and neuroprotective therapies. Stroke 2011:STROKEAHA. 111.618850.

154. Dijkland SA, Voormolen DC, Venema E, et al. Utility-Weighted Modified Rankin Scale as Primary Outcome in Stroke Trials: A Simulation Study. Stroke 2018;49:965-71.

155. Rivero-Arias O, Ouellet M, Gray A, Wolstenholme J, Rothwell PM, Luengo-Fernandez R. Mapping the modified Rankin scale (mRS) measurement into the generic EuroQol (EQ-5D) health outcome. Medical Decision Making 2010;30:341-54.

156. Nogueira RG, Jadhav AP, Haussen DC, et al. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. New England Journal of Medicine 2018;378:11-21.

157. Langhorne P, Ramachandra S. Organised inpatient (stroke unit) care for stroke: network meta-analysis. Cochrane Database of Systematic Reviews 2020.

158. Indredavik B, Bakke F, Solberg R, Rokseth R, Haaheim LL, Holme I. Benefit of a stroke unit: a randomized controlled trial. Stroke 1991;22:1026-31.

159. Ronning OM, Guldvog B. Stroke units versus general medical wards, I: twelve- and eighteen-month survival: a randomized, controlled trial. Stroke 1998;29:58-62.

160. Candelise L, Gattinoni M, Bersano A, Micieli G, Sterzi R, Morabito A. Stroke-unit care for acute stroke patients: an observational follow-up study. Lancet 2007;369:299-305.

161. Organised inpatient (stroke unit) care for stroke. Cochrane Database Syst Rev 2007:Cd000197.

162. Bernhardt J, English C, Johnson L, Cumming TB. Early Mobilization After Stroke. Stroke 2015;46:1141-6.

163. Bernhardt J, Dewey H, Thrift A, Donnan G. Inactive and Alone. Stroke 2004;35:1005-9.

164. Sheedy R, Kramer SF, Johnson L, et al. Acute Hospital Admission for Stroke Is Characterised by Inactivity. Stroke Research and Treatment 2020;2020:5879295.

165. Naito Y, Kamiya M, Morishima N, Ishikawa T. Association between out-of-bed mobilization and complications of immobility in acute phase of severe stroke: A retrospective observational study. Journal of Stroke and Cerebrovascular Diseases 2020;29:105112.

166. Langhorne P, Stott DJ, Robertson L, et al. Medical Complications After Stroke. Stroke 2000;31:1223-9.

167. Askim T, Bernhardt J, Salvesen Ø, Indredavik B. Physical activity early after stroke and its association to functional outcome 3 months later. Journal of stroke and cerebrovascular diseases 2014;23:e305-e12.

168. Murphy TH, Corbett D. Plasticity during stroke recovery: from synapse to behaviour. Nature Reviews Neuroscience 2009;10:861-72.

169. Zeiler SR, Krakauer JW. The interaction between training and plasticity in the poststroke brain. Curr Opin Neurol 2013;26:609-16.

170. Johansson BB. Brain Plasticity and Stroke Rehabilitation. Stroke 2000;31:223-30.

171. Krakauer JW, Carmichael ST, Corbett D, Wittenberg GF. Getting neurorehabilitation right: what can be learned from animal models? Neurorehabilitation and neural repair 2012;26:923-31.

172. Pekna M, Pekny M, Nilsson M. Modulation of neural plasticity as a basis for stroke rehabilitation. Stroke 2012;43:2819-28.

173. Anderson CS, Arima H, Lavados P, et al. Cluster-Randomized, Crossover Trial of Head Positioning in Acute Stroke. New England Journal of Medicine 2017;376:2437-47.

174. Carvalho LB, Kramer S, Borschmann K, Chambers B, Thijs V, Bernhardt J. Cerebral haemodynamics with head position changes post-ischaemic stroke: A systematic review and meta-analysis. Journal of Cerebral Blood Flow & Metabolism;0:0271678X20922457.

175. Skarin M, Bernhardt J, Sjöholm A, Nilsson M, Linden T. ‘Better Wear Out Sheets than Shoes’: A Survey of 202 Stroke Professionals' Early Mobilisation Practices and Concerns. International Journal of Stroke 2011;6:10-5.

176. Muhl L, Kulin J, Dagonnier M, et al. Mobilization after thrombolysis (rtPA) within 24 hours of acute stroke: what factors influence inclusion of patients in A Very Early Rehabilitation Trial (AVERT)? BMC Neurol 2014;14:163.

177. Collaboration SUT. Organised inpatient (stroke unit) care for stroke. The Cochrane database of systematic reviews 2002:CD000197.

178. Adams Jr HP, Brott TG, Crowell RM, et al. Guidelines for the management of patients with acute ischemic stroke. A statement for healthcare professionals from a

special writing group of the Stroke Council, American Heart Association. Circulation 1994;90:1588-601.

179. Xu T, Yu X, Ou S, Liu X, Yuan J, Chen Y. Efficacy and Safety of Very Early Mobilization in Patients with Acute Stroke: A Systematic Review and Meta-analysis. Scientific reports 2017;7:6550-.

180. Craig LE, Bernhardt J, Langhorne P, Wu O. Early Mobilization After Stroke. Stroke 2010;41:2632-6.

181. Bernhardt J, Churilov L, Ellery F, et al. Prespecified dose-response analysis for A Very Early Rehabilitation Trial (AVERT). Neurology 2016;86:2138-45.

182. Rothwell PM. AVERT: a major milestone in stroke research. Lancet 2015;386:7-9.

183. Herisson F, Godard S, Volteau C, Le Blanc E, Guillon B, Gaudron M. Early Sitting in Ischemic Stroke Patients (SEVEL): A Randomized Controlled Trial. PLoS One 2016;11:e0149466.

184. Barer D, Watkins C. Could upright posture be harmful in the early stages of stroke? Lancet 2015;386:1734.

185. Shen J, Huber M, Zhao EY, et al. Early rehabilitation aggravates brain damage after stroke via enhanced activation of nicotinamide adenine dinucleotide phosphate oxidase (NOX). Brain research 2016;1648:266-76.

186. Li F, Pendy JT, Jr., Ding JN, et al. Exercise rehabilitation immediately following ischemic stroke exacerbates inflammatory injury. Neurological research 2017;39:530-7.

187. Ward NS, Kitago T. Getting the right prescription for rehabilitation after stroke. AAN Enterprises; 2016.

188. Arias M, Smith LN. Early mobilization of acute stroke patients. Journal of Clinical Nursing 2007;16:282-8.

189. Sjöholm A, Skarin M, Linden T, Bernhardt J. Does evidence really matter? Professionals' opinions on the practice of early mobilization after stroke. Journal of multidisciplinary healthcare 2011;4:367.

190. Ha J, Churilov L, Linden T, Bernhardt J. Bed Rest or Mobilization after rt-PA? A Case-Crossover Study of Factors Influencing Clinical Decision Making in Stroke Services. International Journal of Stroke 2011;8:172-9.

191. Arnold SM, Dinkins M, Mooney LH, et al. Very early mobilization in stroke patients treated with intravenous recombinant tissue plasminogen activator. Journal of Stroke and Cerebrovascular Diseases 2015;24:1168-73.

192. Tubergen T, LaCroix H, Packard L, et al. Abstract WP485: Early Mobilization After 12 Hours Post Thrombolysis is Feasible and Safe in Minor Stroke Patients. Stroke 2019;50:AWP485-AWP.

193. Silver B, Hamid T, Khan M, et al. 12 versus 24 h bed rest after acute ischemic stroke thrombolysis: a preliminary experience. Journal of the Neurological Sciences 2020;409:116618.

194. Boddice G, Brauer S, Gustafsson L, Kenardy J, Hoffmann T. Clinical Guidelines for Stroke Management 2010. 2010.

195. Party ISW. National clinical guideline for stroke. London: Royal College of Physicians; 2012.

196. Committee ESOE, Committee EW. Guidelines for management of ischaemic stroke and transient ischaemic attack 2008. Cerebrovascular diseases 2008;25:457-507.

197. Broderick J, Connolly S, Feldmann E, et al. AHA/ASA Guideline. Stroke 2007.

198. Swain S, Turner C, Tyrrell P, Rudd A. Guidelines: Diagnosis and initial management of acute stroke and transient ischaemic attack: Summary of NICE guidance. BMJ: British Medical Journal 2008;337:291-3.

199. Indredavik B, Salvesen R, Næss H, Thorsvik D. Nasjonal retningslinje for behandling og rehabilitering ved hjerneslag. Oslo: Helsedirektoratet 2010:1-196.

200. Smith L. Management of Patients With Stroke: Rehabilitation, Prevention and Management of Complications, and Discharge Planning: a National Clinical Guideline. SIGN; 2010.

201. Ringelstein EB, Chamorro A, Kaste M, et al. European Stroke Organisation recommendations to establish a stroke unit and stroke center. Stroke 2013;44:828-40.

202. Lindsay M, Donnan G, Langhorne P, Davis S. on Behalf of the Global Stroke Quality and Guidelines Advisory Committee, the Global Stroke Guidelines Working Group, and the Global Stroke Quality Working Group of the World Stroke Organsation. Global stroke guidelines and action plan: a road map for quality stroke care http://www world-stroke org 2016.

203. Lindsay P, Furie KL, Davis SM, Donnan GA, Norrving B. World Stroke Organization global stroke services guidelines and action plan. International Journal of Stroke 2014;9:4-13.

204. Middleton S, Grimley R, Alexandrov AW. Triage, Treatment, and Transfer: Evidence-Based Clinical Practice Recommendations and Models of Nursing Care for the First 72 Hours of Admission to Hospital for Acute Stroke. Stroke 2015;46:e18-e25.

205. Bowen A, James M, Young G. Royal College of Physicians 2016 National clinical guideline for stroke. 2016: RCP.

206. Hebert D, Lindsay MP, McIntyre A, et al. Canadian stroke best practice recommendations: Stroke rehabilitation practice guidelines, update 2015. International Journal of Stroke 2016;11:459-84.

207. Winstein CJ, Stein J, Arena R, et al. Guidelines for adult stroke rehabilitation and recovery. Stroke 2016;47:e98-e169.

208. Clinical Guidelines for Stroke Management. Melbourne Australia: National Stroke Foundation; 2017.

209. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Annals of internal medicine 2007;147:573-7.

210. Rethnam V, Langhorne P, Churilov L, et al. Early mobilisation post-stroke: a systematic review and meta-analysis of individual participant data. Disability and Rehabilitation 2020:1-8.

211. Lindsay M, Norrving B, Furie K, Donnan G, Langhorne P, Davis S. Global stroke guidelines and action plan: a road map for quality stroke care. Geneva: World Stroke Organization 2016.

212. Owolabi MO, Thrift AG, Johnson W, et al. The Lancet Neurology Commission on Stroke: Evidence-Based Pragmatic Solutions to Reduce the Global Burden of Stroke (in press). The Lancet Neurology 2020.

213. Balk EM, Chung M, Chen ML, Chang LKW, Trikalinos TA. Data extraction from machine-translated versus original language randomized trial reports: a comparative study. Systematic Reviews 2013;2:97.

214. Jackson JL, Kuriyama A, Anton A, et al. The Accuracy of Google Translate for Abstracting Data From Non-English-Language Trials for Systematic Reviews. Ann Intern Med 2019;171:677-9.

215. Etikan I, Musa SA, Alkassim RS. Comparison of convenience sampling and purposive sampling. American journal of theoretical and applied statistics 2016;5:1-4.

216. Hsieh H-F, Shannon SE. Three approaches to qualitative content analysis. Qualitative health research 2005;15:1277-88.

217. Potter WJ, Levine-Donnerstein D. Rethinking validity and reliability in content analysis. 1999.

218. Group MoSRW. VA/DoD Clinical Practice Guideline for the Management of Stroke Rehabilitation. In: Affairs DoDDoV, ed.2019:1-170.

219. Zhang T, Zhao J, Li X, et al. Chinese Stroke Association guidelines for clinical management of cerebrovascular disorders: executive summary and 2019 update of clinical management of stroke rehabilitation. Stroke and Vascular Neurology 2020:svn-2019-000321.

220. Clinical Guidelines for Stroke Management. Melbourne, Australia: Stroke Foundation; 2019.

221. Pigretti SG, Alet MJ, Mamani CE, et al. [Consensus on acute ischemic stroke]. Medicina (B Aires) 2019;79 Suppl 2:1-46.

222. Boulanger J, Lindsay M, Gubitz G, et al. Canadian Stroke Best Practice Recommendations for Acute Stroke Management: Prehospital, Emergency Department, and Acute Inpatient Stroke Care, 6th Edition, Update 2018. International Journal of Stroke 2018;13:949-84.

223. Teasell R, Salbach NM, Foley N, et al. Canadian Stroke Best Practice Recommendations: Rehabilitation, Recovery, and Community Participation following Stroke. Part One: Rehabilitation and Recovery Following Stroke; 6th Edition Update 2019. Int J Stroke 2020:1747493019897843.

224. Ahmed N, Audebert H, Turc G, et al. Consensus statements and recommendations from the ESO-Karolinska Stroke Update Conference, Stockholm 11‐13 November 2018. European Stroke Journal 2019;4:307-17.

225. Ischemic stroke and TIA: Current Care Guidelines. Working group set up by the Duodecim of the Finnish Medical Association and the Finnish Neurological Association; 2016.

226. Guidelines for Prevention and Management of Stroke. Directorate General of Health Services Ministry of Health and Family Welfare, Government of India; 2019.

227. Ictus cerebrale: Linee guida italiane di prevenzione e trattamento: SPREAD - Stroke Prevention and Educational Awareness Diffusion; 2016.

228. Kim DY, Kim Y-H, Lee J, et al. Clinical practice guideline for stroke rehabilitation in Korea 2016. Brain & Neurorehabilitation 2017;10.

229. Herseninfarct en hersenbloeding. Dutch Society for Neurology; 2017.

230. National guidelines for treatment and rehabilitation of stroke. Oslo, Norway: Norwegian Directorate of Health; 2017.

231. AGUDA E, ISQUÉMICO AC. GUÍA DE PRÁCTICA CLÍNICA. 2018.

232. Stroke and transient ischaemic attack in over 16s: diagnosis and initial management. London: Nice 2019.; 2019.

233. Powers WJ, Rabinstein AA, Ackerson T, et al. 2018 Guidelines for the Early Management of Patients With Acute Ischemic Stroke: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. Stroke 2018;49:e46-e99.

234. Winstein CJ, Stein J, Arena R, et al. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2016;47:e98-e169.

235. Sall J, Eapen BC, Tran JE, Bowles AO, Bursaw A, Rodgers ME. The management of stroke rehabilitation: a synopsis of the 2019 US Department of Veterans Affairs and US Department of Defense clinical practice guideline. Annals of Internal Medicine 2019;171:916-24.

236. Jolliffe L, Lannin NA, Cadilhac DA, Hoffmann T. Systematic review of clinical practice guidelines to identify recommendations for rehabilitation after stroke and other acquired brain injuries. BMJ Open 2018;8:e018791.

237. Langlois EV, Montekio VB, Young T, Song K, Alcalde-Rabanal J, Tran N. Enhancing evidence informed policymaking in complex health systems: lessons from multi-site collaborative approaches. Health research policy and systems 2016;14:20.

238. Bernhardt J, Dewey H, Thrift A, Collier J, Donnan G. A very early rehabilitation trial for stroke (AVERT) phase II safety and feasibility. Stroke 2008;39:390-6.

239. Langhorne P, Stott D, Knight A, Bernhardt J, Barer D, Watkins C. Very early rehabilitation or intensive telemetry after stroke: a pilot randomised trial. Cerebrovascular Diseases 2010;29:352-60.

240. Poletto SR, Rebello LC, Valença MJM, et al. Early mobilization in ischemic stroke: a pilot randomized trial of safety and feasibility in a public hospital in Brazil. Cerebrovascular diseases extra 2015;5:31-40.

241. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS medicine 2009;6:e1000097.

242. Herisson F, Godard S, Volteau C, et al. Early sitting in ischemic stroke patients (SEVEL): a randomized controlled trial. PLoS One 2016;11:e0149466.

243. Tong Y, Cheng Z, Rajah GB, et al. High Intensity Physical Rehabilitation Later Than 24 h Post Stroke Is Beneficial in Patients: A Pilot Randomized Controlled Trial (RCT) Study in Mild to Moderate Ischemic Stroke. Frontiers in neurology 2019;10.

244. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ : British Medical Journal 2014;348:g1687.

245. Quinn T, Dawson J, Walters M. Dr John Rankin; his life, legacy and the 50th anniversary of the Rankin Stroke Scale. Scottish medical journal 2008;53:44-7.

246. Debray TPA, Moons KGM, Abo-Zaid GMA, Koffijberg H, Riley RD. Individual participant data meta-analysis for a binary outcome: one-stage or two-stage? PloS one 2013;8:e60650-e.

247. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. Statistics in Medicine 2017;36:855-75.

248. Bernhardt J, Hayward KS, Dancause N, et al. A stroke recovery trial development framework: Consensus-based core recommendations from the Second Stroke Recovery and Rehabilitation Roundtable. International Journal of Stroke 2019;14:792-802.

249. Taichman DB, Sahni P, Pinborg A, et al. Data sharing statements for clinical trials: a requirement of the International Committee of Medical Journal Editors. Jama 2017;317:2491-2.

250. Team RC. R: Language and Environment for Statistical Computing (3.4. 4). R Foundation for Statistical Computing. 2020.

251. Wickham H. ggplot2: elegant graphics for data analysis: springer; 2016.

252. Chang W, Cheng J, Allaire J, Xie Y, McPherson J. shiny: Web Application Framework for R. 1.5.0 ed: R package; 2020.

253. Johnson CO, Nguyen M, Roth GA, et al. Global, regional, and national burden of stroke, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Neurology 2019;18:439-58.

254. Indredavik B, Rohweder G, Naalsund E, Lydersen S. Medical Complications in a Comprehensive Stroke Unit and an Early Supported Discharge Service. Stroke 2008;39:414-20.

255. Boone M, Chillon J-M, Garcia P-Y, et al. NIHSS and acute complications after anterior and posterior circulation strokes. Ther Clin Risk Manag 2012;8:87-93.

256. Capo-Lugo CE, Askew RL, Muldoon K, et al. Longer Time Before Acute Rehabilitation Therapy Worsens Disability After Intracerebral Hemorrhage. Archives of Physical Medicine and Rehabilitation 2020;101:870-6.

257. Yen H-C, Jeng J-S, Chen W-S, et al. Early Mobilization of Mild-Moderate Intracerebral Hemorrhage Patients in a Stroke Center: A Randomized Controlled Trial. Neurorehabilitation and Neural Repair 2019;34:72-81.

258. Luker JA, Bernhardt J, Grimmer-Somers KA. Demographic and stroke-related factors as predictors of quality of acute stroke care provided by allied health professionals. Journal of multidisciplinary healthcare 2011;4:247-59.

259. Sewell K, Tse T, Harris E, et al. Pre-existing Comorbidity Burden and Patient Perceived Stroke Impact. International Journal of Stroke 2020:1747493020920838.

260. Ho E, Cheung SHC, Denton M, et al. The practice and predictors of early mobilization of patients post-acute admission to a specialized stroke center. Topics in Stroke Rehabilitation 2018;25:541-7.

261. Luker JA, Craig LE, Bennett L, et al. Implementing a complex rehabilitation intervention in a stroke trial: a qualitative process evaluation of AVERT. BMC Medical Research Methodology 2016;16:52.

262. Lees KR, Khatri P, Alexandrov AV, et al. Stroke treatment academic industry roundtable recommendations for individual data pooling analyses in stroke. Stroke 2016;47:2154-9.

263. Lenert L, Kaplan RM. Validity and interpretation of preference-based measures of health-related quality of life. Medical Care 2000:II138-II50.

264. Post PN, Stiggelbout AM, Wakker PP. The utility of health states after stroke: a systematic review of the literature. Stroke 2001;32:1425-9.

