Apathy in Stroke: Conceptualization, Measurement, and Impact

Christiane Boaventura Lourenço

B.Sc. (Physical Therapy), M.Sc (Biomedical Sciences)

School of Physical and Occupational Therapy

Faculty of Medicine, McGill University

Montreal, Quebec, Canada

August 15th , 2014

A thesis submitted to McGill University in a partial fulfillment of the requirement of the degree of the Doctor in Philosophy

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

TABLE OF ABBREVIATIONS

Abbreviations	Meaning
AD	Alzheimer's disease
ADL	Activities of Daily Living
AES	Apathy Evaluation Scale
AES- I	Apathy Evaluation Scale – informant version
AES- S	Apathy Evaluation Scale – self-rated version
AES-C	Apathy Evaluation Scale – clinician version
AI	Apathy Inventory
AIC	Akaike's information criterion
ANOVA	Analysis of variance
AS	Apathy Scale
AVC	Accident Cerebral Vascular
AVC	Accident Cerebral Vascular
BIC	Bayesian Information criterion
CCT	Case controlled trials
CFA	Confirmatory Factor Analysis
CIF	Classification Internationale du Fonctionnement, du Handicap et de la Sante
CO-OP	Cognitive Orientation to daily Occupational Performance
CPC	Category probability curve
CT	Computed tomography
CTT	Classical Test Theory
DIF	Differential item functioning
EA	Echelle d'apathie
FIM	Functional Independence Measure
fMRI	Functional magnetic resonance imaging

GBTM	Group-based trajectory modelling
GDS	Geriatric Depression Scale
ICC	Item characteristic curves
ICF	International Classification of Functioning, Disability and Health
KMO	Kaiser-Meyer-Olkin measure of sample adequacy
MHI	Mental Health Index
MOCA	Montreal Cognitive Assessment
MRI	Magnetic Resonance image
mRS	Modified Rankin Scale
NPI	Neuropsychiatric Inventory
OMS	Organisation Modiale de la Sante
PA	Parallel Analysis
PCA	Principal component analysis
PD	Parkinson's disease
PD	Proton density
PEDro	Physiotherapy Evidence Database
PET	Positron emission tomography
PSI	Person separation index
R-AS	Rasch version of Apathy Sacle
RCT	Randomized controlled trials
RMA	Rasch Measure of Apathy
RMT	Rasch Measurement Theory
RUMM	Rasch Unidimensional Measurement Model
SD	Standard deviation
SIS	Stroke Impact Scale
SMD	Standardized mean difference
SPECT	Single-photon emission computed tomography

SPGR	Spoiled gradient recalled
ssBIC	Sample size adjusted BIC
TBI	Traumatic Brain Injury
TIF	Test information function
TPC	Threshold probability curve
WHO	World Health Organization
VBM	Voxel-based morphometric
VLSM	Voxel-based lesion symptom method
VAL	Voxel-based analysis of lesions

ABSTRACT

Apathy is a primary motivational syndrome characterized by a decrease in all three domains of goal direction: behaviour, cognition, and emotion. The incidence of apathy in stroke is variable ranging from 15 – 50%. Apathetic patients are more likely to have cognitive impairment, depression, slower recovery of function, and higher dependence on others. Patients with apathy are distinguishable from those without, and are challenging for the stroke rehabilitation process, yet this syndrome is understudied in the context of stroke rehabilitation. There is a need for proper identification and measurement in order to be able to adapt rehabilitation interventions to the apathetic stroke patient's needs and evaluate their effects. Therefore, the overall objective of this study is to contribute to the understanding of the role of motivation and apathy in stroke rehabilitation by taking a longitudinal view of the key construct apathy.

There are four distinct components to this thesis. The first was a formal review of the literature to support a conceptual framework. This review indicated that the four components of apathy (openness to experience, energy level, motivation, and emotional function) mapped to the World Health Organization's (WHO) International Classification of Functioning, Disability, and Health (ICF) situating apathy within disability.

The second, third and fourth components used an existing data set of 82 people with stroke who were followed longitudinally, at four time points, from onset to one year post

stroke. Manuscript 2 contributed to the understanding of the conceptual and psychometric properties of the most used measure of Apathy in Stroke, the Apathy Scale (AS) [9]. Rasch analysis identified important psychometric limitations and the steps taken to improve these limitations resulted in a new version of the AS comprising 10 of the original 14 items reflecting only the three components of the motivation continuum. There are no items reflecting the emotional component, therefore this new version is best described as a measure of motivation symptoms.

The third study aimed to identify items from closely related constructs that could improve the conceptual limitations identified in Manuscript 2 and form a valid measure of apathy. The measure emerged from this study spans the construct of apathy with improved psychometric properties.

The final study aimed at providing preliminary evidence on how these measures of motivation and apathy behave longitudinally. A systematic review conducted within this thesis identified that apathy is potentially modifiable with intervention, notwithstanding measurement limitations. Thus, this study takes an observational view of longitudinal change in apathy. Specifically, the objectives were to (i) estimate the extent to which apathy or motivation change in the first year post stroke; and (ii) estimate the extent to which motivation and apathy impacts participation in the first year post stroke. Group-based trajectory analysis revealed five trajectories of motivation and four trajectories of apathy. The majority of stroke participants showed moderate levels of apathy and

motivation that remained stable over time and had an important impact on achieving appropriate levels of participation.

In summary, both motivation and apathy are relevant constructs for rehabilitation of stroke patients. Further work in this area should focus on supplementing expert opinion on apathy by with input from patients and caregivers on the important elements and drafting questions that use the words of potential respondents. This new measure could then undergo further validation.

ABRÉGÉ

L'apathie est un syndrome de motivation primaire caractérisé par une diminution dans les trois domaines de la direction des objectifs : le comportement, la cognition ainsi que l'émotion[1]. L'incidence de l'apathie chez les personnes ayant eu un accident cérébral vasculaire (AVC) varie entre 15 et 50% [2-4]. Les patients apathiques sont plus susceptibles d'avoir des troubles cognitifs, des dépressions, une récupération plus lente de la fonction ainsi qu'une plus grande dépendance des autres [5-8]. Les patients atteints d'apathie se distinguent de ceux sans et représentent un plus grand défi dans le processus de réadaptation suite à un AVC, mais ce syndrome demeure encore peu étudié dans ce contexte. Il y a donc un besoin pour une identification et une mesure adéquate afin de pouvoir adapter les interventions en réadaptation aux besoins des patients atteints d'apathie suite à un AVC et d'évaluer ses effets. Par conséquent, l'objectif global de cette étude est de contribuer à la compréhension du rôle de la motivation et de l'apathie dans la réadaptation d'un AVC en considérant une vue longitudinale du concept clé de l'apathie.

Il y a quatre composantes distinctes à cette thèse. La première était une revue formelle de la littérature afin de supporter le cadre conceptuel. Cette revue a indiqué que les quatre composantes de l'apathie (l'ouverture à l'expérience, le niveau d'énergie, la motivation et la fonction émotionnelle) sont représentées dans la Classification

internationale du fonctionnement, du handicap et de la santé (CIF) de l'Organisation mondiale de la Santé (OMS) et situe l'apathie comme étant un handicap.

Les deuxième, troisième et quatrième composantes ont utilisé une base de données existante incluant 82 personnes ayant eu un AVC et qui ont été suivis longitudinalement, à quatre points dans le temps, de l'apparition jusqu'à un an post-AVC. Le manuscrit 2 à contribuer à la compréhension des propriétés conceptuelles et psychométriques de la mesure la plus utilisée d'apathie suite à un AVC, l'échelle d'apathie (EA) [9]. Les analyses Rasch ont été appliquées et de la démarche prise afin d'améliorer les propriétés de mesure de l'EA a résulté une nouvelle mesure comprenant 10 des 14 items originaux, reflétant trois des quatre composantes essentielles du concept de l'apathie. Ainsi, le concept latent de la motivation est mieux représenté.

La troisième étude visait à identifier les items étroitement liés au concept qui pourraient améliorer les limites conceptuelles identifiées dans le manuscrit 2 et former une mesure valide de l'apathie. La mesure qui a émergé de cette étude couvre le concept de l'apathie et ce, avec de meilleures propriétés psychométriques.

La dernière étude visait à procurer des évidences préliminaires sur comment ces mesures de la motivation et de l'apathie se comportent longitudinalement. Une revue systématique de la littérature réalisée dans le cadre de cette thèse a identifié que l'apathie peut potentiellement être modifiable avec une intervention, nonobstant les limitations de la mesure. Cette étude observe les changements longitudinaux de

l'apathie. Spécifiquement, les objectifs étaient (i) d'estimer la mesure dans laquelle l'apathie ou la motivation change dans la première année post-AVC; et (ii) d'estimer la mesure dans laquelle la motivation et l'apathie ont un impact sur la participation dans la première année post-AVC. Des analyses de trajectoire par groupe ont révélé cinq trajectoires pour la motivation et quatre trajectoires pour l'apathie. La majorité des participants démontraient des niveaux modérés d'apathie et de motivation qui sont demeurés stables au fil du temps, mais qui n'ont eu aucun impact sur la participation.

En résumé, tant la motivation que l'apathie sont des concepts pertinents pour la réadaptation de patients atteints d'un AVC. Parce que les deux nouvelles mesures présentées ici sont une amélioration de l'EA originale, l'une ou l'autre serait facile à incorporer dans la pratique clinique ou la recherche. La poursuite des travaux dans ce domaine devrait se concentrer sur le complément de l'opinion des experts avec la contribution des patients et de leur aidant-naturel sur les éléments importants et rédiger des questions qui utilisent les mots de répondants potentiels. Une mesure moderne pourrait ensuite soutenir une validation supplémentaire.

PREFACE

Motivation has been identified by clinicians as one of the most important factors predicting return to social activities [1], functional abilities [2-4], outcomes of physical therapy [4], rehabilitation [5], and of long-term functional recovery [6-8]. However, there is a lack of consensus on a definition and classification of patients' level of motivation. In the scientific field, impaired motivation is identified as apathy, however, the definitions, the nature of the syndrome, such as its cause and course after stroke, and the extent to which it impacts functional recovery is still not clear for rehabilitation professionals.