265. Hawthorne G, Richardson J, Day NA. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. Annals of Medicine 2001;33:358-70.

266. Ali M, MacIsaac R, Quinn TJ, et al. Dependency and health utilities in stroke: Data to inform cost-effectiveness analyses. European Stroke Journal 2017;2:70-6.

267. International classification of impairments, disabilities, and handicaps: a manual of classification relating to the consequences of disease, published in accordance with resolution WHA29. 35 of the Twenty-ninth World Health Assembly, May 1976. World Health Organization; 1980.

268. Hawthorne G, Richardson J, Osborne R. The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health-related quality of life. Quality of Life Research 1999;8:209-24.

269. Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian Journal of Internal Medicine 2013;4:627.

270. Rabe-Hesketh S, Skrondal A. Multilevel and longitudinal modeling using Stata: STATA press; 2008.

271. Richardson J, Iezzi A, Khan MA. Why do multi-attribute utility instruments produce different utilities: the relative importance of the descriptive systems, scale and 'micro-utility' effects. Quality of Life Research 2015;24:2045-53.

272. Richardson J, Khan MA, Iezzi A, Maxwell A. Comparing and explaining differences in the magnitude, content, and sensitivity of utilities predicted by the EQ-5D, SF-6D, HUI 3, 15D, QWB, and AQoL-8D multiattribute utility instruments. Medical Decision Making 2015;35:276-91.

273. Dorman PJ, Waddell F, Slattery J, Dennis M, Sandercock P. Is the EuroQol a valid measure of health-related quality of life after stroke? Stroke 1997;28:1876-82.

274. Hunger M, Sabariego C, Stollenwerk B, Cieza A, Leidl R. Validity, reliability and responsiveness of the EQ-5D in German stroke patients undergoing rehabilitation. Quality of Life Research 2012;21:1205-16.

275. Golicki D, Niewada M, Buczek J, et al. Validity of EQ-5D-5L in stroke. Quality of Life Research 2015;24:845-50.

276. Sturm JW, Osborne RH, Dewey HM, Donnan GA, Macdonell RA, Thrift AG. Brief comprehensive quality of life assessment after stroke: the assessment of quality of life instrument in the North East Melbourne Stroke Incidence Study (NEMESIS). Stroke 2002;33:2888-94.

277. Goyal M, Demchuk AM, Menon BK, et al. Randomized Assessment of Rapid Endovascular Treatment of Ischemic Stroke. New England Journal of Medicine 2015;372:1019-30.

278. Hacke W, Kaste M, Bluhmki E, et al. Thrombolysis with Alteplase 3 to 4.5 Hours after Acute Ischemic Stroke. New England Journal of Medicine 2008;359:1317-29.

279. Furlan A, Higashida R, Wechsler L, et al. Intra-arterial Prourokinase for Acute Ischemic StrokeThe PROACT II Study: A Randomized Controlled Trial. Jama 1999;282:2003-11.

280. Saver JL, Goyal M, Bonafe A, et al. Solitaire™ with the Intention for Thrombectomy as Primary Endovascular Treatment for Acute Ischemic Stroke (SWIFT PRIME) Trial: Protocol for a Randomized, Controlled, Multicenter Study Comparing the Solitaire Revascularization Device with IV tPA with IV tPA Alone in Acute Ischemic Stroke. International Journal of Stroke 2015;10:439-48.

281. Shuaib A, Lees KR, Lyden P, et al. NXY-059 for the Treatment of Acute Ischemic Stroke. New England Journal of Medicine 2007;357:562-71.

282. Jovin TG, Chamorro A, Cobo E, et al. Thrombectomy within 8 Hours after Symptom Onset in Ischemic Stroke. New England Journal of Medicine 2015;372:2296-306.

283. Berkhemer OA, Fransen PSS, Beumer D, et al. A Randomized Trial of Intraarterial Treatment for Acute Ischemic Stroke. New England Journal of Medicine 2014;372:11-20.

284. Hanley DF, Lane K, McBee N, et al. Thrombolytic removal of intraventricular haemorrhage in treatment of severe stroke: results of the randomised, multicentre, multiregion, placebo-controlled CLEAR III trial. The Lancet 2017;389:603-11.

285. Tissue Plasminogen Activator for Acute Ischemic Stroke. New England Journal of Medicine 1995;333:1581-8.

286. Anderson CS, Robinson T, Lindley RI, et al. Low-Dose versus Standard-Dose Intravenous Alteplase in Acute Ischemic Stroke. New England Journal of Medicine 2016;374:2313-23.

287. Qureshi AI, Palesch YY, Barsan WG, et al. Intensive Blood-Pressure Lowering in Patients with Acute Cerebral Hemorrhage. New England Journal of Medicine 2016;375:1033-43.

288. Rethnam V, Bernhardt J, Dewey H, et al. Utility-weighted modified Rankin Scale: Still too crude to be a truly patient-centric primary outcome measure? International Journal of Stroke 2019;0:1747493019830583.

289. Kwon S, Hartzema AG, Duncan PW, Min-Lai S. Disability Measures in Stroke. Stroke 2004;35:918-23.

290. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. Journal of clinical epidemiology 1989;42:703-9.

291. Makin SD, Doubal FN, Quinn TJ, Bath PM, Dennis MS, Wardlaw JM. The effect of different combinations of vascular, dependency and cognitive endpoints on the sample size required to detect a treatment effect in trials of treatments to improve outcome after lacunar and non-lacunar ischaemic stroke. European Stroke Journal 2018;3:66-73.

292. Kapoor A, Lanctôt KL, Bayley M, et al. "Good outcome" isn't good enough: cognitive impairment, depressive symptoms, and social restrictions in physically recovered stroke patients. Stroke 2017;48:1688-90.

293. Green TL, King KM. Functional and Psychosocial Outcomes 1 year after Mild Stroke. Journal of Stroke and Cerebrovascular Diseases 2010;19:10-6.

294. Edwards DF, Hahn M, Baum C, Dromerick AW. The impact of mild stroke on meaningful activity and life satisfaction. Journal of Stroke and Cerebrovascular Diseases 2006;15:151-7.

295. Carlsson G, Möller A, Blomstrand C. A qualitative study of the consequences of'hidden dysfunctions' one year after a mild stroke in persons<? 75 years. Disability and rehabilitation 2004;26:1373-80.

296. Verbraak ME, Hoeksma AF, Lindeboom R, Kwa VIH. Subtle Problems in Activities of Daily Living after a Transient Ischemic Attack or an Apparently Fully Recovered Non-disabling Stroke. Journal of Stroke and Cerebrovascular Diseases 2012;21:124-30.

297. Tse T, Binte Yusoff SZ, Churilov L, et al. Increased work and social engagement is associated with increased stroke specific quality of life in stroke survivors at 3 months and 12 months post-stroke: a longitudinal study of an Australian stroke cohort. Topics in sTroke rehabiliTaTion 2017;24:405-14.

298. Rangaraju S, Haussen D, Nogueira RG, Nahab F, Frankel M. Comparison of 3-Month Stroke Disability and Quality of Life across Modified Rankin Scale Categories. Interventional Neurology 2017;6:36-41.

299. Bath PMW, Lees KR, Schellinger PD, et al. Statistical Analysis of the Primary Outcome in Acute Stroke Trials. Stroke 2012;43:1171-8.

300. Bernhardt J, Raffelt A, Churilov L, et al. Exploring threats to generalisability in a large international rehabilitation trial (AVERT). BMJ open 2015;5:e008378-e.

301. Hannah D, Lindholm B, Maisch L. Certain Uncertainty. Circulation: Cardiovascular Quality and Outcomes 2014;7:968-9.

302. Hennekens CH, DeMets D. The Need for Large-Scale Randomized Evidence Without Undue Emphasis on Small Trials, Meta-analyses, or Subgroup Analyses. JAMA 2009;302:2361-2.

303. Ventresca M, Schünemann HJ, Macbeth F, et al. Obtaining and managing data sets for individual participant data meta-analysis: scoping review and practical guide. BMC Medical Research Methodology 2020;20:113.

304. Simmonds MC, Higginsa JP, Stewartb LA, Tierneyb JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. Clinical Trials 2005;2:209-17.

305. Smith CT, Marcucci M, Nolan SJ, et al. Individual participant data meta-analyses compared with meta-analyses based on aggregate data. Cochrane Database of Systematic Reviews 2016.

306. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. Evaluation & the health professions 2002;25:76-97.

307. Lees KR, Khatri P. Stroke Treatment Academic Industry Roundtable (STAIR) Recommendations for Individual Data Pooling Analyses in Stroke. Stroke; a journal of cerebral circulation 2016;47:2154.

308. Vale CL, Rydzewska LH, Rovers MM, Emberson JR, Gueyffier F, Stewart LA. Uptake of systematic reviews and meta-analyses based on individual participant data in clinical practice guidelines: descriptive study. bmj 2015;350.

309. Tierney JF, Pignon J-P, Gueffyier F, et al. How individual participant data meta-analyses have influenced trial design, conduct, and analysis. Journal of clinical epidemiology 2015;68:1325-35.

310. Hoffmann TC, Walker MF, Langhorne P, Eames S, Thomas E, Glasziou P. What's in a name? The challenge of describing interventions in systematic reviews: analysis of a random sample of reviews of non-pharmacological stroke interventions. BMJ Open 2015;5.

311. Bernhardt J, Borschmann K, Boyd L, et al. Moving rehabilitation research forward: developing consensus statements for rehabilitation and recovery research. International Journal of Stroke 2016;11:454-8.

312. Boyd LA, Hayward KS, Ward NS, et al. Biomarkers of stroke recovery: consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. International Journal of Stroke 2017;12:480-93.

313. Corbett D, Carmichael ST, Murphy TH, et al. Enhancing the alignment of the preclinical and clinical stroke recovery research pipeline: Consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable translational working group. Neurorehabilitation and Neural Repair 2017;31:699-707.

314. Saver JL, Warach S, Janis S, et al. Standardizing the Structure of Stroke Clinical and Epidemiologic Research Data. Stroke 2012;43:967-73.

315. McNutt RA, Livingston EH. Evidence-Based Medicine Requires Appropriate Clinical Context. JAMA 2010;303:454-5.

316. ENRICH: Early Minimally-invasive Removal of Intracerebral Hemorrhage (ICH) (Accessed 31 August, 2020, at https://clinicaltrials.gov/ct2/show/NCT02880878.)

317. Benefits of Stroke Treatment Delivered Using a Mobile Stroke Unit (BEST-MSU). (Accessed 31 August, 2020, at https://clinicaltrials.gov/ct2/show/NCT02190500.)

318. Blood Pressure After Endovascular Stroke Therapy-II (BEST-II). (Accessed 31 August, 2020, at https://clinicaltrials.gov/ct2/show/NCT04116112.)

319. Taylor-Rowan M, Wilson A, Dawson J, Quinn TJ. Functional Assessment for Acute Stroke Trials: Properties, Analysis, and Application. Frontiers in Neurology 2018;9.

320. van der Weijden T, Pieterse AH, Koelewijn-van Loon MS, et al. How can clinical practice guidelines be adapted to facilitate shared decision making? A qualitative key-informant study. BMJ quality & safety 2013;22:855-63.

321. Andrews JC, Schünemann HJ, Oxman AD, et al. GRADE guidelines: 15. Going from evidence to recommendation−determinants of a recommendation's direction and strength. Journal of clinical epidemiology 2013;66:726-35.

322. Guyatt GH, Oxman AD, Kunz R, et al. Going from evidence to recommendations. BMJ 2008;336:1049-51.

Appendices

Appendix A

Published Papers

Check for updates

International Journal of Stroke WSO

International Journal of Stroke
0(0) 1–10
© 2019 World Stroke Organization
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1747493019830583
journals.sagepub.com/home/wso
SAGE

Utility-weighted modified Rankin Scale: Still too crude to be a truly patient-centric primary outcome measure?

Venesha Rethnam[1], Julie Bernhardt[1], Helen Dewey[2], Marj Moodie[3], Hannah Johns[1], Lan Gao[3], Janice Collier[1], Fiona Ellery[1] and Leonid Churilov[1]; on behalf of AVERT Trial Collaboration Group

Abstract

Background: The utility-weighted modified Rankin Scale (UW-mRS) is an outcome measure recently proposed to improve statistical efficiency and interpretability of the mRS. Statistical properties of the UW-mRS have been well investigated, but construct validity has yet to be established.

Aims: To investigate the construct validity of the UW-mRS as a primary outcome measure by assessing variability in utility values within and between mRS categories, over time post-stroke, and by different derivation methods.

Methods: UW-mRS was derived using assessment of quality of life (AQoL-4D) and mRS scores at 3 and 12 months ($n = 2030$) from a large randomized controlled trial, A Very Early Rehabilitation Trial (AVERT). Receiver operator characteristic (ROC) analysis of AQoL-4D was conducted to differentiate between sequential mRS categories. Intraclass correlation was used to explore variability in utility values over time post-stroke, UW-mRS values, and derivation methods from multiple studies.

Results: UW-mRS values for mRS categories 0–6 at three months were 0.80, 0.78, 0.63, 0.37, 0.11, 0.03, and 0. Based on AQoL-4D utility values, areas under the ROC curve varied from 0.54 to 0.87. Time post-stroke explained 42%–56% of variability in AQoL-4D utility values in patients with no change in mRS between 3 and 12 months. The choice of the derivation method contributed to 25% of the variability in UW-mRS values.

Conclusions: The high variability in utility values between and within mRS categories, over time post-stroke, and using different derivation methods is not adequately reflected in the UW-mRS. These threats to construct validity warrant caution when using UW-mRS as a primary outcome measure.

Trial registration: Australian New Zealand Clinical Trials Registry (ACTRN12606000185561).

Keywords

Utility-weighted modified Rankin scale, primary outcome measure, stroke, assessment of quality of life, AQoL-4D, AVERT

Received: 17 September 2018; accepted: 4 December 2018

Introduction

The recently proposed UW-mRS is derived by translating ordinal mRS categories into a utility-weighted scale.[1] It is an emerging outcome measure which aims to reflect patient perception of quality of life to improve interpretability of the mRS.[2,3] The UW-mRS also claims to improve statistical efficiency by providing a continuous alternative to either an ordinal or dichotomized analysis of the mRS.[3]

[1]NHMRC Centre for Research Excellence in Stroke Rehabilitation and Brain Recovery, Florey Institute of Neuroscience and Mental Health, University of Melbourne, Melbourne, Australia
[2]Eastern Health and Eastern Health Clinical School, Monash University, Clayton, Australia
[3]Deakin Health Economics, Deakin University, Burwood, Australia

Corresponding author:
Venesha Rethnam, Florey Institute of Neuroscience and Mental Health – Austin Campus, 245 Burgundy Street Heidelberg, Victoria 3084, Australia.
Email: venesha.rethnam@florey.edu.au

The development of a UW-mRS has been advocated by several associations including the Stroke Therapy Academic Industry Roundtable.[4] Utility, centered on health-related quality of life, is a widely accepted measure of a patient's preference for a specific health outcome.[5,6] UW-mRS values are obtained by assigning a mean utility weight to each mRS category based on patient, clinician, or societal valuation of various disability health states.[7]

Since its introduction, research efforts have focused on the statistical efficiency of the UW-mRS to detect treatment effects in stroke intervention trials.[2,3] These investigations showed similar efficiency compared to the ordinal mRS, and greater efficiency compared to the dichotomized mRS. While prior investigations have focused on statistical properties, construct validity of the UW-mRS has yet to be established.

It is unclear whether a mean utility weight per mRS category sufficiently reflects the variability of utility values within mRS categories, and whether the difference in utility values adequately discriminates between mRS categories. Similarly, it is unclear whether utility weights reflect changes in patient perspectives over time post-stroke, and consequently, the long-term outcomes of an intervention to a stroke survivor. Lastly, uncertainty concerning the most appropriate utility derivation method for a stroke population persists challenging the use of the UW-mRS as a primary outcome measure in stroke trials.[8,9]

Aims

The overall objective was to investigate the construct validity of the UW-mRS as a primary outcome measure for stroke trials. The aims of this study were:

1. Derive AQoL-4D predicted UW-mRS values per mRS category from the AVERT population at 3- and 12-months post-stroke;
2. Investigate between- and within-patient variability of individual utility values;
3. Compare the derived UW-mRS values and utility derivation method with published UW-mRS studies.

Methods

Conceptual model for assessing threats to UW-mRS validity

We conceptualized the potential validity threats as belonging to the following two set of circumstances. Firstly, there is the potential inability of UW-mRS values to reflect the observed variability in utility values between and within individual mRS categories. Under this scenario, considerable mismatch between the UW-mRS value and the true utility value may exist. Secondly, UW-mRS values can be derived using various elicitation methods, country-specific value sets, multi-attribute utility instruments, and stroke population datasets.[10,11] Obtaining a utility involves defining and valuing health states of interest to develop a *value set*. Utility values can be elicited using two methods: direct (e.g. time trade-off method, standard gambling, visual analogue scale) or indirect (e.g. patient's response to a multi-attribute utility instrument: AQoL-4D, European Quality of Life Scale [EQ-5D-3L] and Health Utilities Index).[10]

As many stroke trials use multi-attribute utility instruments to measure quality of life in parallel to the mRS, a UW-mRS can be derived using a stroke population dataset containing responses to both measures.[1–3,7,11] To produce UW-mRS values, utility values are stratified by mRS category and then averaged. We, therefore, will refer to individual patient AQoL-4D utility values as *utility values* and the mean of AQoL-4D utility values per mRS category as *UW-mRS values* in this study.

Study population

We used individual patient data mRS and AQoL-4D data from AVERT.[12] AVERT was a parallel-group, single-blind, randomized controlled trial conducted at 56 acute stroke units in five jurisdictions: Australia, New Zealand, Malaysia, Singapore, and the UK. The aim was to compare the effectiveness of frequent, higher dose, very early mobilization (sitting, standing or walking <24 h post-stroke onset) with usual care after stroke. mRS data were obtained face-to-face by a trained and accredited blinded assessor. AQoL-4D was completed by the patient (88%) or a proxy (12%) when communication or cognitive impairments precluded direct questioning of the patient. The primary and secondary outcomes of the trial have been previously reported.[12] Ethics approval was obtained from institutional review boards at all sites.

AQoL-4D derived utilities

The UW-mRS values per mRS category were determined using AQoL-4D utility values and mRS scores from AVERT at 3 months and 12 months. The descriptive system for AQoL-4D was developed within the framework of the World Health Organisation's classification of impairment, disability, and handicaps.[13] AQoL-4D consists of four dimensions and three levels: independent living – self-care, household tasks, and mobility; relationships – friendships, isolation, and family role; mental health – sleeping, worrying, and pain; senses – seeing, hearing, and communication.

The utility scoring algorithm is based on a stratified sample of the Australian population using the time trade-off method. This algorithm weights the items and applies a multiplicative function to achieve an index producing a life-death utility scale. The scale is bounded by the following points: 1.00 (best possible health state), 0.00 (death equivalent) and −0.04 (worse-than-death states).[14] AVERT participants who died before their 3 or 12 month follow-up were assigned a utility value of 0.

Statistical analyses

To analyze the variability in utility values between individual mRS categories, we conducted the following tests. A one-way ANOVA was performed to estimate differences between mean utility values across mRS categories at 3- and 12-months post-stroke, respectively. ROC curves were generated to differentiate between sequential mRS categories based on utility values. An ROC curve area of 0.5 indicates no discrimination, while 1.0 indicates perfect discrimination.[15] Finally, a misclassification matrix was used to classify each patient into a "matched" mRS category based on the closest UW-mRS value to their observed individual AQoL-4D utility value. The formal agreement between the "actual" and "matched" mRS was estimated using weighted Cohen's Kappa with quadratic weights.

To explore the variability in utility values within individual patients over time post-stroke, intraclass correlations (ICC) were calculated using a random effects linear regression model using patients (random effect), time, i.e. 3 vs. 12 months (input), and utility values (output). The ICC obtained from this model indicates the degree of variability that can be explained by within- and between-patient differences. The ICC is 0 when there are no between-patient differences and increases up to 1 when the between-patient variance increases relative to within-patient variance.[16] Variability in utility values over time post-stroke in patients who did not change in mRS score from 3 months to 12 months was calculated using the previously described random effects linear regression model. If the UW-mRS is a valid representation of the mRS, particularly in patients with the same mRS score at 3 and 12 months and covariates such as age, stroke severity, and co-morbidities are accounted for, similar utility values would be expected, and any observed difference should be attributed to time.

To explore the variability in AVERT UW-mRS values and previously published studies, all UW-mRS values from 9 UW-mRS sets at 3 months, and 2 UW-mRS sets at 12 months were applied to individual patients from the AVERT dataset.[1–3,7,11] ICCs were calculated using a variance component model implemented as random effects linear regression using age and baseline stroke severity (inputs), utility values (output), and patients (random effects). Statistical

Table 1. Characteristics of 2030 AVERT stroke patients.

Variables	Total, $n = 2030$
Recruitment region	
Australia and New Zealand	1216 (60%)
Asia	240 (12%)
UK	574 (28%)
Age (years)	
<60	593 (29%)
60–80	906 (45%)
>80	531 (26%)
Sex	
Male	1238 (61%)
Female	792 (39%)
Risk factors	
Hypertension	1410 (68%)
Ischemic heart disease	483 (23%)
Hypercholesterolemia	832 (40%)
Diabetes mellitus	459 (22%)
Atrial fibrillation	465 (22%)
Premorbid modified Rankin scale	
0	1528 (75%)
1	292 (15%)
2	210 (10%)
Stroke history	
NIHSS	7 (4–12)
Ischemic stroke	1784 (88%)
Intracerebral hemorrhage	246 (12%)
rtPA treatment	503 (24%)

Note: Data are *n* (%) or median (IQR).
NIHSS: National Institutes of Health Stroke Scale; rtPA: recombinant tissue plasminogen.

analyses were performed with Stata (version 14.2; StataCorp LP, College Station, TX, USA).

Results

Study population

All 2104 participants from AVERT were included in our analysis. The utility values and mRS scores at three months were available in 2030 participants (96.5%) including 160 participants who died before follow-up. Utility values and mRS scores at 12 months were available in 2017 participants (95.7%), including 257 participants who died before follow-up. The study population had a mean (SD) age of 70.7 (12.7) years and 61% of participants were male (Table 1).

AVERT UW-mRS values

The UW-mRS values, i.e. mean AQoL-4D utility values (SD), derived for mRS categories 0–6 at three months include: 0.80 (0.22), 0.78 (0.21), 0.63 (0.23), 0.37 (0.21), 0.11 (0.12), 0.03 (0.06), and 0 (0), respectively (one-way ANOVA $p < 0.0001$). The UW-mRS values for mRS categories 0–6 at 12 months include 0.85 (0.19), 0.77 (0.21), 0.63 (0.24), 0.34 (0.21), 0.11 (0.12), 0.03 (0.07), and 0 (0), respectively (one-way ANOVA $p < 0.0001$).

Utility variability within and between mRS

We observed considerable variation in utility values within each mRS category at both 3 months and 12 months (Figure 1). The misclassification matrix showed excellent statistical agreement between the "actual" and "matched" mRS categories at 3 months (kappa, 0.81; CI, 0.79–0.84) and 12 months (kappa, 0.86; CI, 0.85–0.87) (Figure 2). However, based on the AQoL-4D utility values, good discrimination between sequential pairs of mRS categories was only observed for mRS 3 vs. 4 at 3 months (area under the curve [AUC], 0.87; 95% CI, 0.84–0.90) and 12 months (AUC, 0.83; 95% CI, 0.80–0.87) (Table 2). In contrast, utility was less discriminant for mRS 1 vs. 2 and 4 vs. 5 which reflect the matching inconsistencies in the misclassification matrix. For example, an individual with a mRS of 2 at 3 months could be classified into mRS 0 (29%), mRS 1 (13%), mRS 2 (28%), and mRS 3 (24%) based on the closest UW-mRS value. Similarly, an "actual" mRS of 4 and 5 displayed a mirrored inability to discriminate between "matched" mRS 4, 5, and 6.

Utility variability over time post-stroke

We observed substantial within-patient variability in utility values for patients with no change in mRS between 3- and 12-months post-stroke ($n = 1189$) (Figure 3). The degree of variability that could be explained by

Figure 1. Distribution of AQoL utility values across mRS categories at 3 and 12 months.

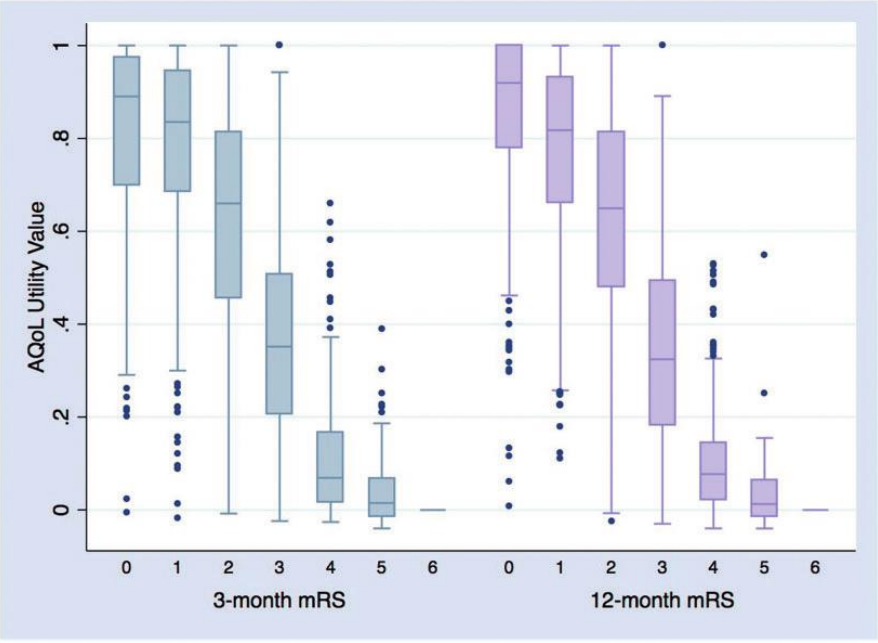
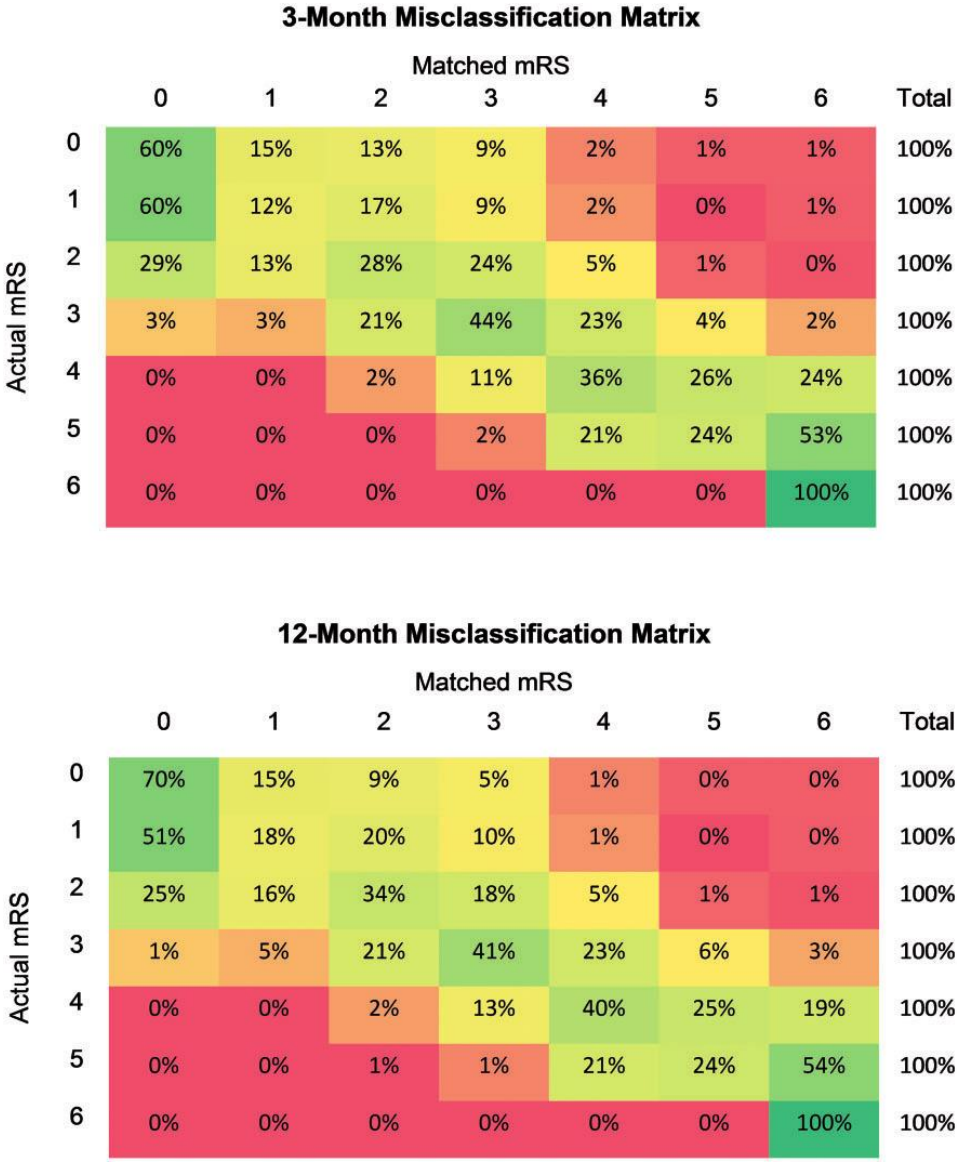


Figure 2. Proportion of patients classified into a matched mRS based on the closest UW-mRS value to their observed AQoL-4D utility value.



within-patient differences (time post-stroke) was 47% (ICC=0.53). The proportion of variability in utility values attributable to time post-stroke for the "no change" groups was: 45% (mRS 0), 42% (mRS 1), 44% (mRS 2), 54% (mRS 3), 56% (mRS 4), and 53% (mRS 5).

Utility variability by derivation method

The UW-mRS values displayed a similar sigmoidal relationship across the mRS categories consistent with previously published utility weights (Table 3). UW-mRS values, however, were generally lower for all categories except mRS 5 and the UW-mRS set based on an applied Singaporean value set by Ali et al.[11] We found variability in UW-mRS values due to the choice of valuing method, the country value set applied, and the stroke population dataset analyzed (Table 3). Taking into consideration age and stroke severity, the choice of the derivation method contributed to 25% and 22% of the variability in UW-mRS values from nine UW-mRS sets at 3 months and two UW-mRS sets at 12 months, respectively (Figure 4).

Discussion

This study is the first to derive an AQoL-4D predicted UW-mRS values which were consistent with prior studies using different derivation methods confirming the unequally ordered intervals of the mRS.[1–3,7,11] Our AQoL-4D UW-mRS values, however, indicated poorer patient perception of their health states compared to previously derived UW-mRS values. This concurs with previous investigations of discrepancies in predicted utilities by different instruments which evaluate related but distinct constructs.[17,18] For example, nearly 80% of the difference in utility weights predicted by AQoL-8D and EQ-5D-3L can be explained by differences in their descriptive systems, i.e. the set of questions and response categories which seek to describe a person's health.[17] The remaining observable difference is due to scale effects and scale-adjusted differences attributable to the utility formula. This is problematic as it compromises the comparability of utility values and, by implication, the UW-mRS.

Extensive research has established the psychometric properties of EQ-5D-3L in a stroke population.[19–21] Accordingly, EQ-5D-3L is an appealing instrument for mapping utility weights to mRS in previously published studies.[2,7,11] However, a criticism of the EQ-5D-3L, and consequently these studies, is its lack of sensitivity to health states.[10,22] The instrument was not designed to assess the needs and deficits of stroke patients. AQoL-4D yields the same generic weakness, but accounts for more stroke- and health-related quality of life domains than the EQ-5D-3L.[10] Given that AQoL-4D has shown predictive validity in a stroke population, it is possible that the lower UW-mRS values observed in this study indicate a more accurate representation of stroke outcomes based on the perspective of stroke patients.[22]

Table 2. ROC curve analyses between sequential mRS categories based on AQoL utility values at 3 and 12 months.

mRS comparisons	3-month AUC (95% CI)	12-month AUC (95% CI)
0 vs. 1	0.54 (0.49–0.60)	0.64 (0.60–0.68)
1 vs. 2	0.70 (0.66–0.74)	0.68 (0.64–0.72)
2 vs. 3	0.80 (0.77–0.83)	0.81 (0.78–0.84)
3 vs. 4	0.87 (0.84–0.90)	0.83 (0.80–0.87)
4 vs. 5	0.72 (0.67–0.76)	0.75 (0.70–0.80)
5 vs. 6	0.63 (0.56–0.70)	0.61 (0.52–0.69)

mRS: modified Rankin Scale; AUC: area under curve; CI; confidence interval.

Figure 3. Distribution of change in AQoL-4D utility values between 3 and 12 months for patients whose mRS did not change ($n = 1189$).

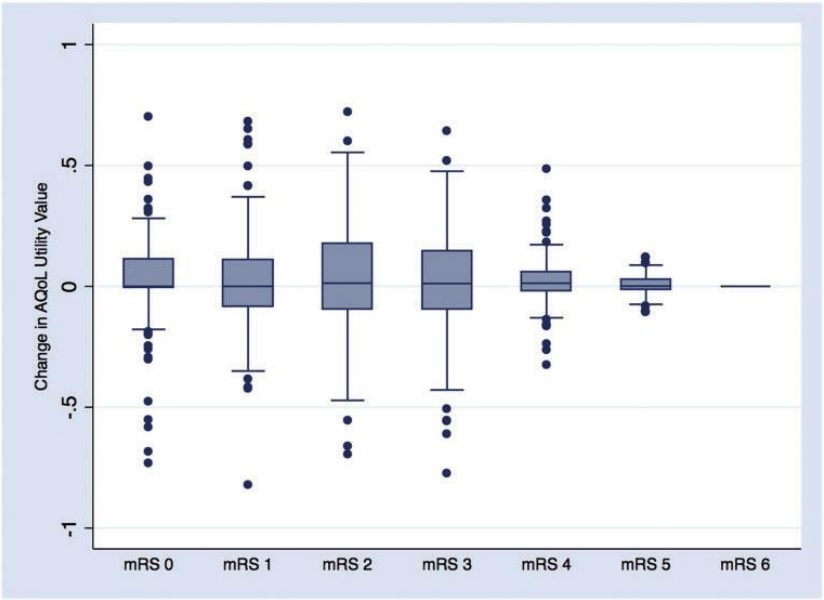
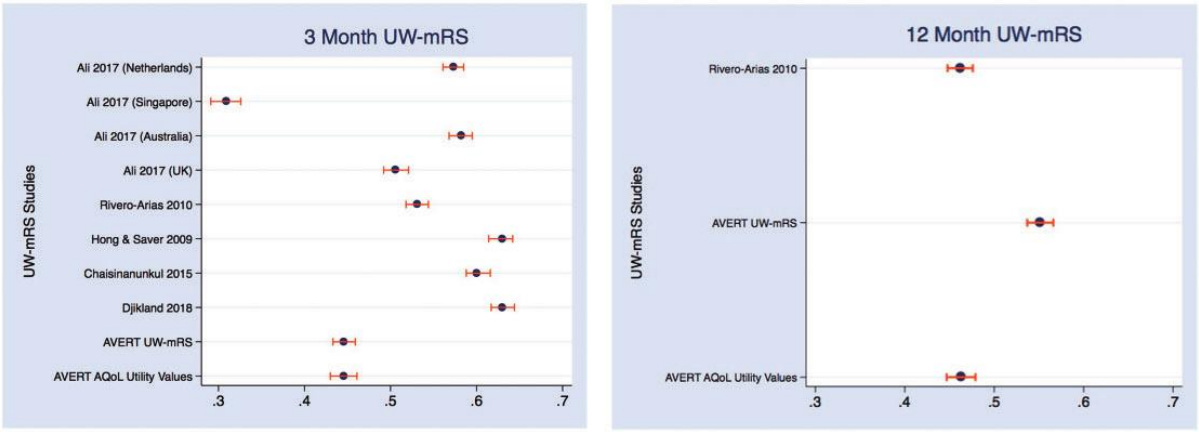


Table 3. Comparison of UW-mRS values and derivation methods between AVERT and previously published studies.

Studies	Valuing method	Country value set applied	Country of population analyzed	*n*	mRS Levels					
					0	1	2	3	4	5
3 month UW-mRS[a]										
AVERT	AQoL TTO	Australia	International	2083	0.80	0.78	0.63	0.37	0.11	0.03
			UK	601	0.78	0.77	0.60	0.35	0.09	0.04
			Australia New Zealand	1238	0.80	0.79	0.64	0.38	0.12	0.03
			Singapore Malaysia	244	0.81	0.76	0.66	0.37	0.09	−0.00
Dijkland et al.[2]	EQ-5D TTO	Netherlands	Netherlands	500	0.95	0.93	0.83	0.62	0.42	0.11
Ali et al.[11]	ED-5D TTO	UK	UK	70	0.90	0.82	0.70	0.53	0.20	−0.15
		Australia	Australia New Zealand	80	0.93	0.86	0.76	0.61	0.35	0.02
		Singapore	Singapore Malaysia Philippines	160	0.88	0.74	0.51	0.23	−0.16	−0.48
		Netherlands	Netherlands France Belgium	215	0.91	0.83	0.73	0.59	0.35	0.12
Chaisinanunkul et al.[3b]	EQ-5D TTO & WHO-GBDP PTO	N/A	International	N/A	1.00	0.91	0.76	0.65	0.33	0.00
Hong and Saver[1c]	WHO-GBDP PTO	N/A	International	N/A	1.00	0.95	0.79	0.67	0.35	0.06
Rivero-Arias et al.[7d]	EQ5D TTO	UK	UK	365/478	0.92	0.81	0.69	0.57	0.24	−0.01
12 Month UW-mRS[e]										
AVERT	AQoL TTO	Australia	International	2052	0.85	0.77	0.63	0.34	0.11	0.03
			UK	585	0.85	0.78	0.54	0.35	0.10	0.02
			Australia New Zealand	1226	0.84	0.76	0.65	0.34	0.11	0.03
			Singapore Malaysia	241	0.91	0.80	0.66	0.34	0.13	0.02
Rivero-Arias et al.[7]	EQ5D TTO	UK	UK	665	0.96	0.83	0.68	0.55	0.32	−0.07

[a]Derived from data collected at three months post-stroke.
[b]Country value set and sample number not specified as UW-mRS was derived by averaging values from TTO and PTO studies.
[c]Country value set and sample number cannot be specified due to PTO valuing method.
[d]UW-mRS derived by averaging one-month and six-month UW-mRS.
[e]Derived from data collected at 12 months post-stroke.
UW-mRS: utility weighted modified Rankin Scale; AQoL: assessment of quality of life; TTO: time trade-off; EQ-5D: European quality of life scale; WHO-GBDP: World Health Organisation burden of disease project; PTO: person trade-off.

Figure 4. UW-mRS values from different studies applied to the AVERT dataset at 3 and 12 months (mean, 95% CI).



Differences in UW-mRS values can also be attributed to the country-specific value set applied and the population sample from which the quality of life and mRS data are collected. The application of different country-specific value sets has previously been shown to result in significantly different utility distributions.[11] Guidance on the most appropriate value set to correctly depict post-stroke outcomes while allowing aggregation of international datasets is still required.