This thesis present work carried out to understand apathy in stroke survivors. A review of the available apathy literature identified the need for a strong conceptual model for the apathy construct, and this work enabled the evaluation of the main available measure of apathy in stroke and structured the analysis that was the basis for this thesis.

Thesis Organization and Overview

This thesis consists of a series of four manuscripts and it is organized according to the regulations of McGill University's Faculty of Graduate and Postdoctoral Studies for manuscript-based thesis. These regulations are provided hereafter:

McGill University Regulations for Manuscript-Based Theses

"As an alternative to the traditional thesis format, the thesis research may be presented as a collection of scholarly papers of which the student is the author or co-author; that is, it can include the text of one or more manuscripts, submitted or to be submitted for publication, and/or published articles reformatted according to thesis requirements as described below. Manuscripts for publication are frequently very concise documents. The thesis is expected to be a more detailed, scholarly work than manuscripts for publication in journals, and must conform to general thesis requirements. Note: These papers cannot alone constitute the thesis;

The thesis must contain additional text that will connect them, producing a cohesive, unitary focus, and documenting a single program of research. A Manuscript- (or Article-) based thesis will be judged by the examiners as a unified, logically-coherent document in the same way a traditional thesis is judged.

The structure for the manuscript-based thesis must conform to the following:

- *Just as in the traditional format, the thesis must be presented as a unified whole with respect to font size, line spacing and margin sizes.*
- *The thesis must conform to all other requirements listed under Thesis components above.*
- *The thesis must be more than a collection of manuscripts. All components must be integrated into a cohesive unit with a logical progression from one chapter to the next,*

providing a cohesive, unitary focus, documenting a single program of research. Connecting text must be provided so that the completed thesis functions as an integrated whole.

- *There is no specified number of manuscripts or articles required for a Master's or a Doctoral thesis, nor is prior publication or acceptance for publication of the manuscripts a requirement. Publication or acceptance for publication of research results before presentation of the thesis in no way supersedes the University's evaluation and judgment of the work during the thesis examination process (i.e., it does not guarantee that the thesis will be found acceptable for the degree).*

An outline of the organization of the thesis follows. The introduction provides a brief overview of the thesis topic, rationale and global objectives.

Chapter 1 provides background on substantive and methodological topics covered within the thesis. It summarizes the scientific literature on apathy construct, more specifically apathy in Stroke: definition, conceptual limitations, incidence, and impact. This chapter also presents background information on theoretical model of rehabilitation and on measurement of apathy, with focus on methodological issues relevant to the current project.

Chapter 2 presents Manuscript1 entitled 'Can we modify Apathy? A structured review'. This manuscript presents the results of the extent to which apathy can be modified by

pharmacological and non-pharmacological interventions in people with neurological disorder. The results are summarized across studies that targeted the outcome of interest.

Chapter 3 presents the objectives of all manuscripts motivated by the material presented in the Introduction and Chapters 1 and 2.

Chapter 4 presents Manuscript 2, entitled, 'Improvement of the measurement properties of the Apathy Scale using Rasch Analysis. This manuscript answered questions about the content validity of the Apathy Scale by classifying the items to a World Health International Classification of Functioning Disability and Health (ICF) and applying Rasch analysis.

Chapter 5 presents Manuscript 3, entitled 'Improving the measurement of Apathy using Rasch analysis: An example from stroke'. This manuscript aimed to estimate the extent to which items from a closely related construct can fulfil the gap in the linear continuum identified by the second manuscript. We comprehensively described the methods applied for the selection and analysis of additional Items using factor analysis and Rasch analysis.

Chapter 6 presents Manuscript 4 entitled "Identifying and characterizing trajectories of apathy in stroke: Impact on functional recovery". Trajectory modelling identified sub-groups of patients with different patterns or trajectories of apathy over the four time

points (3, 6, 9, and 12 months). Joint trajectory modelling allowed the estimation of the impact of apathy on functional recovery post stroke, more specifically on participation.

Chapter 7 presents a linking chapter with the results of preliminary Brain lesion analysis. This is a chapter, not a manuscript.

Chapter 8 presents the Summary of Results, Discussion and Conclusion. The findings are summarized across studies and discussed in relation to previously reported research. The lessons learned are presented and avenues for future research are suggested.

Tables, figures, and appendices are presented at the end of each manuscript, and are embedded within the text in each chapter. A list of all tables and figures are listed in an index following the table of contents. A table of abbreviations is located just before the thesis introduction. References for all manuscripts are found following Chapter 8.

Contribution of Co-Authors

For all four manuscripts (chapter 2,4-6), the candidate conceptualized the research questions, performed all statistical analysis, and wrote the manuscripts with feedback provided from the supervisor, Dr. Nancy Mayo, and from Dr. Lois Finch. The data from Manuscripts 2, 3, and 4 came from previous study that evaluated stroke depression.

The candidate used these data for the concept analysis, the development of the measure, and to model the longitudinal change over the first year post stroke.

As supervisor, Dr. Nancy Mayo over saw all aspects of the PhD project design, analyses and writing of the manuscripts. She also provided expertise regarding advanced statistical techniques such as Rasch analyses and group based trajectory modelling.

Dr. Lois Finch collected the original data used in Manuscripts 2-4 and provided expertise on the methods and procedure, qualitative information obtained during the data collection, and clinical experience in stroke rehabilitation. She also provided expertise regarding Rasch Analysis.

Dr. Lesley Fellows provided expertise on neurology, on the methodology and analysis of brain lesion mapping. Alyssa Herzing, a PhD candidate in the health outcomes unite provided expertise feedback on the conceptual development and selection of emotional function items analysed in Manuscript 3.

Statement of Originality

The studies presented in the thesis' manuscripts are the result of my original work with guidance and feedback from my thesis supervisory and graduate committee. The theme for my doctoral research came from my experience as a physiotherapist and my

masters' research on motivation and virtual reality. I have always had interest in the topic of motivation, so this thesis topic was a fascinating one to pursue. Although clinicians and researchers acknowledge the importance of motivation to rehabilitation, they have not adequately defined and quantified motivation.

The main contribution of this thesis was to increase the awareness and understanding of the impact of apathy and motivation on stroke rehabilitation. My comprehensive review of the apathy literature identified important limitations regarding the construct definition and measurement which led to the development of a new version of apathy scale that is in accordance with the conceptual model of Apathy syndrome and shows good psychometric properties according to the Rasch measurement model. It is this comprehensiveness and conceptual precision that renders the manuscripts and the thesis fairly unique.

In order to develop a more adequate measure of apathy and handle longitudinal behavioural data, I acquired the necessary training in Rasch Analysis through a 2-week introductory and intermediate course at the University of Leeds, England. I also learned the necessary analytical skills in longitudinal modelling, more specifically, trajectory modelling that allowed me to model patterns of behaviour over time.

By appropriately measuring the level of apathy and motivation in individuals after stroke and understanding its impact on stroke recovery, effective interventions can be developed.

I also learned theVoxel-based lesion–symptom mapping methodology, with the help of Dr Fellows, to run the analysis required to investigate the relationship between lesion location and the apathy-motivation continuum,

This thesis constitutes original work as the studies presented herein were designed and overseen by the doctoral candidate to answer questions not previously addressed within the apathy literature. The originality was enhanced by the incorporation of health behaviour theory and measurement methodology, and the use of advanced analytical techniques for longitudinal data. The doctoral candidate was responsible for the development of the study research question, statistical analysis, data interpretation, and the writing of the manuscript. The candidate used an existing data source generated from a funded study on depression post-stroke. Using existing data is both efficient and ethical as time and money has already been spent to amass these data; as much could be learned from this existing data source, it was not ethically justified to collect new data from this vulnerable and ill population solely for the purposes of research. The results presented in this thesis would guide future studies.

Acknowledgments

I would like to acknowledge the financial support that I received during my PhD program from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil.

I would like to thank first of all my thesis supervisor, Dr. Nancy Mayo, for all the guidance, support and the great knowledge in research, methodology and statistical analysis. But most importantly, I would like to thank her for accepting me as a new student in a very delicate situation, giving me the chance to learn and be a part of her research team. This was the best thing that happened to me during my PhD. I can't even estimate how much I have learned!

I would like to thank Dr. Johanne Higgins for her input and feedback on the research project.

I sincerely thank, Dr. Lois Finch for all the guidance, knowledge and time you she gave me since the beginning; and Susan Scott for all the help with statistical analysis, especially for always being available to save me from SAS, even from far away.

A special thank you to Sabrina Figueiredo, for being there for me in the past 6 years, you were the best friend I could have away from home, always helping me to stay emotionally under control in all the areas of my life; Melanie Baninã, thank you for your amazing friendship and big heart and for helping me get through my PhD, especially when I needed the most, the first year; Riany Sena for making this journey much more fun and enjoyable, and for always being there for me; Skye Barbic for all your help, support, and input on the thesis process; Ana Maria for enriching my cultural knowledge, and for always saving me when I needed the most, even with so many stupid questions. I have learned a lot from you. Special thanks to all my colleagues and

fellow graduate students for support and fun moment in the lab: Ayse Kuspinar, Shahnaz Shahrbanian, Marie-Eve Letellier, Alaa Arafah, Stanley Hum, and Revital Hacmon.

To Dr. Eva Kehayia, for being like a mom to all of us; and giving me the strength and support when I needed the most, and for helping me find a new way to get through my PhD.

To my parents, for always believing in me and providing all the necessary support, and more, to make sure I achieved my dreams and goals. I wouldn't have made it if it wasn't for them. I can't thank enough!

Finally, a very special thanks to my husband, Lucas. First of all for accepting to follow me in this McGill journey and leave everything behind in Brazil, not knowing what was expecting us here in Montreal. For all the patience, for understanding all the times I couldn't be there (and there were many), and for being my *"porto seguro"*, helping me get through the most difficult obstacles and moments.

INTRODUCTION

Annually in Canada, there are 50,000 new cases of stroke [9], making stroke the third leading cause of death among adults. Although statistics show a decrease in stroke incidence and mortality over the past 10 years[9-11], the occurrence is still high. As a result of the increase in survival after stroke, the prevalence of people living with stroke related disabilities has increased, only 10% of stroke survivors recover completely and 75% of the survivors will have minor to severe impairments and disabilities [9].