Another threat to construct validity was the inability of UW-mRS values to reflect variability in AQoL-4D utility values between and within individual mRS categories. This may be explained by individual patient differences in the affected dimensions of the AQoL-4D, which has been previously postulated.[11] Neurological deficits vary with the region of sustained damage (e.g. motor vs. cognition), as such, the affected dimensions of AQoL-4D can also vary (e.g. selfcare vs. relationships). These dimensional differences, however, are often classified as the same level of functional impairment by the mRS.

Our findings demonstrate an inability to sufficiently discriminate levels of functional disability based on a patient's perspective of their health. This highlights an obvious issue concerning the simple yet inadequate choice to average utility values per mRS category. Weighting the mRS in this way does not reflect the variability of individual utility values within each category, reduces the sensitivity of the UW-mRS, and disregards crucial information about quality of life. If the UW-mRS only captures a marginally greater understanding of patient perspectives, the justification of its superiority over the mRS as a primary outcome measure is somewhat questionable, particularly when statistical superiority has not been established.[2]

Based on the derived utility values, mRS does not effectively capture patient perceptions of their stroke outcome or final state health outcome. The substantial variability in utility values over time post-stroke, particularly for patients whose mRS score did not change between 3 and 12 months is indicative of the mRS's oversimplification as an outcome measure and the potential disregard of changes in patient perspectives of their health over time. Similarly, we should question whether the UW-mRS is too crude to reflect long-term effects of stroke. Perhaps, utility values should exclusively supplement the mRS rather than be integrated to produce what our data suggest may be a primary outcome measure with limited validity.

The are several strengths of this study; it is the first to derive AQoL-4D UW-mRS values, AVERT provides a large, broad sample of patients early post-stroke, and the longitudinal data facilitated the exploration of patient perspectives over time post-stroke. An outcome measure should consider whether an intervention has long-term benefits for a stroke survivor and reflect changes in quality of life over time. The first limitation of this study is the selection criteria of AVERT. However, compared to other acute stroke trial datasets, it is reasonably non-restrictive, and the large, multinational sample promotes the generalizability of the conclusions. The second limitation is the application of an Australian value set to an international population which, ignores country-specific differences in quality of life.[11] However, the application of customized value sets is problematic when pooling data for analyses. Standardization of UW-mRS derivation methods is required to preserve the validity of data and the ability for data aggregation.

By introducing a meaningful concept of the difference between mRS categories for an otherwise ordinal

scale, the UW-mRS provides an imperfect solution to an important problem. Findings from this study should not discourage the quest to overcome the shortcomings of the mRS. The need to employ outcome metrics devoid of crudity while reflecting patient perspectives of quality of life is still critical.[23] Further investigation of the UW-mRS's added value, including validation on longitudinal data, is time-critical given its recent approval as a co-primary outcome measure by the Food and Drug Administration in a thrombectomy trial.[24] Nonetheless, parallel collection of longitudinal quality of life and mRS data is still warranted. However, the sources of variability shown in this study need to be considered to ensure the use of UW-mRS as a primary outcome measure is justifiable, valid, standardized, statistically efficient, and accurately represents final health states of stroke patients.

Acknowledgments

We thank the hard-working investigators of the AVERT Collaboration, participants, and families for supporting the trial.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: AVERT was supported by the National Health and Medical Research Council of Australia [386201, 1041401]; Chest Heart and Stroke Scotland [Res08/A114]; Northern Ireland Chest Heart and Stroke; Singapore Health [SHF/FG401P/2008]; the UK Stroke Association [TSA2009/09]; and UK National Institute of Health Research [HTA Project 12/01/16]. JB received fellowship funding from the Australia Research Council [0991086] and the National Heart Foundation. The Florey Institute of Neuroscience and Mental Health acknowledge the support of the Victorian Government's Operational Infrastructure Support Grant.

ORCID iD

Venesha Rethnam http://orcid.org/0000-0001-6118-3005
Lan Gao http://orcid.org/0000-0001-6118-3005
Janice Collier http://orcid.org/0000-0001-6118-3005

References

1. Hong K-S and Saver JL. Quantifying the value of stroke disability outcomes: WHO global burden of disease project disability weights for each level of the modified Rankin Scale. *Stroke* 2009; 40: 3828–3833.
2. Dijkland SA, Voormolen DC, Venema E, et al. Utility-weighted modified rankin scale as primary outcome in stroke trials: a simulation study. *Stroke* 2018; 49: 965–971.
3. Chaisinanunkul N, Adeoye O, Lewis RJ, et al. Adopting a patient-centered approach to primary outcome analysis of acute stroke trials using a utility-weighted modified Rankin scale. *Stroke* 2015; 46: 2238–2243.
4. Lees KR, Khatri P, Alexandrov AV, et al. Stroke treatment academic industry roundtable recommendations for individual data pooling analyses in stroke. *Stroke* 2016; 47: 2154–2159.
5. Torrance GW. Utility approach to measuring health-related quality of life. *J Chronic Dis* 1987; 40: 593–600.
6. Whitehead SJ and Ali S. Health outcomes in economic evaluation: the QALY and utilities. *Br Med Bull* 2010; 96: 5–21.
7. Rivero-Arias O, Ouellet M, Gray A, Wolstenholme J, Rothwell PM and Luengo-Fernandez R. Mapping the modified Rankin scale (mRS) measurement into the generic EuroQol (EQ-5D) health outcome. *Med Decis Making* 2010; 30: 341–354.
8. Lenert L and Kaplan RM. Validity and interpretation of preference-based measures of health-related quality of life. *Med Care* 2000; 38: II138–II50.
9. Post PN, Stiggelbout AM and Wakker PP. The utility of health states after stroke: a systematic review of the literature. *Stroke* 2001; 32: 1425–1429.
10. Hawthorne G, Richardson J and Day NA. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. *Ann Med* 2001; 33: 358–370.
11. Ali M, MacIsaac R, Quinn TJ, et al. Dependency and health utilities in stroke: data to inform cost-effectiveness analyses. *Eur Stroke J* 2017; 2: 70–76.
12. Bernhardt J, Langhorne P, Lindley RI, et al. Efficacy and safety of very early mobilisation within 24 h of stroke onset (AVERT): a randomised controlled trial. *Lancet* 2015; 386: 46–55.
13. International classification of impairments, disabilities, and handicaps: a manual of classification relating to the consequences of disease, published in accordance with resolution WHA29. 35 of the Twenty-ninth World Health Assembly, May 1976. World Health Organization, 1980.
14. Hawthorne G, Richardson J and Osborne R. The Assessment of quality of life (AQoL) instrument: a psychometric measure of health-related quality of life. *Qual Life Res* 1999; 8: 209–224.
15. Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med* 2013; 4: 627.
16. Rabe-Hesketh S and Skrondal A. *Multilevel and longitudinal modeling using Stata*. Texas, United States of America: STATA Press, 2008.
17. Richardson J, Iezzi A and Khan MA. Why do multi-attribute utility instruments produce different utilities: the relative importance of the descriptive systems, scale and 'micro-utility'effects. *Qual Life Res* 2015; 24: 2045–2053.
18. Richardson J, Khan MA, Iezzi A and Maxwell A. Comparing and explaining differences in the magnitude, content, and sensitivity of utilities predicted by the EQ-5D, SF-6D, HUI 3, 15D, QWB, and AQoL-8D

multiattribute utility instruments. *Med Decis Making* 2015; 35: 276–291.
19. Dorman PJ, Waddell F, Slattery J, Dennis M and Sandercock P. Is the EuroQol a valid measure of health-related quality of life after stroke? *Stroke* 1997; 28: 1876–1882.
20. Hunger M, Sabariego C, Stollenwerk B, Cieza A and Leidl R. Validity, reliability and responsiveness of the EQ-5D in German stroke patients undergoing rehabilitation. *Qual Life Res* 2012; 21: 1205–1216.
21. Golicki D, Niewada M, Buczek J, et al. Validity of EQ-5D-5L in stroke. *Qual Life Res* 2015; 24: 845–850.
22. Sturm JW, Osborne RH, Dewey HM, Donnan GA, Macdonell RA and Thrift AG. Brief comprehensive quality of life assessment after stroke: the assessment of quality of life instrument in the North East Melbourne Stroke Incidence Study (NEMESIS). *Stroke* 2002; 33: 2888–2894.
23. Broderick JP, Adeoye O and Elm J. Evolution of the modified Rankin scale and its use in future stroke trials. *Stroke* 2017; 48: 2007–2012.
24. Nogueira RG, Jadhav AP, Haussen DC, et al. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *N Engl J Medi* 2018; 378: 11–21.

Check for updates

International Journal of Stroke
0(0) 1–9
! 2020 World Stroke Organization
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1747493020951941
journals.sagepub.com/home/wso
SAGE

Look closer: The multidimensional patterns of post-stroke burden behind the modified Rankin Scale

Venesha Rethnam[1,2], Julie Bernhardt[1,2], Hannah Johns[1,2], Kathryn S Hayward[1,2,3], Janice M Collier[1,2], Fiona Ellery[1,2], Lan Gao[4], Marj Moodie[4], Helen Dewey[5], Geoffrey A Donnan[2,6] and Leonid Churilov[2,6,7]

Abstract

Background: The utility-weighted modified Rankin Scale, representing patient perspectives of quality of life, is a newly proposed measure to improve the interpretability of the modified Rankin Scale. Despite obvious advantages, such weighting imperfectly reflects the multidimensional patterns of post-stroke burden.

Aims: To investigate multidimensional patterns of post-stroke burden formed by individual domains of Assessment of Quality of Life and Barthel Index for each modified Rankin Scale category.

Methods: In the A Very Early Rehabilitation Trial (n¼2104), modified Rankin Scale scores and modified Rankin Scale-stratified Barthel Index scores of Self-care and Mobility, and Assessment of Quality of Life scores of Independent Living, Senses, Mental Health and Relationships were collected at three months. The multivariate relationship between individual Assessment of Quality of Life and Barthel Index domains, and modified Rankin Scale was investigated using random effects linear regression models with respective interaction terms.

Results: Of 2104 patients, simultaneously collected Assessment of Quality of Life, Barthel Index and modified Rankin Scale scores at three months were available in 1870 patients. While individual Assessment of Quality of Life and Barthel Index domain scores decreased significantly as modified Rankin Scale increased ($p < 0.0001$), the patterns of decrease differed by domains ($p < 0.0001$). Patients with modified Rankin Scale 0–1 had the largest post-stroke burden in the Mental Health and Relationship domains, while patients with modified Rankin Scale > 3 showed the greatest burden in Independent Living, Mobility and Self-care domains.

Conclusions: Across the modified Rankin Scale, individual domains are varyingly impacted demonstrating unique patterns of post-stroke burden, which facilitates appropriate assessment, articulation and interpretation of the modified Rankin Scale and utility-weighted modified Rankin Scale.

Keywords

Modified Rankin Scale, primary outcome measure, stroke, assessment of quality of life, A Very Early Rehabilitation Trial, Barthel Index

Received: 21 April 2020; accepted: 23 June 2020

Introduction

The modified Rankin Scale (mRS), initially described as a measure of "post-stroke disability,"[1] includes six ordered grades of functional outcome and a death state. While all mRS categories consistently and explicitly refer to a generic degree of post-stroke disability, some categories are varyingly described in terms of the presence of symptoms, ability to perform activities of daily living (ADL), walking ability and self-care. The variability in description across the scale may explain

[1]Florey Institute of Neuroscience and Mental Health, Heidelberg, Australia
[2]NHMRC Centre for Research Excellence in Stroke Rehabilitation and Brain Recovery, Melbourne, Australia
[3]Melbourne School of Health Sciences, University of Melbourne, Parkville, Australia
[4]Deakin Health Economics, Deakin University, Burwood, Australia
[5]Eastern Health and Eastern Health Clinical School, Monash University, Clayton, Australia
[6]Melbourne Brain Centre, Royal Melbourne Hospital, Melbourne, Australia
[7]Melbourne Medical School, University of Melbourne, Parkville, Australia

Corresponding author:
Venesha Rethnam, Florey Institute of Neuroscience and Mental Health—Austin Campus, 245 Burgundy Street, Heidelberg, Victoria 3084, Australia.
Email: venesha.rethnam@florey.edu.au

the multidimensionality of post-stroke disability. Modern stroke trials also adopt different interpretations of the mRS, inconsistently referring to the scale as a measure of neurologic functional disability,[2] disability,[3,4] global disability,[5–7] functional outcome,[8–10] global functioning[11] and disability or dependence in daily activities[12] post-stroke. Consequently, it is challenging for a clinician or trialist to consistently interpret and communicate the multidimensional disease burden associated with different mRS categories.

The International Classification Functioning, Disability and Health (ICF) Core Set for Stroke was developed to help evaluate and classify the multidimensionality of post-stroke disability, function and health.[13] The Comprehensive and Brief ICF Core Sets for Stroke highlight the necessity for trial outcomes, such as the mRS, to describe a more representative spectrum of problems in the functioning of patients post-stroke. The core set includes domains, not limited to, neurological deficits, inability to perform ADL and self-care tasks, neuropsychological dysfunction and neuropsychiatric deficits.[13–15] Despite this, different mRS categories do not uniformly reflect post-stroke burden, which is further amplified by the "ordered, but unequally spaced" nature of the scale.[16] The utility-weighted mRS (UW-mRS) has been proposed as a solution to this important problem.[16,17] The UW-mRS aims to reflect patient perspectives of quality of life in a single measure that provides a continuous alternative to the mRS. Despite the advantages of the UW-mRS, patients with a given mRS may exhibit a wide range of utility values, and patients with different mRS scores can have similar utility values, making the UW-mRS not sufficiently granular to represent post-stroke burden.[18]

This limitation may be due to the insufficient ability of the single dimensional UW-mRS to represent the multiple domains of post-stroke burden across mRS categories. In A Very Early Rehabilitation Trial (AVERT), mRS, Assessment of Quality of Life (AQoL-4D) and Barthel Index (BI) scores were simultaneously collected as primary and secondary outcomes. Collectively, the mRS, AQoL-4D and BI can measure several aspects of the Brief ICF Core for Stroke. AQoL-4D captures patient perspectives of post-stroke burden and can provide valuable information to supplement the mRS. Similarly, BI can reflect intricacies of physical function that are not represented in the mRS.[19] The aim of this study was to formulate and investigate the *multidimensional patterns* of post-stroke burden for individual mRS categories using AQoL-4D and BI.

Methods

All data that support the findings of this study are available from the corresponding author on reasonable request. Individual participant's three-month mRS, AQoL-4D and BI scores from A Very Early Rehabilitation Trial (AVERT) were obtained.[20] AVERT was a parallel-group, single-blind, randomized controlled trial conducted in 54 acute stroke units across Australia, New Zealand, Malaysia, Singapore and the UK with a total sample of 2104 patients.[20] The aim of the trial was to compare the effectiveness of frequent, higher dose, very early mobilization (sitting, standing or walking <24 h post-stroke) with usual care after stroke. The three-month mRS scores were obtained from an unstructured interview (in person or by telephone) by a trained assessor who was not working in the study stroke unit and was blinded to treatment allocation. Detailed description of baseline demographics, and primary and secondary outcomes of AVERT have been previously published.[20] Ethics approval was obtained from institutional review boards at all sites.

Assessment of quality of life-4D and BI outcomes

The descriptive system for the AQoL-4D was developed within the framework of the World Health Organization's definition of health-related quality of life.[21] AQoL-4D consists of four domains each measured on a scale from 0.00 (worst possible health state) to 1.00 (best possible health state) and containing three items: Independent Living—self-care, household tasks and mobility; Relationships—friendships, isolation and family role; Mental Health—sleeping, worrying and pain; and Senses—seeing, hearing and communication.[21]

The BI measures a patient's performance in 10 ADL to indicate the level of functional independence, ability to mobilize in ADL and the need for assistance in care.[22,23] These items can be divided into two domains: Self-care—feeding, grooming, bathing, bowel and bladder care, and toilet use; and Mobility—ambulation, transfers and stair climbing.[24] Each of these functional categories may be scored from 0 to 1, 0 to 2 or 0 to 3, depending on the item. The summed score ranges from 0 to 20.

The use of composite outcome measures to increase trial efficiency and to capture numerous relevant patient factors have been previously investigated and shown to improve trial power compared to unadjusted mRS analysis alone.[25] In this study, the inclusion of BI provides a clinician assessment of function and independence post-stroke that supplements the patient-reported Independent Living domains of AQoL-4D. To enable the comparison of individual domains within and between mRS categories, AQoL-4D and BI domain scores were scaled to 0% to 100% for BI and 0.00 to 1.00 for AQoL-4D. A percentage of the total achievable score within each BI domain was utilized, e.g., a patient

with a score of 6 of the 8 for Mobility received a score of 75%. Similarly, AQoL-4D domains are bounded by the range of the entire AQoL-4D scale, i.e., 0 (worst health state) to 1 (best health state).

Statistical analyses

A Kruskal–Wallis test was used to investigate the differences in individual AQoL-4D and BI domains across mRS categories. Receiver Operator Characteristics (ROC) curve analyses were conducted to examine the ability of individual AQOL-4D and BI domain scores to discriminate between sequential mRS categories. A ROC curve area of 0.5 indicates no discrimination, while 1.0 indicates perfect discrimination.[26] The comparative speed of decline in AQoL-4D and BI domain scores across increasing mRS categories were investigated with random effects linear regression models using individual AQOL-4D and BI domains and mRS categories as independent variables, individual scores as dependent variables, domain-by-mRS category interaction terms and patients treated as random effects. Patient who died at three months (mRS 6) were not included in the analyses. Statistical analyses were performed with Stata (version 14.2; StataCorp LP, College Station, TX).

Results

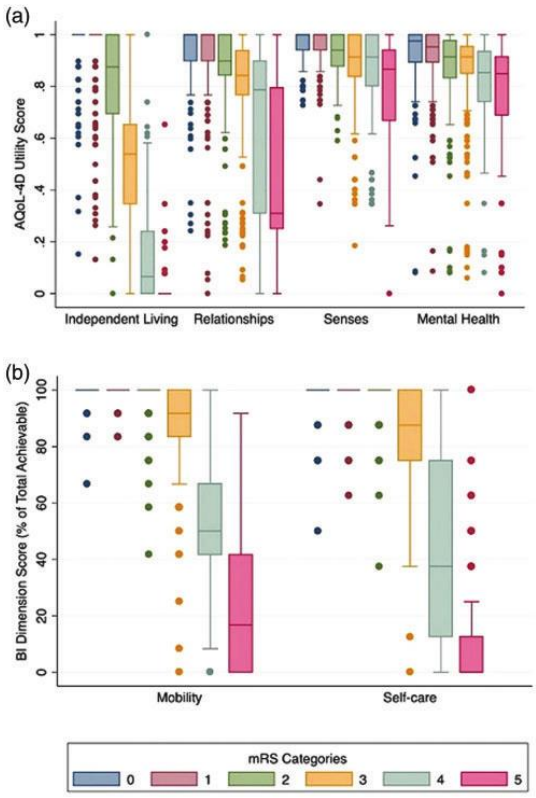
Study population

From 2104 participants recruited in AVERT, 160 participants died at months post-stroke. AQoL-4D, BI and mRS scores were available in 1870 stroke survivors (96.2%) at three months. The proportions within each category are as follows: mRS 0 (n = 182), mRS 1 (n = 393), mRS 2 (n = 404), mRS 3 (n = 448), mRS 4 (n = 261) and mRS 5 (n = 182). The selected sample had a mean (SD) age of 70.1 (12.7) years, median (interquartile range, IQR) baseline National Institutes of Health Stroke Scale (NIHSS) score of 7 (4–12), 61% were males and 88% were ischemic strokes. AQoL-4D was completed by the patient (88%) or a proxy (12%) when communication or cognitive impairments precluded direct questioning of the patient. At three months, the median mRS was 3 (IQR: 1–4).