The most common and well known impairments after stroke affect the motor, sensory and language systems [12]. However, impairments of emotion, mood, and behaviour are also common after stroke, both as primary *sequelae* and arising from stroke related disability. These are as important as classical motor impairment and are usually under recognized by clinicians [13]. One of the consequences of stroke is lack of motivation or apathy, which have been reported to negatively affect return to social activities [1], functional abilities [2-4], outcomes of physical therapy [4], rehabilitation [5], and long-term functional recovery [6-8].

The primary goal of stroke rehabilitation is to achieve a high level of independent functioning, either by recovery of motor control or by developing new movement pattern to compensate for impairments limitations[14]. The process of recovery is usually slow and lengthy, lasting from 2 months to more than a year. It also involves increase in

range of rehabilitation services, from acute care rehabilitation hospitals, home, and community. As a result, people with stroke experience many new situations (environmental, function, uncertainty) which can be quite overwhelming. It is, therefore, reasonable to expect a decreased interest in participating actively in therapeutic interventions. It is not difficult to understand why it is so hard to motivate overwhelmed patients and engage them in therapy.

However, studies have shown that the patient has an important role in the recovery process, and that recovery success is dependent on his/her dedication, persistence, motivation and endurance with regard to therapy. People need to want to recover and need to put a great deal of effort into therapy [15]. Rehabilitation professionals believe that motivation is a very strong determinant of rehabilitation outcome [15, 16]. A motivated patient is expected to have a positive attitude towards the recovery process, no matter how slow, frustrating and painful the rehabilitation process may be; or how much psychological and social stress is caused by their disability [15]. In fact, some professionals may avoid working with unmotivated patients. Kaufman and Becker [15] showed that when patients were identified as highly motivated, they received more time and effort from rehabilitation professionals than did unmotivated patients who were often characterized as lazy. In addition, unmotivated patients may be discharged early for lack of participation and might not have the same opportunity for functional recovery as the more motivated ones.

Even though there is enough support in the literature that motivation plays an important role in functional recovery, it is not clear how rehabilitation professionals define the concept. According to Maclean et al, [17, 18] rehabilitation professionals define level of motivation on the basis of the patient's behaviour and engagement with rehabilitation. Patients are considered motivated when they are proactive, ask relevant questions, understand the rehabilitation process, take initiative, for therapeutic exercises and work with the therapist to achieve rehabilitation goals. On the other hand, unmotivated patients are passive, pessimists, lack interaction with staff, lack interest in rehabilitation, or refuse to participate. These classification criteria are not very accurate. There is a need for better methods for identifying motivation.

The scientific literature uses the term apathy to define impaired motivation, and a variety of different definitions regarding the essential features of apathy have been proposed. The definitions, nature, cause, and course after stroke remain unclear as does how it impacts on functional recovery . Rehabilitation professionals treating people with stroke are always challenged to identify motivating strategies and effective interventions that will result in functional gains in real-world activities. A better understanding of apathy can lead to proper assessment and prevent misjudging patients as lazy, insensitive, uncaring, challenging, depressed, or even demented. It will also help plan rehabilitation according to the patient needs; increasing their chances to achieve optimal functioning[19].

RATIONALE

Motivation drives the choice of certain behaviours or actions over others. It directly affects the initiation, direction, intensity, and persistence of behaviour and can contribute to the achievement of desired goals in life. Studies have shown that stroke patients with apathy have decreased initiative or goal-directed behaviour and an absence of emotional response to events related to their needs or goals [20]. In order to benefit fully from stroke rehabilitation patients need to be an active participant. Therefore, an apathetic stroke patient will pose challenges for adapting a rehabilitation program to his/her needs. However, more needs to be learned about the nature of this syndrome, its cause and course after stroke, and to what extent stroke functional recovery is impacted by apathy. Before we can think of how to improve apathy or motivation, measures are needed. The results from this study provide two measurement options that are superior to the Apathy Scale (AS).

CHAPTER1–LITERATURE REVIEW

To ensure the literature review was comprehensive, a scoping review was carried out based on articles obtained from an extensive literature search strategy, initially planned to answer another specific structured review study question (see Chapter 2). The search was conducted in 4 different databases (Medline, Embase, Psychoinfo and CINAHL) using the following keywords and terms specific to each database: cerebrovascular disorder, Parkinson disease, Alzheimer disease, Brain injury, TBI, physical therapy, rehabilitation, physical therapy modalities, recovery of function, behaviour therapy, cognitive therapy, goal setting, electrotherapy, motivation therapy and apathy (Table 2.1).

In order to identify all the articles on construct definition, measurement, prevalence, lesion location, and clinical correlates, papers were selected from a total of 566 results prioritizing most recent publications of observational or experimental studies and review studies that involved apathy and cerebrovascular disorders. For the structured review that will be published, Manuscript 1, updates prior to publications will be carried out.

Construct Definition

The word apathy originated from the Greek word “apathes” (απαθής) which means “without feeling”. The first entry of the word apathy in the Oxford English Dictionary is from1603, where it is defined as “freedom from, or insensibility to, passion or feeling;

passionless existence". The current Oxford Dictionary definition reflects a more modern understanding: "indifference to what is calculated to move feelings; or to excite interest or actions".

Apathy is the term used to describe the negative side of motivation. When symptoms of lack of initiative, energy, persistence, and drive are observable in a person, he/she is identified as apathetic. Apathy can be a trait, usual characteristic of the individual (i.e. a history of lifelong passivity, low role activity, low self-stem, and low life satisfaction). Or apathy can be a state arising from a temporary adaptation to major changes in life (e.g. personal tragedy, natural catastrophe, social loss, and environmental deprivation).

There is also a movement to declare apathy as a syndrome with a specific set of diagnostic criteria [21]. Apathy as a syndrome was first operationalized by Marin [21], who defined it as a primary motivational syndrome characterized by impairments in all three domains that direct a goal: 1) goal-directed behaviour manifested by a lack of productivity, effort, initiative, and persistence in activities; 2) goal-directed cognition evidenced by a lack of interest and concern about their needs and goals; 3) emotional responses to goal-directed behaviour characterized by the absence of feelings. He emphasized that some clinical disorders can exhibit symptoms similar to the apathy syndrome and lead to misdiagnoses. Distinguishing the apathy syndrome from other psychiatric disorders (e.g. dementia, depression, and delirium) is of great importance. The key factor is to identify whether the symptoms are not a consequence of cognitive

impairment, emotional distress, or diminished level of consciousness. For example, if symptoms of lack of motivation are associated with emotional distress (i.e. the person is dysphoric about his decreased interest towards activities) then apathy should be regarded as a symptom of the syndrome of depression, rather than the apathy syndrome.

Based on Marin's criteria, a new standardized set of diagnostic criteria was proposed which required the presence of at least one symptom fitting to each of the three goal-directed domains: 1) behaviour (e.g. lack of effort or energy and dependency on others to structure daily activities), cognition (e.g. lack of interest in new experience and lack of concern about personal problems), and concomitants of behaviour (e.g. flat affect and lack of emotional responsiveness to events). Additionally, these symptoms should persist for a minimum of four weeks during most of the day and should not be a consequence of decreased level of consciousness or substance reaction.

There have been two main modifications of Marin's original criteria. Starkstein [22] disagreed with the idea that the apathy syndrome should be excluded in the presence of cognitive impairments as in Alzheimer's disease (AD) and Parkinson's disease (PD). In fact, recent studies indicate that the incidence of apathy is higher in individuals with neurological disorders with some degree of cognitive impairment [23]. Thus, the criterion for diagnosing apathy only in the absence cognitive impairment has been abandoned. Another change concerns the emotional domain of goal-directed behaviour. Starkstein

[22] pointed out the lack of consensus among researchers regarding the emotional component of the apathy syndrome, and that apathy has been described more often as a disorder of behavioural and cognitive dimensions. Drawing on these ideas, he defines a third domain, labelled “concomitants of goal directed behaviour”.

More recently, Robert et al. [24], with the endorsement of the Association Françoise de Psychiatrie Biologique and the European Psychiatric Association, suggested a new definition and description of the apathy syndrome, one that aims to facilitate communication, research, and treatment. The authors draw attention to the need for new criteria as there is diversity in the definition of the apathy syndrome across studies and no supported classification system. The authors identified strong agreement among researchers on the inclusion of the domains of interest, action initiation, energy, and emotional response. An introspective domain (e.g. lack of self-awareness) was also proposed but not agreed upon. The authors revised Starkstein & Leentjens [25] criteria and proposed the most recent diagnostic criteria. In order to be diagnosed with the apathy syndrome, a patient’s clinical presentation needs to correspond to four main criteria: (i) decreased motivation in relation to previous level of functioning; (ii) impairment in at least two of three psychological functions associated with goal-directed behaviour, goal directed cognitive activity, and loss or diminished emotion; (iii) significant impairments in activity and participation; (iv) ruled out diminished level of consciousness, emotional distress, and dementia (Table 1.1).

The process of conceptualizing a complex construct such as apathy, follows a number of steps from concept identification to concept refinement (e.g. concept development, delineation, comparison, clarification, and concept correction) [26]. Apathy syndrome has been developed and submitted to critical assessment and, although there was some confusion in terminology among researchers, the qualitative conceptual analysis conducted by Robert et al. [24] provided further development and clarification of the concept [26]. Therefore, Apathy Syndrome now seems to have well established characteristics, restrictions, preconditions, and outcomes.

Measurement

In accordance with the difficulties in properly operationalizing the apathy construct, there are several limitations and challenging factors associated with its measurement. First, apathy is difficult to measure as it must be inferred from behaviours or response to rewards. These manifestations can be reported by the patient (self-report), a proxy, or clinician. However, since apathy is strongly associated with impaired insight and cognition, the reports from apathetic stroke patients might not be very reliable [27]. An assessment by a proxy (family member, friend, or caregiver) or a clinician might be a better solution to address this problem. Certainly, the proxies know more about the patients' previous behaviour and level of pre-stroke motivation, thus allowing for a more sensitive perception of changes. On the other hand, proxies' estimations can be

susceptible to bias given that they cannot have direct access to the inner thoughts of another person and can only depend on observable factors. The amount of inference that can be drawn from another person's subjective content will depend on the amount of subjective-state information shared between proxies and patients, the number of experiences proxies have in detecting changes in patient's emotional state and the proxies' level of empathy [27]. There is evidence that reports from proxies are more accurate when the construct is objective (e.g., mobility); if it is subjective, such as well-being (e.g., feeling), there is a higher chance for misinterpretation [27]. In addition, as proxies familiarize themselves with patients' conditions, their internal standards, values, and concepts may change (i.e., response shift), which can influence the appraisal of the condition when evaluated over time [28].