Post-stroke burden across the mRS using AQoL-4D and BI domains

Kruskal–Wallis test showed significant differences in scores for individual domains of AQoL-4D and BI across mRS categories ($p < 0.0001$; Figure 1). From the results of the ROC curve analysis, the scores for the Independent Living domain demonstrated excellent discrimination for mRS 3 vs. 4 and mRS 2 vs. 3 at three months post-stroke (Table 1). The scores for the Relationships domain could moderately discriminate mRS 2 vs. 3 and 3 vs. 4. Discrimination by other AQoL-4D domains was generally poor for the remaining comparisons. BI Self-care and Mobility domains demonstrated excellent discrimination between mRS 3 vs. 4 and 4 vs. 5 (Table 1).

Figure 1. Distribution of AQoL-4D and BI domain scores stratified by mRS. BI: Barthel Index; AQoL-4D: Assessment of Quality of Life; mRS: modified Rankin Scale.



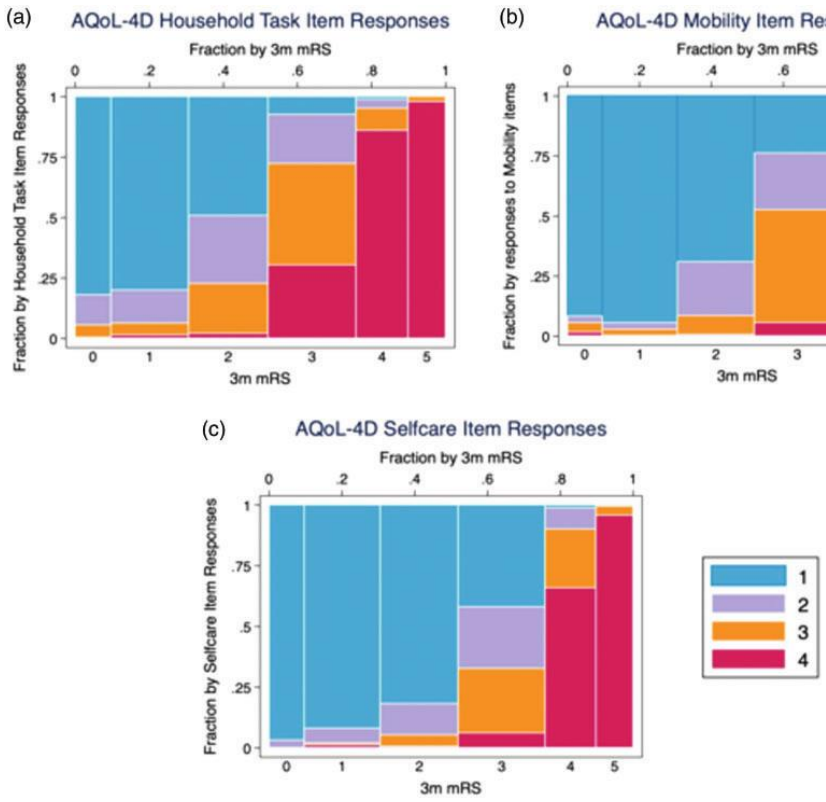
Since the overall pattern of decline in scores of AQoL-4D Independent Living domain across the mRS was not matched by the comparable decline in scores of Mobility and Self-care domains of BI, we undertook an additional exploratory analysis of individual items which contribute to the AQoL-4D Independent Living domain (household tasks, self-care and mobility items). A higher proportion of patients exhibiting poor performance was found in the household tasks items compared to mobility and self-care items of AQoL-4D's Independent Living domain across the mRS (Figure 2). A significant Independent Living individual items-by-mRS interaction effect was found ($p < 0.0001$).

Table 1. AUC values (95% CI) for discrimination between sequential mRS categories for individual domains of post-stroke burden at three months

	mRS 0 vs. 1	mRS 1 vs. 2	mRS 2 vs. 3	mRS 3 vs. 4	mRS 4 vs. 5
Independent living	0.51 (0.47–0.55)	0.69 (0.66–0.73)	0.85 (0.83–0.88)	0.92 (0.89–0.94)	0.72 (0.68–0.75)
Relationships	0.54 (0.50–0.59)	0.60 (0.56–0.64)	0.65 (0.61–0.69)	0.65 (0.60–0.69)	0.63 (0.58–0.69)
Senses	0.53 (0.49–0.58)	0.61 (0.57–0.64)	0.59 (0.55–0.62)	0.53 (0.48–0.57)	0.60 (0.54–0.65)
Mental health	0.53 (0.49–0.58)	0.62 (0.58–0.66)	0.51 (0.48–0.55)	0.59 (0.55–0.64)	0.56 (0.51–0.62)
Self-care	0.51 (0.49–0.53)	0.55 (0.53–0.57)	0.71 (0.68–0.74)	0.91 (0.89–0.93)	0.85 (0.81–0.88)
Mobility	0.50 (0.48–0.52)	0.54 (0.52–0.56)	0.71 (0.68–0.74)	0.89 (0.86–0.91)	0.83 (0.87–0.80)

AUC: area under the receiver operating characteristic curve; CI: confidence interval; mRS: modified Rankin Scale.

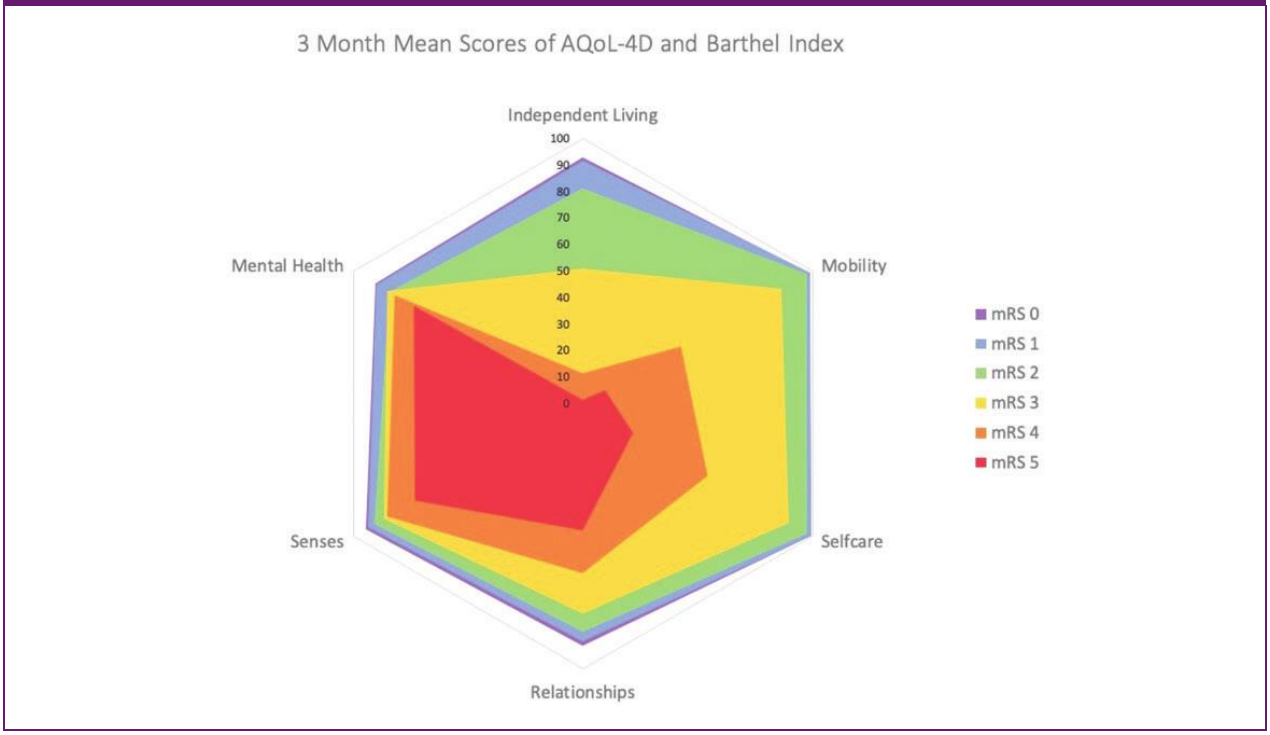
Figure 2. Level of performance in Household Tasks, Mobility and Self-care items at three months. The performance level ranged from 1 (best) to 4 (poor). AQoL-4D: Assessment of Quality of Life; mRS: modified Rankin Scale.



Multidimensional patterns of post-stroke burden across the mRS

A significant AQoL-4D/BI domain-by-mRS interaction effect was found ($p < 0.0001$), i.e., different domains displayed different patterns of change across mRS categories. The multidimensional patterns of post-stroke burden encompassing mean AQoL-4D and BI domain scores at three months for each mRS category are represented in a radar plot (Figure 3). The mean scores for all AQoL-4D and BI domains declined with increasing mRS categories. All independence-related

Figure 3. Mean AQoL-4D and BI domain scores per mRS category. The individual axes are scaled from 0 (poor performance) to 100 (best performance) and correspond to respective domains of post-stroke burden. The level of performance for all six domains is depicted using a colored polygon whereby a smaller area of the polygon is indicative of poorer performance. mRS: modified Rankin Scale.



domains (Independent Living, Self-care and Mobility) displayed a large decline in scores across the mRS while Relationships was the most impacted non-independence-related domain (Figure 3). Based on the mean scores in patients clinically classified as symptomless (mRS 0) or with no significant disability (mRS 1), the most impacted domains post-stroke were Mental Health (sleep, worrying and pain) and Relationships (friendships, isolation and family role; Figure 4). For patients with an mRS 2–5, Independent Living was the most impacted domain.

The multidimensional *patterns of stroke burden* encompassing mean AQoL-4D and BI domain scores at three months for dichotomized mRS 0–2 and 3–5 outcomes are represented in Figure 5. Radar plots for mRS 0–1 vs. 2–5 and 0–3 vs. 4–5 are presented in Supplemental Figures 1 and 2.

Discussion

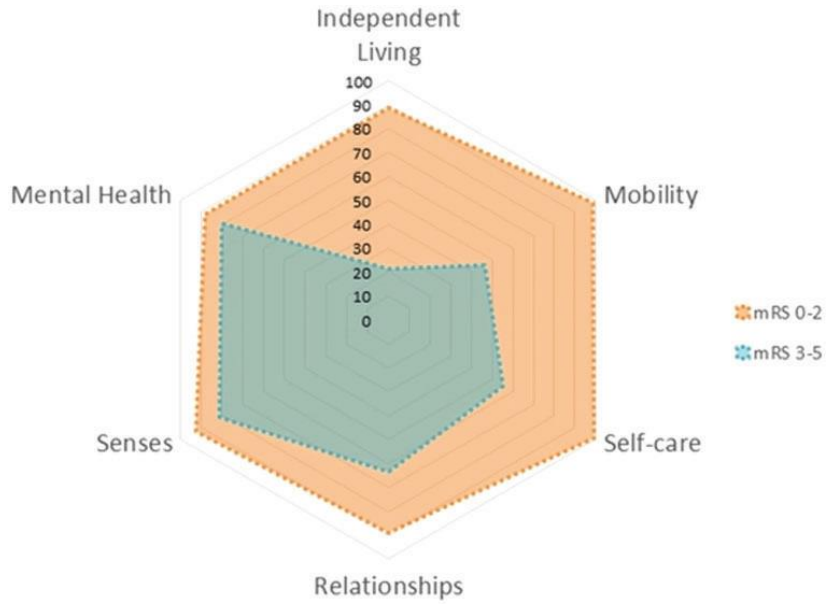
This study has systematically investigated both objective and patient-perceived measures of post-stroke burden at each level of the mRS. We found a significant domain-by-mRS interaction effect, i.e., different post-stroke burden domains displayed different patterns of decline across mRS categories. This contextualizes the outcome of an individual patient that receives a particular mRS score allowing better judgment about the way the mRS should be assessed and interpreted. Our findings support a recent observational cohort study, which examined patient-reported domains of health across levels of the mRS and found that as disability increased, symptoms in multiple disability domains increased in variable degrees.[14] The present study, however, supplements this work by systematically investigating the interactions between, and compensatory nature of, domains of stroke burden at various levels of the mRS in a sample nearly twice the size, and across multiple centers internationally.

We observed a more pronounced decline in mean scores for the domains Mental Health and Relationships compared to Independent Living, Mobility and Self- in patients clinically classified as mRS 0 or 1. This is interesting given several studies have demonstrated difficulties with participation, social roles and reintegration, mental health, cognition and feelings of isolation even after physical (basic self-care and mobility) post-stroke recovery appears complete.[14,27–31] Thus, caution should be taken when generalizing patients with an outcome of mRS 0–1 as

Figure 4. The most impacted domain (lowest mean score) to the least impacted domain (highest mean score) per mRS category. mRS: modified Rankin Scale.

	Most Impacted Domain					Least Impacted Domain
mRS 0	Mental Health	Relationships	Independent Living	Senses	Mobility	Self-care
mRS 1	Relationships	Mental Health	Independent Living	Senses	Mobility	Self-care
mRS 2	Independent Living	Mental Health	Relationships	Senses	Mobility	Self-care
mRS 3	Independent Living	Relationships	Mental Health	Mobility	Senses	Self-care
mRS 4	Independent Living	Mobility	Self-care	Relationships	Mental Health	Senses
mRS 5	Independent Living	Mobility	Self-care	Relationships	Mental Health	Senses

Figure 5. mRS 0–2 vs. 3–5 patterns of post-stroke burden. mRS: modified Rankin Scale.



exhibiting "no symptoms at all" or "no significant disability" particularly when it is evident that non-motor-related domains of stroke burden contribute to long-term outcomes post-stroke.

Patients with an mRS>3 displayed a deterioration in the domains Independent Living, Mobility, Self-care and Relationships. The decline in mean scores for Independent Living, however, were more marked than the decline in mean scores for Mobility and Self-care domains across most mRS categories. Two potential explanations exist: (1) since BI Mobility and Self-care are clinician-rated and AQoL-4D Independent Living is patient-rated, patients may perceive poorer outcomes compared to what is observed and scored by a clinician for the same health state; and/or (2) the more pronounced decline in the Independent Living domain is due to the deterioration in the household tasks item that is not explicitly measured by both domains of the BI. This complements a study which mapped the mRS to the ICF Core Set for stroke and found the strongest association between Self-care and Mobility, followed by Domestic life.[15] Unlike our study, they did not

explicitly consider patient perspectives, therefore difficulties in domestic living, rather than mobility and self-care, may be more important to a stroke survivor. These findings ultimately highlight the multifaceted nature of post-stroke outcomes, i.e., a stroke patient with hemiparesis (impairment) may have poor mobility (activity restriction), that has significant implications on social interactions, ability to return to work, driving, cooking, cleaning, and so on (participation restrictions).[13,32] Further work is required to discriminate between mRS health states by capturing greater information about participation restrictions post-stroke.

Our findings displayed good discrimination between several sequential mRS comparisons in Independent Living, Mobility and Self-care domain scores, particularly mRS 3 vs. 4. This aligns with a recent suggestion that mRS 0–3, rather than 0–2, to be considered a "good" outcome category in moderate-to-severe ischemic stroke.[33] The mechanisms of an intervention and type of patients enrolled in clinical trials are important in determining the "right" statistical approach to analyze the mRS.[34,35] The recommended ordinal analysis approach still raises questions about the ability to interpret the treatment effects and its clinical importance.[36] Thus, by attempting to understand the clinical and societal impact of stroke at each categorical level of the mRS, we are able to understand the monotonic yet unequal impact of stroke burden across the mRS.

There are several strengths of this study. We included domains that do not contribute to the description of each mRS category, i.e., Mental Health (sleeping, worrying and pain), Senses (seeing, hearing and communication) and Relationships (friendships, isolation and family role). These domains formed a more representative spectrum of post-stroke burden of that transcends the current description, which only measure the presence of symptoms, ability to perform ADL, walking ability and self-care.[23] This is important since many trials describe the functional outcome of a patient in terms of the impairment, but it is arguably equally as important to describe domains that are more meaningful to a stroke survivor. Furthermore, the use of AQoL-4D provided rich information about a patient's perception of their health and wellbeing post-stroke that is not explicitly measured by the mRS or BI. This is essential to progress toward a more patient-centric approach to interpreting outcome measures that accurately reflect long-term post-stroke outcomes.

There are also limitations. Clinical trial selection criteria limit the generalizability of the findings. This sample did not include patients admitted to a stroke unit >24 h of stroke onset. However, the sample is broadly similar to population-based norms including a representative proportion of moderate-to-severe ischemic and hemorrhagic strokes (44%).[37] The reasonably non-restrictive, large and multinational sample may reduce external validity concerns about our conclusions. The ceiling effect of BI is a well-known limitation and may have influenced our findings. Additionally, since only two available measures from AVERT were used, our findings are naturally within the scope of AQoL-4D and BI, and only covers a fraction of the Comprehensive ICF Core Set for Stroke. Important post-stroke outcomes, e.g., cognition, were not included in this study. It is possible that other measures could provide additional domains to enrich the clinical picture of three-month stroke outcome. The number of measures selected, the feasibility of data collection, and the interpretability of those results are important trade-offs worth considering for future stroke trials.

The findings from this study have given us a closer look at the multidimensional burden of stroke at each mRS category. There are three main implications that arise from this research. Firstly, the observed non-uniform change in different domains of post-stroke burden across the mRS needs to be considered when interpreting individual scores for future stroke trials. Secondly, the "hidden" domains of burden (i.e., mental health and relationships) in patients classified at symptomless may indicate areas of unmet need(s), which should be considered when developing targeted clinical strategies to support patient's post-stroke. Lastly, the demonstrated importance of domestic living (e.g. driving, cooking and cleaning) to the decline in Independent Living across the mRS may suggest that it is an important measure of post-stroke burden from a patient's perspective. Domestic living is not explicitly captured by the mRS but may be a pragmatic measure to assess post-stroke burden.

Since the superiority of the UW-mRS or other measures over the mRS has not been established, the mRS is not likely to be challenged in its position. Therefore, the findings from our study may provide some important considerations about the multidimensionality of post-stroke burden from a clinical and patient perspective, and the ability (and inability) of the mRS to measure those domains of post-stroke burden.

Acknowledgements

The authors thank the AVERT Trial Collaboration Group Investigators for granting access and use of the AVERT dataset to conduct this study, the blinded assessors who collected the three-month outcome data and participants and their families for supporting the trial. The Florey Institute of Neuroscience and Mental Health acknowledges the support of the Victorian Government's Operational Infrastructure Support Grant.

Declaration of conflicting interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: AVERT was funded by National Health and Medical Research Council (NHMRC) of Australia (386201, 1041401); Chest Heart and Stroke Scotland (Res08/A114); Northern Ireland Chest Heart and Stroke (AVERT-NI 2008); Singapore Health (SHF/FG401P/2008); the UK Stroke Association (TSA2009/09); and UK National Institute of Health Research (HTA Project 12/01/16). VR supported by Post-Graduate Scholarship from NHMRC Centre of Research Excellence in Stroke Recovery and Brain Rehabilitation (1077898). NHMRC fellowship funding was provided to KH (1088449), HD (336102) and JB (1058635, 1154904). JB also received fellowship funding from Australian Research Council (0991086) and the National Heart Foundation. The Florey Institute of Neuroscience and Mental Health acknowledges the support of the Victorian Government's Operational Infrastructure Support Grant.

ORCID iDs

Venesha Rethnam https://orcid.org/0000-0002-0493-8536
Julie Bernhardt https://orcid.org/0000-0002-2787-8484
Hannah Johns https://orcid.org/0000-0003-2135-0504
Kathryn S Hayward https://orcid.org/0000-0001-5240-3264
Janice M Collier https://orcid.org/0000-0003-2950-4870
Lan Gao https://orcid.org/0000-0001-9734-1140

Supplemental Material

Supplemental material for this article is available online.