Another important challenge is related to the available measure. Currently, the instrument most used to assess apathy in stroke is the Apathy Scale (AS) proposed by Starkstein [29, 30]. Several other scales have been used in stroke studies such as the Apathy Evaluation Scale (AES) [31], Apathy Inventory (AI) [32], and the Neuropsychiatric Inventory (NPI) [33].

Based on the conceptual definition in the literature and extensive clinical experience, the AES was developed by Marin et al. [31] to assess the affective, behavioural, and cognitive domains of apathy in adolescent and adult populations in a variety of disorders through multiple rater sources and versions: clinician (AES-C), informant (AES-I) and

self-rated (AES-S). A high score indicates more severe apathy. The AES has been used and validated for different clinical disorders (e.g. AD, PD, stroke and TBI), and shows good internal consistency ranging from 0.86 to 0.94; good test-retest reliability; good interrater reliabilities; fair to good convergent validity and discriminant validity for the informant and clinician versions but not as favourable for the self-rated version [31, 34, 35].

The AS was developed by Starkstein et al [29] as an edited version of Marin's AES18 items version. This scale contains 14 items adapted from the AES's of which 6 are in common with the AES; however, the authors did not provide any description of the methods used for selection and verification of the item pool. It was first designed to measure apathy in PD but it has also been applied in AD, Huntington's disease and stroke. Seven out of ten studies investigating apathy in stroke have used the AS [29, 36-41]. The questions are read by the examiner to the patient who should choose one of the four response options in a four point likert-type scale: not at all, slightly, some and a lot [23] . The scores range from 0 to 42, where higher scores indicate more severity of apathy. Reliability has been validated in PD patients, good 1-week test-retest(r= 0.90), interrrater reliabilities (r= 0.81), and Cronbach's α= 0.76[29]. No other studies have attempted to replicate these results [34]. The only type of validity assessed was the predictive criterion validity in which a neurologist rated 12 patients into apathetic and non-apathetic. The results showed that the apathetic patients had higher AS score than

non-apathetic patients. The sensitivity was 66% and specificity 100%, with a cut-off score of 14. The convergent and discriminant validity has not been assessed [34].

The NPI was primarily designed for use in dementia, but has also shown applicability for other neurological disorders, to assess and quantify neurobehavioral disturbances of patients and caregiver distress [33, 34]. Apathy is one of ten psychopathologic disorders examined (e.g., delusion, agitation, dysphoria, anxiety, euphoria, disinhibition, irritability and aberrant motor activity).The NPI-apathy has four screening questions that assess frequency and severity of patient's changes in interest, engagement in activities, and responsiveness (apathetic or indifferent) since the disease onset. The screening questions are directed to the informant (caregiver, family). The score on the NPI apathy subscale range from 0-12, with higher scores indicating more severe apathy. The full version of the scale has been validated in many studies with different samples, showing good internal consistency (α = 0.81-0.88), test-retest (r=0.74), and interrater reliabilities (r=0.89). However, this was not done for the apathy subscale [34]. The validity of the subscale was assessed by Cummings et al. [33] using ratings from a panel of international experts in the assessment and treatment of neurobehavioral disturbances in dementia. Discriminant validity have not been reported [34].

The AI was designed by Robert et al [32] to assess apathy in people with brain disorders. It is based on the model of the NPI scale in combination with the three dimensions suggested by Marin's criteria (e.g., goal-directed behaviour, goal-directed cognition, and

emotional responses) that were here denominated emotional blunting, lack of initiative, and lack of interest, respectively. There are three versions of the scale: caregiver, patient and clinician. Good reliabilities have been demonstrated for the caregiver (internal consistency = 0.84; interrater reliability, r=0.99; test-retest reliability ranging from 0.97-0.99 for each domain of apathy being assessed) but not for the patient version. Convergent validity with NPI-apathy was moderate for the caregiver version and poor for the patient version.

Although all these available measures have some psychometric support, there is still no gold standard measure of apathy that meets all psychometric requirements, particularly with respect to content validity and reliability. In addition, they were developed on the basis of the classical test theory (CTT).The classical test theory (CTT) is the more traditional psychometric method adopted by the majority of health researchers which are based on statistical approaches such as correlational or descriptive analysis to assess scaling assumptions, reliability and validity [42]. The methods include descriptive statistics, missing data, correlations, item redundancy, endorsement frequencies, factor analysis, Cronbach's alpha per scale, among others. Although widely used, this psychometric approach has some important limitations that should be taken into consideration. Cano et al.[43] has outlined four main limitations: (i) the metric originates from ordinal data that are summed to a total score which remains ordinal, assuming that each item contributes equally to the total score; (ii) the score for the person and the

sample under study dependent on what measure was used; (iii) reliability and validity are sample dependent; and (iv) data is suitable for group use and not individual patient assessment [43].