References

1. Quinn T, Dawson J and Walters M. Dr John Rankin; his life, legacy and the 50th anniversary of the Rankin Stroke Scale. *Scottish Med J* 2008; 53: 44–47.
2. Goyal M, Demchuk AM, Menon BK, et al. Randomized assessment of rapid endovascular treatment of ischemic stroke. *New Engl J Med* 2015; 372: 1019–1030.
3. Hacke W, Kaste M, Bluhmki E, et al. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. *New Engl J Med* 2008; 359: 1317–1329.
4. Furlan A, Higashida R, Wechsler L, et al. Intra-arterial prourokinase for acute ischemic stroke the PROACT II study: a randomized controlled trial. *JAMA* 1999; 282: 2003–2011.
5. Saver JL, Goyal M, Bonafe A, et al. Solitaire™ with the Intention intention for Thrombectomy as Primary Endovascular Treatment for Acute Ischemic Stroke (SWIFT PRIME) Trial: protocol for a randomized, controlled, multicenter study comparing the solitaire revascularization device with IV tPA with IV tPA alone in acute ischemic stroke. *Int J Stroke* 2015; 10: 439–448.
6. Shuaib A, Lees KR, Lyden P, et al. NXY-059 for the treatment of acute ischemic stroke. *New Engl J Med* 2007; 357: 562–571.
7. Jovin TG, Chamorro A, Cobo E, et al. Thrombectomy within 8 hours after symptom onset in ischemic stroke. *New Engl J Med* 2015; 372: 2296–2306.
8. Berkhemer OA, Fransen PSS, Beumer D, et al. A randomized trial of intraarterial treatment for acute ischemic stroke. *New Engl J Med* 2014; 372: 11–20.
9. Hanley DF, Lane K, McBee N, et al. Thrombolytic removal of intraventricular haemorrhage in treatment of severe stroke: results of the randomised, multicentre, multiregion, placebo-controlled CLEAR III trial. *Lancet* 2017; 389: 603–611.
10. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *New Engl J Med* 1995; 333: 1581–1588.
11. Anderson CS, Robinson T, Lindley RI, et al. Low-dose versus standard-dose intravenous alteplase in acute ischemic stroke. *New Engl J Med* 2016; 374: 2313–2323.
12. Qureshi AI, Palesch YY, Barsan WG, et al. Intensive blood-pressure lowering in patients with acute cerebral hemorrhage. *New Engl J Med* 2016; 375: 1033–1043.
13. Geyh S, Cieza A, Schouten J, et al. ICF core sets for stroke. *J Rehabil Med* 2004; 36: 135–141.
14. Katzan IL, Thompson NR, Uchino K and Lapin B. The most affected health domains after ischemic stroke. *Neurology* 2018; 90: e1364–e1371.
15. Berzina G, Sveen U, Paanalahti M and Sunnerhagen KS. Analyzing the modified Rankin scale using concepts of the international classification of functioning, disability and health. *Eur J Phys Rehabil Med* 2016; 52: 203–213.
16. Hong K-S and Saver JL. Quantifying the value of stroke disability outcomes: WHO global burden of disease project disability weights for each level of the modified Rankin Scale. *Stroke* 2009; 40: 3828–3833.
17. Chaisinanunkul N, Adeoye O, Lewis RJ, et al. Adopting a patient-centered approach to primary outcome analysis of acute stroke trials using a utility-weighted modified Rankin scale. *Stroke* 2015; 46: 2238–2243.
18. Rethnam V, Bernhardt J, Dewey H, et al. Utility-weighted modified Rankin Scale: still too crude to be a truly patient-centric primary outcome measure? *Int J Stroke* 2019; 0: 1747493019830583.
19. Kwon S, Hartzema AG, Duncan PW and Min-Lai S. Disability measures in stroke. *Stroke* 2004; 35: 918–923.
20. Bernhardt J, Langhorne P, Lindley RI, et al. Efficacy and safety of very early mobilisation within 24 h of stroke onset (AVERT): a randomised controlled trial. *Lancet* 2015; 386: 46–55.
21. Hawthorne G, Richardson J and Osborne R. The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health-related quality of life. *Qual Life Res* 1999; 8: 209–224.
22. Shah S, Vanclay F and Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. *J Clin Epidemiol* 1989; 42: 703–709.

23. Harrison JK, McArthur KS and Quinn TJ. Assessment scales in stroke: clinimetric and clinical considerations. *Clin Interv Aging* 2013; 8: 201.
24. Sulter G, Steen C and Keyser JD. Use of the Barthel Index and modified Rankin Scale in acute stroke trials. *Stroke* 1999; 30: 1538–1541.
25. Makin SD, Doubal FN, Quinn TJ, Bath PM, Dennis MS and Wardlaw JW. The effect of different combinations of vascular, dependency and cognitive endpoints on the sample size required to detect a treatment effect in trials of treatments to improve outcome after lacunar and non-lacunar ischaemic stroke. *Eur Stroke J* 2018; 3: 66–73.
26. Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Internal Med* 2013; 4: 627.
27. Kapoor A, Lanctôt KL, Bayley M, et al. "Good outcome" isn't good enough: cognitive impairment, depressive symptoms, and social restrictions in physically recovered stroke patients. *Stroke* 2017; 48: 1688–1690.
28. Green TL and King KM. Functional and psychosocial outcomes 1 year after mild stroke. *J Stroke Cerebrovas Dis* 2010; 19: 10–16.
29. Edwards DF, Hahn M, Baum C, et al. The impact of mild stroke on meaningful activity and life satisfaction. *J Stroke Cerebrovas Dis* 2006; 15: 151–157.
30. Carlsson G, Möller A and Blomstrand C. A qualitative study of the consequences of'hidden dysfunctions' one year after a mild stroke in persons<? 75 years. *Disabil Rehabil* 2004; 26: 1373–1380.
31. Verbraak ME, Hoeksma AF, Lindeboom R and Kwa VIH. Subtle problems in activities of daily living after a transient ischemic attack or an apparently fully recovered non-disabling stroke. *J Stroke Cerebrovas Dis* 2012; 21: 124–130.
32. Tse T, Binte Yusoff SZ, Churilov L, et al. Increased work and social engagement is associated with increased stroke specific quality of life in stroke survivors at 3 months and 12 months post-stroke: a longitudinal study of an Australian stroke cohort. *Top Stroke Rehabil* 2017; 24: 405–414.
33. Rangaraju S, Haussen D, Nogueira RG, Nahab F and Frankel M. Comparison of 3-month stroke disability and quality of life across modified Rankin Scale categories. *Interv Neurol* 2017; 6: 36–41.
34. Saver JL. Optimal end points for acute stroke therapy trials. *Stroke* 2011; 42: 2356–2362.
35. Bath PMW, Lees KR, Schellinger PD, et al. Statistical analysis of the primary outcome in acute stroke trials. *Stroke* 2012; 43: 1171–1178.
36. Nunn A, Bath PM and Gray LJ. Analysis of the modified Rankin Scale in randomised controlled trials of acute ischaemic stroke: a systematic review. *Stroke Res Treat* 2016; 2016: 7.
37. Bernhardt J, Raffelt A, Churilov L, et al. Exploring threats to generalisability in a large international rehabilitation trial (AVERT). *BMJ Open* 2015; 5: e008378.

APPENDIX B

Supplementary Material: Chapter 2

AGREE-Recommendation EXcellence (AGREE-REX) Reporting Checklist

This checklist is intended to guide the reporting of clinical practice guideline recommendations.

Manuscripts related to the AGREE-REX are being submitted to peer-reviewed journals for publication. Citations will be added here when they become available.

For more information about the AGREE-REX Reporting Checklist, please visit the AGREE Enterprise website at www.agreetrust.org.

CHECKLIST ITEM	REPORTING CRITERIA	Page #
ITEM 1: CLINICAL APPLICABILITY		
1. Evidence The following criteria are related to the evidence supporting the recommendations.	**To be reported:** ☐ Risk of bias related to the study designs of the supporting evidence ☐ Consistency of the results (i.e., similarity of results across studies) ☐ Directness of evidence to the clinical/health problem (i.e., addresses the exact interventions, populations and outcomes of interest) ☐ Precision of the results (e.g., width of confidence intervals of individual studies or meta-analyses) ☐ Magnitude of the benefits and harms ☐ Likelihood of publication bias ☐ Possibility of confounding variable (if applicable) ☐ Dose-response gradient (if applicable)	
2. Applicability to Target Users The following criteria are related to the applicability of the recommendations to target users.	**To be considered during development and reporting of the recommendations and supporting text:** ☐ A clinical/health problem that is relevant to the intended target users ☐ Alignment between the target user's scope of practice and targeted patients/populations ☐ Alignment between the target user's scope of practice and recommended actions ☐ Alignment between the direction of the recommendations (i.e., in favour of or against a particular action) and the trade-offs between harms and benefits ☐ Alignment between the definitiveness or strength of the recommendations and the trade-offs between harms and benefits	
3. Applicability to Patients/Population The following criteria are related to the applicability of the recommendations to patients/populations.	**To be reported:** ☐ Outcomes relevant to the targeted patients/populations that were considered in the development of evidence base ☐ Recommended actions that have the potential to impact outcomes relevant to patients/populations ☐ How the importance of outcomes to patients was determined ☐ How to tailor recommendations for application to individual (or subsets of) patients or populations (e.g., based on age, sex, ethnicity, comorbidities)	

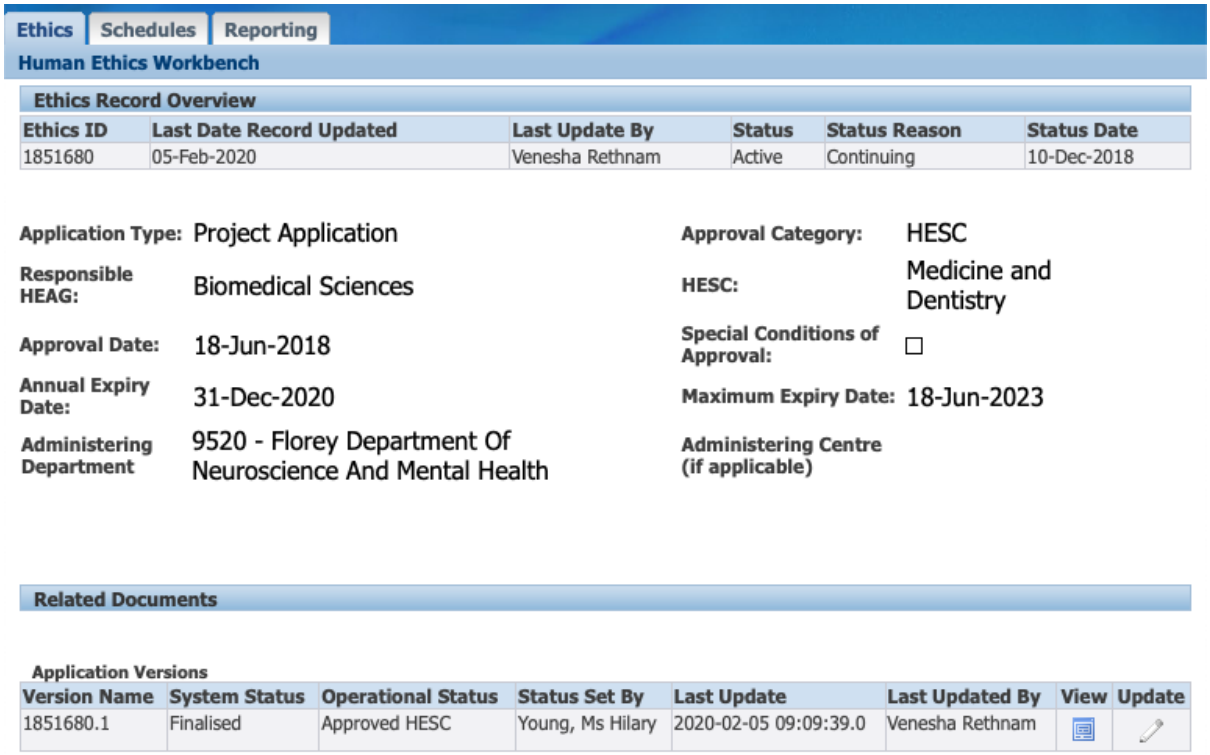
DOMAIN 2: VALUES AND PREFERENCES		
4. Values and Preferences of Target Users The following criteria are related to the target users' values and preferences.	**To be reported:** ☐ Values and preferences of the target users that were considered in relation to the recommended actions ☐ Factors related to target user acceptability of the recommended actions that were considered (e.g., the acceptability of learning new clinical skills or the need to adapt current routine) ☐ Differentiation between recommended actions for which clinical flexibility and individual patient tailoring is more or less appropriate in the decision-making process ☐ Range of recommended actions that are acceptable in the clinical community, including the preferred option (if relevant), and why it is the preferred choice	
5. Values and Preferences of Patients/Population The following criteria are related to the values and preferences of patients/populations.	**To be reported:** ☐ Values and preferences of the target population (e.g., patients, family, caregivers) that were considered in relation to the recommended actions ☐ Factors related to patient/population acceptability of the recommended actions that with considered (e.g., motivation, ability to achieve outcomes, expectations, perceived effectiveness) ☐ Differentiation between recommended actions for which patient choice and/or values are likely to play a large or small part in the decision-making process ☐ Statement about whether tools for assisting in patient decision-making would be beneficial	
6. Values and Preference of Policy/Decision-Makers The following criteria are related to the values and preferences of policy/decision-makers.	**To be reported (if applicable):** ☐ Needs of policy and decision-makers that were considered in the formulation of the recommendations ☐ Impacts of the recommendations on policy and system-level decision-making that were considered in the formulation of recommendations ☐ Impacts of the recommendations on health equities that were considered in the formulation of recommendations ☐ Description of any required changes to policy to align with the recommendations	
7. Values and Preferences of Guideline Developers The following criteria are related to guideline developers' values and preferences and the integration of values and preferences from other stakeholders.	**To be reported:** ☐ Values and preferences that guideline developers brought to the development process ☐ How guideline developer values and preferences influenced their interpretation of the balance between benefits and harms ☐ Methods used to integrate values and preferences, especially when they differ between stakeholders	
DOMAIN 3: IMPLEMENTABILITY		

8. Purpose The following criteria are related to the implementation goals and intended impacts of the guideline.	**To be considered during the development and reporting of the recommendations and supporting text:** ☐ Alignment of guideline recommendations with the implementation goals (e.g., for advocacy, policy change) ☐ Anticipated impacts of recommendation adoption on individuals (e.g., patients, populations, target users), organizations, and/or systems	
9. Local Application and Adoption The following criteria are related to the local application and adoption of the recommendations.	**To be reported:** ☐ Types and degree of change required from current practice ☐ Differentiation between recommendations for which local adaptation may be more or less relevant ☐ Factors that are important for successful dissemination ☐ Issues that can influence the adoption of the recommendations and tools and/or advice for implementers, such as: o Advice on how-to tailor recommendations for the local setting o Resources needed to implement the recommendations and their associated costs o Economic analysis o Required competencies or training to implement recommendations o Data required to implement and monitor the adoption of recommended actions o Strategies to overcome barriers related to provider acceptability and/or patient/population and/or policy acceptability of the recommended action(s). o Criteria that can be used to measure recommendation implementation and quality improvement	

APPENDIX C

Supplementary Material: Chapter 3

Supplementary Material: Ethics approval for Study 1 and Study 3



Ethics | Schedules | Reporting

Human Ethics Workbench

Ethics Record Overview

Ethics ID	Last Date Record Updated	Last Update By	Status	Status Reason	Status Date
1851680	05-Feb-2020	Venesha Rethnam	Active	Continuing	10-Dec-2018

Application Type: Project Application

Approval Category: HESC

Responsible HEAG: Biomedical Sciences

HESC: Medicine and Dentistry

Approval Date: 18-Jun-2018

Special Conditions of Approval: ☐

Annual Expiry Date: 31-Dec-2020

Maximum Expiry Date: 18-Jun-2023

Administering Department 9520 - Florey Department Of Neuroscience And Mental Health

Administering Centre (if applicable)

Related Documents

Application Versions

Version Name	System Status	Operational Status	Status Set By	Last Update	Last Updated By	View	Update
1851680.1	Finalised	Approved HESC	Young, Ms Hilary	2020-02-05 09:09:39.0	Venesha Rethnam		

Supplementary Material: Interview questions on opinions of early mobilisation clinical practice guidelines (Study 1)

1. **Can you talk me through your decision-making process when thinking about mobilising a person early after their stroke?**
    - Rationale: To get the clinician to articulate their decision-making process in order to prepare them for the following questions about how CPGs support this process.
    - Linked to AGREE-REX: *Evidence & Applicability to Patients*

2. **How do clinical practice guidelines influence your decision about early mobilisation?**
    - Rationale: To understand if CPGs are utilised by the recruited clinicians and understand how CPGs support decision-making.
    - Linked to AGREE-REX: *Applicability to Target Users*

3. **In an ideal world, what would you like guidelines to have that might support your decision-making?**
    - Rationale: To understand which aspects of current CPGs are useful and what is needed to better support their decision-making.
    - Linked to AGREE-REX: *Values and Preferences of Target Users*

4. **What issues or situations create uncertainty in your decision-making about early mobilisation?**
    - Rationale: To understand whether uncertainties in decision-making align with the suggestions to improve CPG recommendations.
    - Linked to AGREE-REX: *Evidence & Applicability to Patients*

5. **How do you manage that uncertainty?**
    - Rationale: To understand how clinicians negotiate risks vs benefits with the existing knowledge that further evidence about early mobilisation is needed.
    - Linked to AGREE-REX: *Evidence & Applicability to Patients*

6. **Do you have any other comments you would like to make?**
    - Rationale: To ensure the clinicians have expressed anything else that may be useful in the interpretation of their responses.
    - Linked to AGREE-REX: N/A

Supplementary Table C1: Excluded early mobilisation clinical practice guidelines

Publication	Year	Organisation	Country	Exclusion Reason	Current Status
Consensus on Diagnosis and Treatment of Acute Ischemic Stroke Council – Argentine Society of Cardiology	2012	Argentine Society of Cardiology	Argentina	Pre-2015	Not updated post-2015
Proposal of Guidelines for Acute Stroke Treatment and Management	2009	Belgian Stroke Council	Belgium	Pre-2015	Not updated post-2015
Diretrizes de Atenção À Reabilitação da Pessoa com Acidente Vascular Cerebral	2013	Ministério da Saúde.	Brazil	Pre-2015	Not updated post-2015
Guidelines for Acute Ischemic Stroke Treatment – Part I and Part 2	2012	Executive Committee from Brazilian Stroke Society and the Scientific Department in Cerebrovascular Diseases of the Brazilian Academy of Neurology	Brazil	Pre-2015	Not updated post-2015
Best practice guidelines for the management and rehabilitation of stroke in the north west region of Cameroon	2013	The NWR Best Practices in Stroke Rehabilitation Group	Cameroon	Pre-2015	Not updated post-2015
Stroke and Transient Ischemic Attack - Acute and Long-Term Management	2015	British Columbia Ministry of Health	Canada	Pre-2015	Not updated post-2015
Nova Scotia Guidelines for Stroke Care	2008	Cardiovascular Health Nova Scotia	Canada	Pre-2015	Not updated post-2015
Guía de Práctica Clínica Hemorragia subaracnoideaa rotura de aneurismas cerebrales	2017	Ministry of Health	Chile	No mobilisation recommendations	NA
Guía Clínica Accidente Cerebro Vascular Isquémico, en personas de 15 años y más.	2013	Ministry of Health	Chile	Pre-2015	Not updated post-2015
Guía de práctica clínica de diagnóstico, tratamiento y rehabilitación del episodio agudo del ataque cerebrovascular isquémico en población mayor de 18 años	2015	Ministerio de Salud y Protección Social – Colciencias	Colombia	Pre-2015	Not updated post-2015
Recommendations for Stroke Management 2006 Update	2006	Croatian Society for Neurovascular Disorders of Croatian Medical Association, Croatian Stroke Society, and University Department of Neurology, Sestre milosrdnice	Croatia	Pre-2015	Not updated post-2015

		University Hospital, Reference Center for Neurovascular Disorders of the Croatian Ministry of Health			
Guías clínicas de medicina interna	2018	Ministry of Health	El Salvador	No mobilisation recommendations	NA
2016 European Guidelines on cardiovascular disease prevention in clinical practice	2016	The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts) Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR)	Europe	No mobilisation recommendations	NA
Evidence-based stroke rehabilitation: an expanded guidance document from the European stroke organisation (ESO) guidelines for management of ischaemic stroke and transient ischaemic attack	2008	European stroke organisation (ESO)	Europe	Pre-2015	Not updated post-2015
European Stroke Organisation (ESO) Guidelines for the Management of Spontaneous Intracerebral Hemorrhage	2014	European Stroke Organisation (ESO)	Europe	Pre-2015	Not updated post-2015
Stroke: Early Management (Alert, Pre-hospital Phase, Initial Hospital Phase, Indications for Thrombolysis), Guidelines	2010	Haute Autorité de Santé	France	Pre-2015	Not updated post 2015
Rehabilitation der Mobilität nach Schlaganfall (ReMoS)	2015	The German Socienty of Neurorehabilitation	Germany	No mobilisation recommendations	NA
Acute therapy of ischemic stroke	2012	German Sociecty of Neurology	Germany	Pre-2015	Not updated post 2015
Manejo Rehabilitativodel ICTUS	2013	Instituto Guatemalteco de Seguridad Social	Guatemala	Pre-2015	Not updated post-2015
Recommendations for the Early Management of Acute Ischemic Stroke: A Consensus Statement for Healthcare Professionals from the Indian Stroke Association	2018	Indian Stroke Association	India	No mobilisation recommendations	NA

National Clinical Guidelines and Recommendations for the Care of People with Stroke and Transient Ischemic Attack	2010	Irish Heart Foundation: Council for Stroke	Ireland	Pre-2015	Not updated post-2015
Stroke in General	2011	Joint Committee on Japanese Guidelines for the Management of Stroke 2009	Japan	Pre-2015	Not updated post-2015
Clinical Practice Guidelines, Management of Ischaemic Stroke, 2 Edition 2012	2012	Ministry of Health Malaysia	Malaysia	Pre-2015	Not updated post-2015
GPC: VIGILANCIA Y PREVENCIÓN SECUNDARIA DE LA ENFERMEDAD VASCULAR CEREBRAL EN EL PRIMER NIVEL DE ATENCIÓN	2015	National Center for Technological Excellence in Health	Mexico	Pre-2015	Not updated post-2015
Management of patients with stroke: Rehabilitation, prevention and management of complications, and discharge planning	2013	MOH and health project of MCC, WHO and professional committee of rehabilitation of MOH, Mongolian society of physical & rehabilitation medicine, University of Health Sciences Mongolia (HSUM), the Shastin and State hospital	Mongolia	Pre-2015	Not updated post-2015
NHG Standard Stroke	2018	Dutch College of General Practitioners	Netherlands	Doesn't include AVERT	
New Zealand Clinical Guidelines for Stroke Management	2010	Stroke Foundation of New Zealand	New Zealand	Pre-2015	Stroke foundation guidelines
Ischemic stroke care guidelines for Pakistan	2009	Pakistan Society of Neurology	Pakistan	Pre-2015	Not updated post-2015
Guidelines for the Prevention, Treatment and Rehabilitation of Stroke	2010	Stroke Society of the Philippines	Philippines	Pre-2015	Not updated post-2015
The diagnosis and management of stroke and transient ischaemic attack	2016	Ministry of Public Health	Qatar	No mobilisation recommendations	NA
Ишемический инсульт и транзиторная ишемическая атака у взрослых	2015	Scientific Council of the Ministry of Health of the Russian Federation	Russia	Pre-2015	Not updated post-2015
118 Management of patients with stroke: rehabilitation, prevention and management of complications, and discharge planning	2010	Scottish Intercollegiate Guidelines Network	Scotland	Pre-2015	Not updated post-2015
Stroke and transient ischaemic attacks. Assessment, investigation, immediate management and secondary prevention	2009	Singapore Ministry of Health	Singapore	Pre-2015	Not updated post-2015

South African guideline for management of ischaemic stroke and transient ischaemic attack 2010	2010	South African Stroke Society (SASS)	South Africa	Pre-2015	Not updated post-2015
Management of Stroke			Sri Lanka	No mobilisation recommendations	NA
https://www.socialstyrelsen.se/regler-och-riktlinjer/nationella-riktlinjer/publicerade-riktlinjer/stroke/	2020	Socialstyrelsen	Sweden	No mobilisation recommendations	NA
General Principles for the Management of Acute Ischemic Stroke	2014	Taiwan Stroke Society	Taiwan	Pre-2015	Not updated post-2015
Guidelines for Stroke Unit Organization	2012	The Neurological Society of Thailand	Thailand	Pre-2015	Not updated post-2015
ІШЕМІЧНИЙ ІНСУЛЬТ (екстрена, первинна, вторинна (спеціалізована) медична допомога, медична реабілітація)	2012	Order of the Ministry of Health	Ukraine	Pre-2015	Not updated post-2015
Guidelines for the Management of Spontaneous Intracerebral Hemorrhage	2010	AHA/ASA	USA	Pre-2015	Not updated post-2015

Supplementary Table C2: Included early mobilisation clinical practice guidelines

Organisation	CPG Document	Country	Year	Recommendation	Evidence
Sociedad Neurológica Argentina (SNA), Sociedad Argentina de Cardiología (SAC), Sociedad Argentina de Terapia Intensiva (SATI), Sociedad Argentina de Medicina (SAM), Sociedad Argentina de Emergencias (SAE), Colegio Argentino de Neurointervencionitas (CANI), Asociación Argentina de Neurocirujanos (AANC), Sociedad Argentina de Radiología (SAR) y Sociedad Argentina de Geriatría y Gerontología (SAGG).	*Consensus On Acute Ischemic Cerebrovascular Accident*	Argentina	2019	It is recommended: - Early mobilization to prevent complications (Class I, Level of evidence C) Expert Opinion It is not recommended: - Intense and early mobilization (within 24 hours of the onset of symptoms), as it may reduce the chances of a good prognosis (Class III, Level of evidence B) From one randomised trial	AVERT Group, 2015

Stroke Foundation	*Clinical Guidelines for Stroke Management*	Australia/New Zealand	2017	1. For stroke patients, starting intensive out-of-bed activities within 24 hours of stroke onset is not recommended. (Strong recommendation against). i. Baseline stroke severity and stroke type should be considered when deciding when and how much to mobilise after stroke. Intensive out-of-bed activities in the AVERT trial were an average of 6 out-of-bed activities per day (Bernhardt et al. 2015). Patients who are independently mobile should not have their mobility restricted during the early phase after stroke. 2. All stroke patients should commence mobilisation (out-of-bed activity) within 48 hours of stroke onset unless otherwise contraindicated (e.g. receiving end-of-life care). (Strong recommendation) i. Patients with baseline NIHSS scores 4-7 have higher odds of a favourable outcome when they are mobilised more than once per day and	1. Bernhardt et al. 2015 2. Bernhardt et al. 2015; Lynch et al. 2014, Langhorne 2010 3. Bernhardt et al. 2015

				spend less than 13.5 minutes per day mobilising with physiotherapy staff (Bernhardt et al. 2016). For patients who have difficulty moving after stroke, an assessment by an appropriately trained health professional as to the most appropriate and safe methods of assisting transfers and out-of-bed activity should be conducted as soon as possible and preferably within 24 hours. 3. For patients with mild and moderate stroke, frequent, short sessions of out-of-bed activity should be provided, but the optimal timing within the 48-hour post-stroke time period is unclear. (Weak recommendation)	

Heart and Stroke Foundation, Canada	*Canadian Stroke Best Practice Recommendations for Acute Stroke Management: Prehospital, Emergency Department, and Acute Inpatient Stroke Care, 6th Edition*	Canada	2018	*9.4 Mobilization (revised for 2018)* *Mobilization is defined as "the process of getting a patient to move in the bed, sit up, stand, and eventually walk."* i. All patients admitted to hospital with acute stroke should have an initial assessment, conducted by rehabilitation professionals, as soon as possible after admission [Evidence Level A]. ii. Initial screening and assessment should be commenced within 48 h of admission by rehabilitation professionals in direct contact with the patient [Evidence Level C]. Refer to Canadian Stroke Best Practice Recommendations Stroke Rehabilitation Module for additional recommendations on mobilization following an acute stroke. iii. Rehabilitation therapy should begin as early as possible once the patient is determined to be medically able to participate in active rehabilitation [Evidence Level A].	AVERT Group, 2015 and cohort studies and other RCTs

				iv. Frequent, brief, out-of-bed activity involving active sitting, standing, and walking, beginning within 24 h of stroke onset is recommended if there are no contraindications [Evidence Level B]. More intense early sessions are not of more benefit. Clinical judgment should be used. Note: Contraindications to early mobilization include, but are not restricted to, patients who have had an arterial puncture for an interventional procedure, unstable medical conditions, low oxygen saturation, and/or lower limb fracture or injury.	
Heart and Stroke Foundation, Canada	*Rehabilitation, Recovery and Community Participation following Stroke Part One: Rehabilitation and Recovery following Stroke Update 2019*	Canada	2019	*i. All patients with stroke should receive rehabilitation therapy as early as possible once they are medically stable and able to participate in active rehabilitation [Evidence Level A]. Refer to Section One, Box One: Eligibility and Criteria for Stroke Rehabilitation for more information.* *ii. Early prolonged mobilization of patients within the first few days*	AVERT Group, 2015 and cohort studies and other RCTs

				after a stroke, especially a severe stroke, is not recommended (Evidence Level A). *iii. Earlier mobilization may be reasonable for select patients with acute stroke (for instance people with more mild strokes or transient ischemic attack) but caution is advised, and clinical judgement should be used (Evidence Level C)*	
European Stroke Organisation	*Consensus statements and recommendations from the ESO Karolinska Stroke Update Conference, Stockholm 11–13 November 2018*	Europe	2018	1. The evidence point to that early mobilization is safe in stroke patients but should not be too intense (Grade B). 2. A progressive adaptation to activities of daily living, such as going to the toilet with assistance (if needed) or sitting in a chair to eat is fine (Grade A). 3. Patient should be clinically observed and monitored closely and in case they present symptoms noted (Grade C). 4. Early mobilization after a stroke should be adapted to patient's clinical and neurological situation (Grade C).	Bernhardt et al, 2015 Anderson et al, 2017

Finnish Medical Society Duodecim and the Finnish Neurological Association	*Cerebral infarction and TIA*	Finland	2016	Postural treatment is started immediately to activate the body's senses to promote rehabilitation and prevent body and limb dysfunction. Early mobilization also reduces the risk of pressure ulcers, pneumonia, deep venous thrombosis and pulmonary embolism [155], but mobilization should not be initiated within the first 24 hours [416].	European Stroke Organisation (ESO) Executive Committee, 2008 AVERT II, VERITAS, AVERT III Kaste M, Roine RO. General stroke management and stroke units. In: Stroke, Pathophysiology, Diagnosis and Management. Mohr JP, Choi D, Grotta JC, Weir B, Wolf P (ed.) Philadelphia, Churchill Livingstone, 2004
Directorate General of Health Services Ministry of Health and Family Welfare Government of India	*Guidelines for Prevention and Management of Stroke*	India	2019	1. All patients should be referred to a physiotherapist/rehabilitation as soon as possible, preferably within 24to 48 hours of admission. 2. Passive full-range-of-motion exercises for paralyzed limbs can be started during the first 24 hours. 3. The patient's need in relation to moving and handling should be assessed within 48 hours of admission Early mobility (Mobility between 24-48 hours of stroke) has been associated with improved functional outcomes. To	Not reported

				mobilize stroke patients following actions should be taken considering contraindications and termination criteria (Annexure XIV): a. Active and passive range of motion (ROM) exercise should be initiated within 24 hours of onset of stroke. b. ROM exercises should be limited to point of resistance in unconscious patients and range of comfort in conscious patients. c. Bedside sitting and standing can be initiated within first 48 hours of stroke. d. In patients who have undergone angiography through femoral artery, range of motion exercise for hip joint should be deferred for 6 hours post angiography. Monitoring of any bleeding and discomfort should be noted for at least 2 days post angiography. e. Functional retraining (i.e. participation in bedside mobility, transfers, and activities of daily living) and active participation in self care	

				should be initiated as soon as patient is able to participate. f. In hemorrhagic stroke, mobilization should be delayed till volume of blood gets stabilized for atleast 24 hours. g. Patients with extra ventricular drain (EVD) can be mobilized provided EVD is clamped and secured and intracranial pressure is kept below 20 mm Hg. h. Vital parameters (Heart rate, Blood pressure, Respiratory rate, Temperature), Intracranial pressure if available should be monitored prior to start early mobility, during the session and post interventions. Level 2: Patients who are unable to stand due to poor sensorium or flaccidity, tilt table may be used to make them stand.	
SPREADS – Stroke Prevention and Educational Awareness Diffusion	*Brain stroke: Italian guidelines for prevention and treatment Recommendations and Synthesis*	Italy	2016	Recommendation 14.2.e Strong in favour In stroke patients it is recommended to carry out interventions as early as the first 24 hours mobilization and rehabilitation activities (moderate intensity), if	AVERT III

				they do not exist contraindications to the program. Recommendation 14.2.f Strong against In the first 24 hours after a stroke, mobilization interventions are not recommended intensive with early verticalization, particularly in severe strokes and strokes bleeding.	
Korea Society for Neurorehabilitation	*Clinical Practice Guideline for Stroke Rehabilitation in Korea 2016*	Korea	2017	1-2-1. It is strongly recommended that rehabilitation of acute stroke patients be started as soon as possible when medically stable. (Recommended level A, evidence level 1 ++) 1-2-2. Patients with acute stroke should begin rehabilitation within 72 hours after stroke. (Recommended level B, evidence level 1+) 2-1-1. Patients should move from bed early unless there are contraindications, and it is strongly recommended to carry out wheelchair movements and standing walking. (Recommended level A, evidence level 1 ++)	Several trials including AVERT II and III, Sundseth, VERITAS

				2-1-2. It is not recommended that all stroke patients undergo high intensity gait training uniformly within 24 hours of onset. (Recommended level B, evidence level 1+) 2-1-3. Stroke patients with movement disorders should consider comprehensive exercise recovery programs early, unless contraindicated. (Recommended level GPP)	
Federation of Medical Specialists (Dutch Association for Neurology)	*Cerebral infarction and bleeding*	Netherlands	2019	Mobilize patients with severe cerebral infarction or major bleeding (NIHSS >16) during the first days after a cerebral infarction briefly (no more than 10 consecutive minutes) and frequently (at least 2 to 3 times a day). Mobilize the remaining patients with cerebral hemorrhage or cerebral infarction to the best of their ability. High GRADE There is strong evidence that patients should be able to be mobilized. There is no evidence that secondary complications can be prevented within 24 hours by early	References AVERT III, AVERT II Bernhardt, 2004 Bernhardt, 2009 (Cochrane Review) Cumming, 2008 Cumming, 2011 Otterman, 2012 Van Wijk, 2012

				mobilization. There is no evidence that by mobilizing patients early, neurological recovery, walking ability and ADL independence can be favorably influenced after three months. There is no evidence that early mobilization has a beneficial effect on the duration of admission to a stroke unit. Mediocre GRADE There is indirect evidence that early mobilization within 24 hours after a cerebral infarction or stroke may be detrimental to patients with stroke or severe neurological failure (NIHSS> 16) when admitted to a stroke unit.	
Norwegian Directorate of Health	*National professional guideline for treatment and rehabilitation of stroke*	Norway	2017	For patients with stroke, early mobilization is recommended. SIGN B 2a. Frequent mobilization of short duration in the early phase is suggested. (Strong Recommendation). Practical - this is how the recommendation can be followed Bed rental disposes for	AVERT III Govan, et al, 2007 Indredavik et al, 1999 Legg et al, 2007 Sundseth et al, 2012 Stroke Unit Trialists'Collaboration, 2013 (Cochrane Review) Holte al, 2016.

				immobilization-related complications and early mobilization in stroke unit seem to reduce such complication, and early mobilization is therefore recommended for most patients with acute stroke. [170] ; [171]. However, very early mobilization <24 hours combined with intensive exercise can have negative effects [169]. Early mobilization should therefore be individualized and adapted to the individual patient. Time of mobilization Most patients with mild stroke (National Institute of Health Stroke Scale (NIHSS) 0-5)) who have stable or decreasing symptoms and stable BT can be mobilized within the first 24 hours after onset. For patients with moderate to severe symptoms (NIHSS> 5) and stable BT, initial mobilization should be performed between 24-48 hours after onset. In unstable patients with fluctuating or increasing symptoms and unstable BT, mobilization should	

				usually be postponed until the patient is stable. First mobilization Prior to initial mobilization, the patient should be well hydrated (saline solution may be given 500-1000 ml intravenously). At initial mobilization, BT should also be measured before starting and after 15 minutes. For BT falls> 20mmHg or increasing symptoms including dizziness, mobilization should be interrupted. Patients should follow an "out-of-bed protocol" adapted to the individual, but with the aim of regaining standing and walking function and ADL function. During the first 24-48 hours, mobilization should consist of short periods several times daily (4-6) out of bed - initially of duration 15-20 minutes and gradually increasing if the patient's condition, including BT, is stable. Content of early mobilization The mobilization consists of the patient getting into a sitting, standing or walking	

				position, depending on the level of function. Exercise should be task-related and include daily activities such as transfer, personal care, eating, dressing and undressing and toilet visits. Patients need motivation for self-activity, and appropriate guidance and help when needed. Everyone in the multidisciplinary team should help to mobilize the patient early, initially with short periods out of bed, and during observation. The amount of exercise and duration of exercise periods out of bed seems to be of greater importance than the time of start of mobilization, at least in the time period up to 48 hours after onset. Therefore, when designing the recommendation and the practical advice, one has chosen to place considerable emphasis on the dose; frequency and duration based on clinical experience and analyzes from the AVERT study [168]. In addition, emphasis has been placed on the	

				documentation available from stroke unit studies indicating that long-term bed rest is negative [1] .	
Este documento debe ser citado como: “Instituto de Evaluación de Tecnologías en Salud e Investigación. Guía de Práctica Clínica para el diagnóstico y tratamiento de la etapa aguda del Ataque CerebrovascularIsquémico: Guía en Versión Corta. Lima: EsSalud; 2018”	GUÍA DE PRÁCTICA CLÍNICAPARA EL DIAGNÓSTICO Y TRATAMIENTO DEL PACIENTE CON ATAQUE CEREBROVASCULAR ISQUÉMICO (ACVi)	Peru	2018	Disability (early start of rehabilitation therapy at 72 hours) The Colombian Guide uses Lynch's SR (142), which identifies clinical trials and cohorts without actually carrying out meta-analyzes, indicating that in two cohort studies the effects of starting physical rehabilitation in the first 3 days (72 hours) were compared with 4 or more days, said intervention improved the probability of reducing disability by 15.3% (p <0.001), very low quality of evidence. Mortality (early start of mobilization at 24 hours) The early mobilization of patients after a stroke (within 24 hours) does not show significant differences in mortality at three months compared to late mobilization (4 studies, n = 2446, RR 1.21, 95% CI 0.76 to 1.65, I2 0%, p = 0.54, quality of evidence moderate).	Li et al, 2018

				Disability (early start of mobilization at 24 hours) The early mobilization of patients after a stroke (within 24 hours) does not show significant differences in disability (Rankin 0 to 2) at three months compared to late mobilization (5 studies, n = 2373, RR 0.80, 95% CI 0.58 to 1.02, I2 45%, p = 0.12, low quality of evidence). It is suggested that stroke patients with some degree of disability include them in physical rehabilitation programs. It is suggested to start physical rehabilitation within 72 hours and mobilization of the patient within 24 hours taking into account the patient's condition and hemodynamic status. RConditional in favorLow (⊕⊕⊖⊖)	
Royal College of Physicians Intercollegiate Stroke Working Party	*Royal College of Physicians 2016 National clinical guideline for stroke.*	UK	2016	*Immobility and/or bed rest are well-documented to have detrimental effects on hospital patients in general. Early mobilisation (e.g. activities such as sitting out of bed, transfers, standing and*	A Working Party consensus B AVERT Trial Collaboration group 2015; Bernhardt et al, 2016

				walking) aims to minimise the risk of the complications of immobility and improve functional recovery. *A Patients with difficulty moving after stroke should be assessed as soon as possible within the first 24 hours of onset by an appropriately trained healthcare professional to determine the most appropriate and safe methods of transfer and mobilisation.* *B Patients with difficulty moving early after stroke who are medically stable should be offered frequent, short daily mobilisations (sitting out of bed, standing or walking) by appropriately trained staff with access to appropriate equipment, typically beginning between 24 and 48 hours of stroke onset. Mobilisation within 24 hours of onset should only be for patients who require little or no assistance to mobilise.*	
National Institute for Health and Clinical Excellence	*Stroke and transient ischaemic attack in*	UK	2019	*1.7.2 Help people with acute stroke to sit out of*	https://www.nice.org.uk/guidance/ng128/evidence/f-very-early-mobilisation-pdf-6777399571

	over 16s: diagnosis and initial management			*bed, stand or walk as soon as their clinical condition permits as part of an active management programme in a specialist stroke unit. [2019]* *1.7.3 If people need help to sit out of bed, stand or walk, do not offer high-intensity mobilisation in the first 24 hours after symptom onset. [2019]*	
American Heart Association/American Stroke Association	*Guidelines for the Early Management of Patients With Acute Ischemic Stroke: 2019 Update to the 2018 Guidelines for the Early Management of Acute Ischemic Stroke: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association*	USA	2019	*High-dose, very early mobilization within 24 hours of stroke onset should not be performed because it can reduce the odds of a favorable outcome at 3 months. (III: Harm; B-R)*	Supplementary: Table LXVIII. Randomized Clinical Trials of Mobility Intervention included: SEVEL, 2016; Morreale, 2016; AVERT III, 2015
American Heart Association/American Stroke Association	*Guidelines for Adult Stroke Rehabilitation and Recovery*	USA	2016	*High-dose, very early mobilization within 24 hours of stroke onset can reduce the odds of a favorable outcome at 3 months and is not recommended. (III; A)*	AVERT II, AVERT III

Appendix D

Supplementary Material: Chapter 4

Supplementary Material: Template of Memorandum of Understanding

THE FLOREY

Early mobilisation after stroke: an IPD meta-analysis

Memorandum of Understanding

You are invited to take part in an individual participant data meta-analysis, which aims to explore the dose effects, safety and efficacy of early mobilisation (EM) post stroke.

1. Please provide the name of the person who will be the primary contact person for any data requests or any data related issues in your group, if such a person can be identified:

   _____

   Email: _____

2. Please obtain approval (if considered necessary) from your institutional review board for such sharing of de-identified data.
3. Sharing of de-identified data does not transfer any ownership rights or other intellectual property rights.
4. The source of the data will be acknowledged in all publications of any form (including electronic) resulting from the use of such data.
5. This Memorandum of Understanding is not a legally binding instrument.

Signature from Investigator:

Yes, as principal investigator of the study: [**Study Title**], I am pleased to be part of this IPD meta-analysis. My co-investigators and I am willing to:

1. Share data at the level of individuals for the specified variables
2. Contribute intellectually to meet the objectives of this IPD meta-analysis

(Signature): _____

(Print Name): _____

SIGNATURE

Professor Julie Bernhardt
Head of AVERT Early Rehabilitation Research Group
THE FLOREY
The Florey Institute of Neuroscience & Mental Health

Appendix E

Supplementary Material: Chapter 5

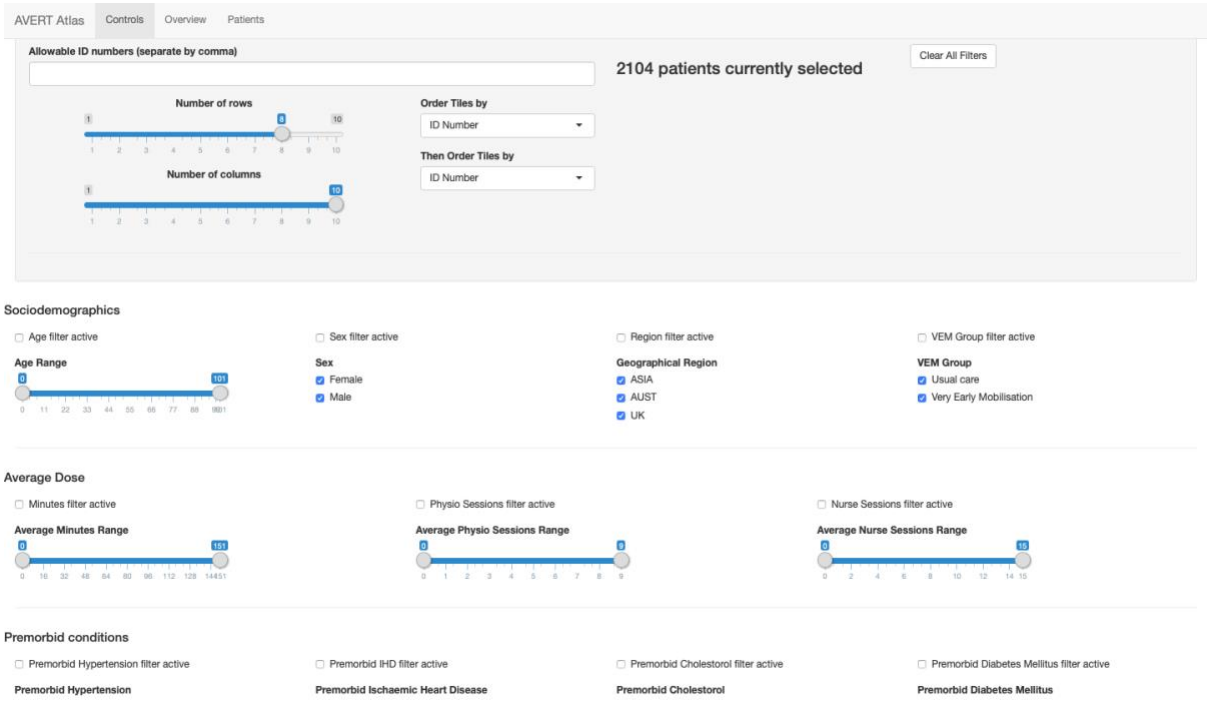
Supplementary Material: Interview Questions on Early Mobilisation Decision-Making (Study 3)

Introductory Interview Questions

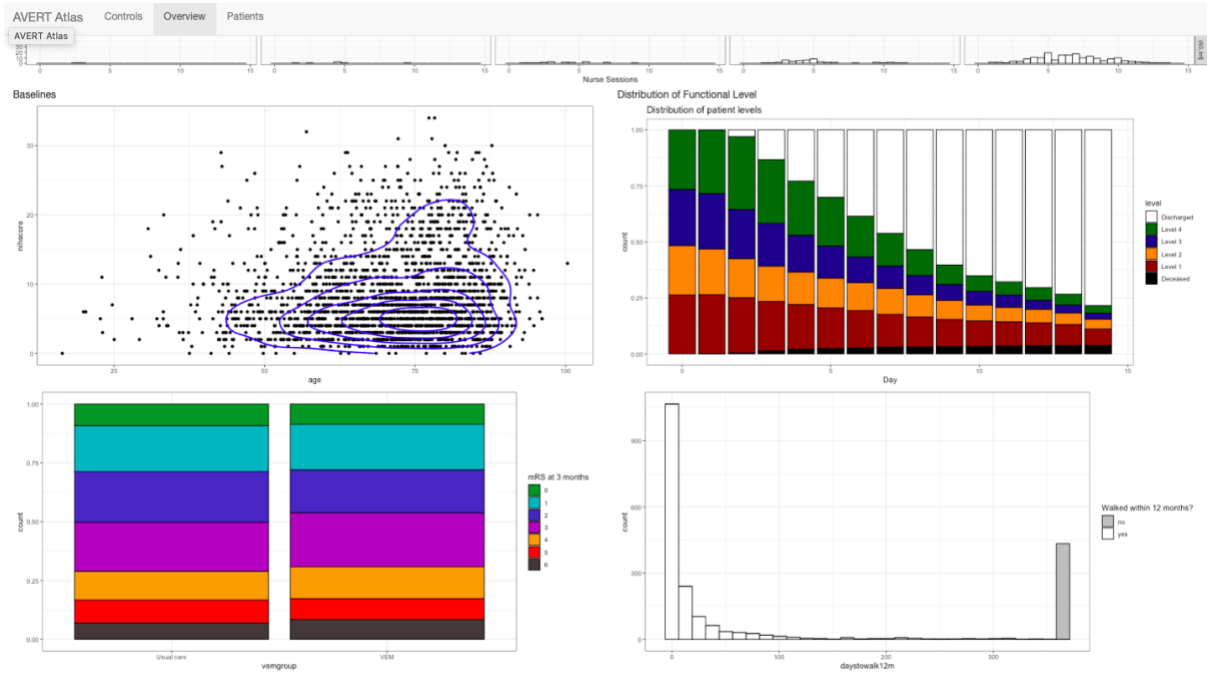
1. Can you talk me through your decision-making process when thinking about mobilising a person early after their stroke?
2. When thinking about mobilising a patient early after stroke, what clinical factors or resources do you rely on for effective decision-making?
3. What non-clinical factors or resources do you rely on for effective decision-making?
4. Who do you think are responders, and why?
5. Who do you think are non-responders, and why?

Interactive Session Questions:

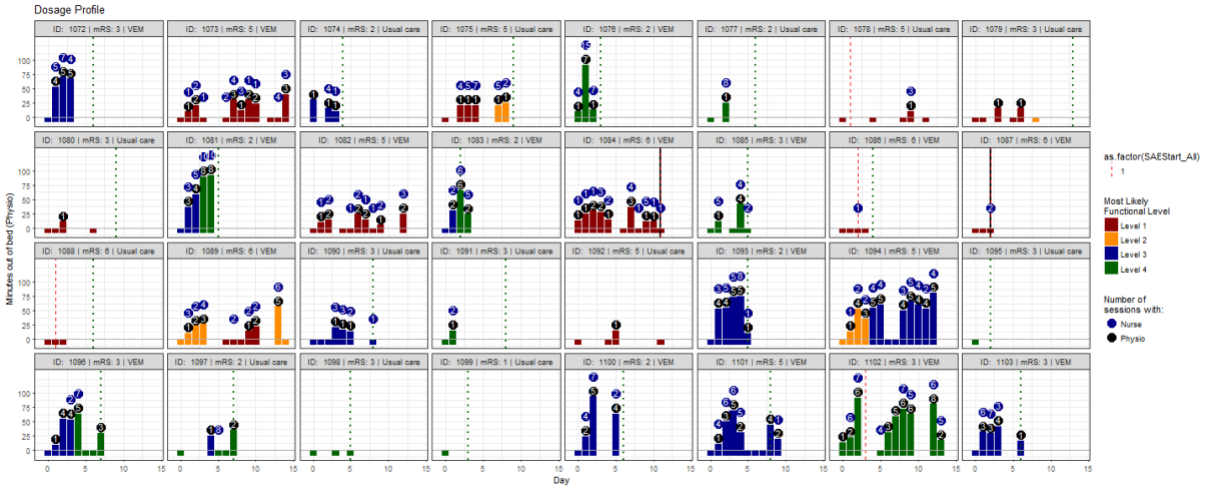
6. Which variables define the patient group of highest interest to you?
    a. Prompt: This could be a group that you believe to be a responders or non-responders or a group that you are uncertain about.
7. Why are you interested in this group of patients?
8. What outcome do you intuitively anticipate for this group and why?



Supplementary Figure E1: AVERT Atlas *Controls* interface. The figure displays the variables collected in AVERT that can be controlled and manipulated by the end user (expert stroke clinician).



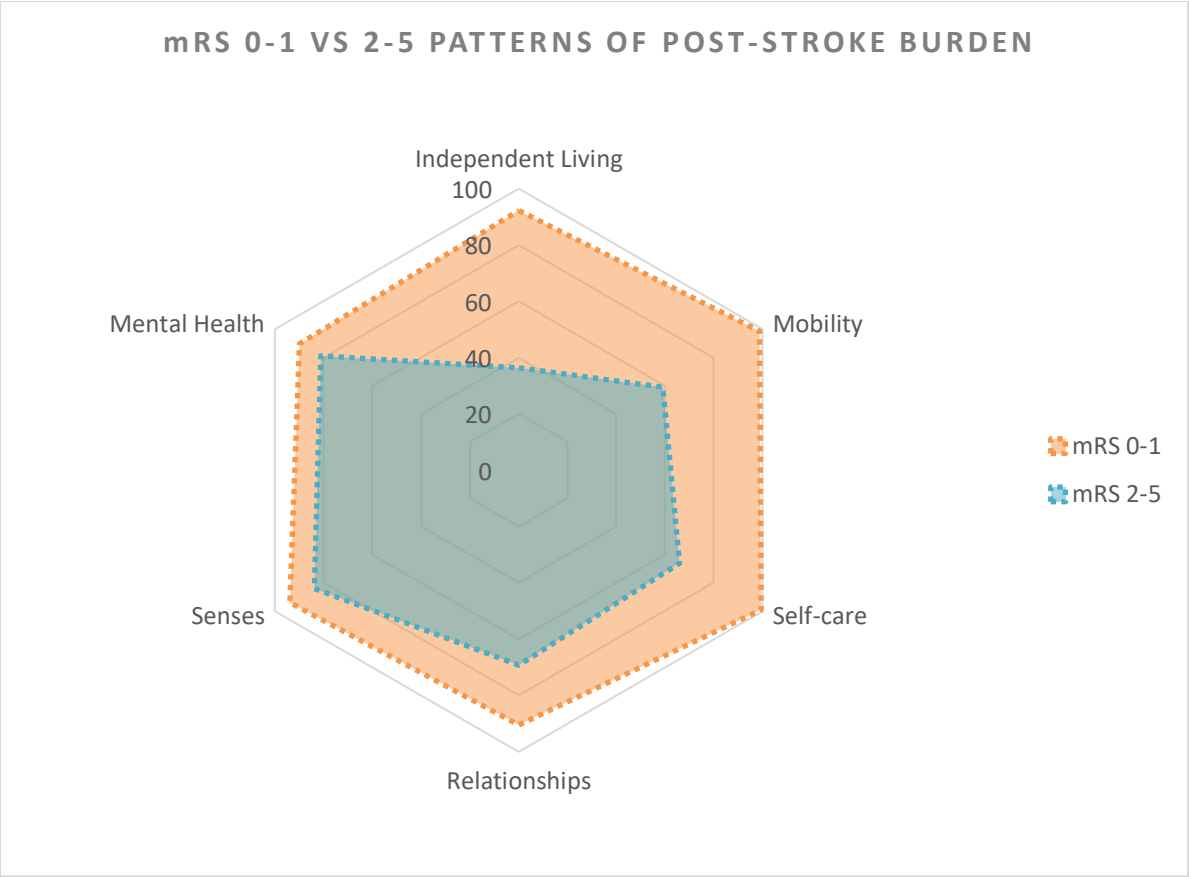
Supplementary Figure E2: AVERT Atlas *Overview* interface. The figure displays the graphical output of aggregated outcomes based on the selection of variables by the end user (expert stroke clinician).



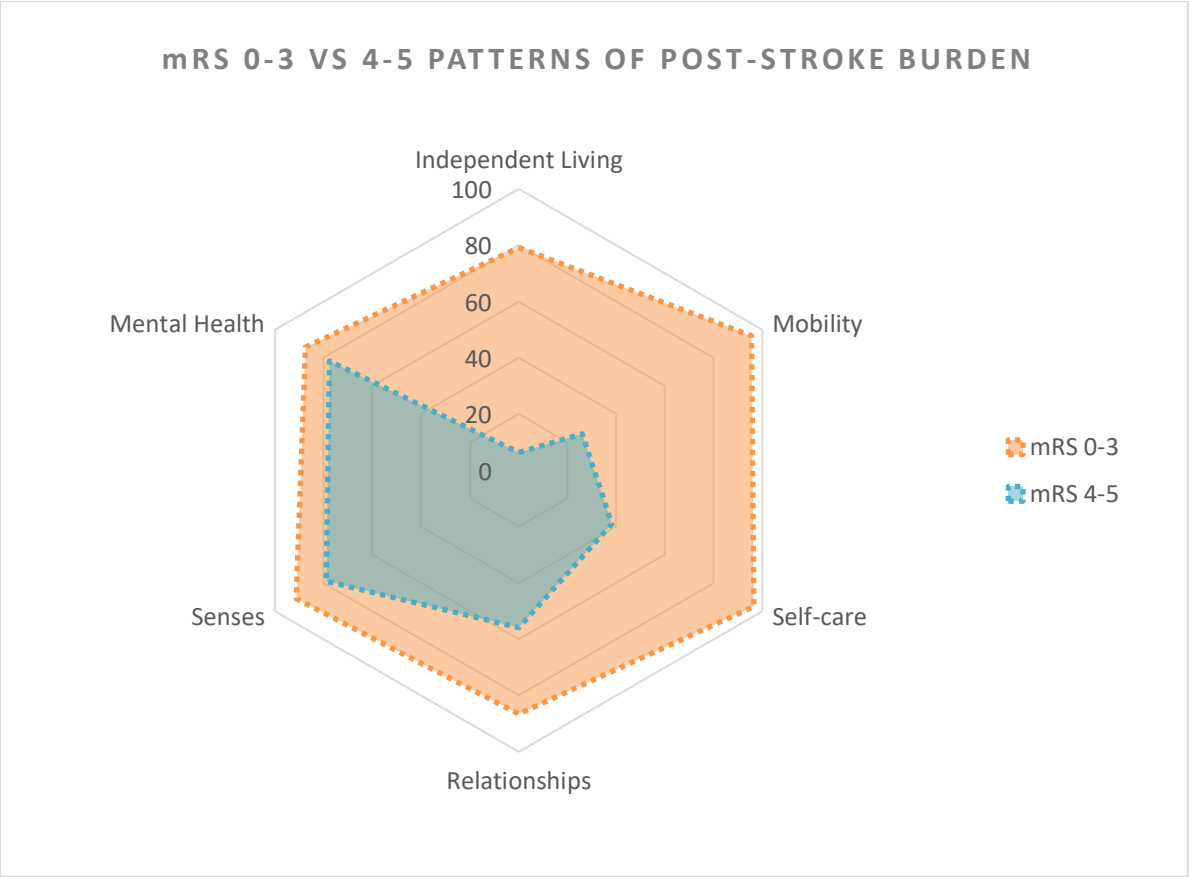
Supplementary Figure C3: AVERT Atlas *Patients* interface. The figure displays the graphical output of individual patient data based on the selection of variables by the end user (expert stroke clinician). Each tile displays the number of minutes and sessions of out-of-bed activity received by a single participant in AVERT each day until discharge or 14 days post-stroke. The colour of each bar graph represents the functional level of the patient

Appendix F

Supplementary Material: Chapter 7



Supplementary Figure F1: Radar plot displays the mean AQoL-4D and BI scores for six domains of post-stroke burden in patients with a 3-month mRS score of 0-1 and 2-5. The individual axes scaled from 0 (poor performance) to 100 (best performance) correspond to respective domains of post-stroke burden. The level of performance for all six domains are depicted using a coloured polygon whereby a smaller area of the polygon is indicative of poorer performance.



Supplementary Figure 2: Radar plot displays the mean AQoL-4D and BI scores for six domains of post-stroke burden in patients with a 3-month mRS score of 0-3 and 4-5. The individual axes scaled from 0 (poor performance) to 100 (best performance) correspond to respective domains of post-stroke burden. The level of performance for all six domains are depicted using a coloured polygon whereby a smaller area of the polygon is indicative of poorer performance.