It is well known that the only type of measure that allows for a total sum of scores is interval scales in which the spaces between units are known to be equal. The interval scales give a more accurate analysis and interpretation of scores differences and changes over time. In ordinal scales, the true distance between items and between response options of the items are not known, they can vary across the index [44]. The items are scored on an ordinal scale (e.g., 3- a lot, 2- some, and 1- not at all), and then the values from each of the items are added up to form a total score. This assumes the values assigned to the response options are quantities when in fact they are numerical labels (albeit ranked) and that each item contributes equally to the total score. With ordered response options, the distance between the categories are not necessarily intervals and hence a change of one unit cannot be interpreted in the same way as change in biophysical units can be interpreted.

Another limitation of the available measures, are the wording of the items. For example, in the AS item: “Would you consider yourself apathetic”, the word apathetic might not be as common in the patient’s vocabulary as it is among clinicians, which can lead to inconsistency of responses. The wording of item response options should also be addressed so that patients are able to respond easily and accurately (e.g., how

accurate can they differentiate between “slightly” and “some”). The literature seems to indicate that response options for level of truth (i.e. response options ranging from “not at all true” to “very true”) are more difficult and challenging to answer as they represent more abstract concepts than response options of time.

All these challenges can interfere with the validity and reliability, jeopardizing the meaning of the total score across groups or individuals over time. To date, there is no single measure of the apathy construct that satisfies the properties of a true interval measure that can be sensitive to detect the broad range of apathetic presentation[19]. This definitely creates a barrier to increase research and knowledge in the field[34].

Prevalence

The prevalence of the apathy syndrome has been extensively investigated in neurological disorders, especially in Alzheimer’s’ Disease (AD) and Parkinson’s disease (PD); but there are fewer studies in stroke or Traumatic Brain Injury (TBI). The literature suggests that TBI and AD patients have the higher prevalence but there is strong evidence for stroke and PD. To estimate the prevalence of apathy in people with neurological disorders a scoping review was carried out, the search strategy is found in Table 2.1. The rates in TBI, PD, AD and stroke identified from this scoping review ranged from 14% – 81% [29, 30, 35, 45-58, 45, 47, 49-51, 53].

For the stroke population, the average rate of apathy derived from eleven different studies [30, 36-41, 59-61], was 30%, with rates ranging from15% to 55% (see Figure 1.2). As can be seen in Figure 1.1, the majority of studies (7 out of 10) [30, 36-41], assessed apathy with the Apathy Scale developed by Starkstein [29].The pooled proportion of those 7 studies resulted in even higher proportion of 38.5% (ranging from 21% to 55%) then the average of all stroke studies (30%). All the other instruments estimated prevalence ≤ 27% (Appendix A1). A 2014 study from China (n=361) estimated a prevalence at 3 months post stroke at 10%, using the AES [62]. These findings suggest that the type of measurement used to assess apathy might influences the prevalence rate in neurological disorders.

Lesion Location

Few studies have investigated the association of lesion location in the apathy syndrome after stroke. The limbic system and basal ganglia are known to regulate motivation while the control of real-life decision making is a function of the prefrontal cortex, in which the left prefrontal cortex is associated with positive goal context and the right prefrontal cortex with withdrawal behaviour [63]. Neurotransmitters such as glutamate, dopamine, and serotonin have also been associated with motivational problems [64]. The literature provides consistent evidence that the apathy syndrome is associated with those areas regulating motivation and decision making, with predominant involvement of the frontal

lobe [37, 38, 60] and the basal ganglia [36, 41, 65]. Other structures have also been associated with apathy, such as the internal capsule [30] and the temporal lobes [37]. The association with apathy and lesions in the basal ganglia has only been reported for studies conducted in Japan [37, 38]. It would be interesting to explore whether this is true for other populations.

Additionally, there were inconsistencies in the measurement techniques used to determine lesion location (e.g. computer tomography, magnetic resonance imaging and Single-photon emission computed tomography).The majority of the studies did not use an appropriate methodology to efficiently answer this question. Table 1.2 summarizes the studies that have investigated the association between stroke lesion locations and apathy, and the methodology used to answer this question.

Clinical correlates

Understanding the impact of apathy on stroke recovery is another important factor that needs attention in rehabilitation settings. Studies have shown that apathetic patients are also more likely to have low cognitive scores, more severe depression, difficulty in performing activities of daily living (ADL), higher dependence on others, and poor functional recovery [36, 61, 66, 67]. Although these effects are quite predominant, little is known about the longitudinal course of the disorder and whether apathy can change over time. Mayo et al. [61] was the first to investigate the longitudinal course and impact of

apathy on stroke functional recovery. Apathy was measured using caregiver report of behaviours. It is interesting to note that the majority of patients included in this study had minor or no apathy and that the apathy syndrome remained mostly constant over time. Only a small group of patients (6%) improved over the first year period, with the peak of recovery occurring within the first 6 months; while another small group of patients (6%) got worse over time. They observed that the main predictors of apathy were poor cognitive status, low functional status prior to discharge and high comorbidity.

This was the first longitudinal study of the course of apathy in stroke, and confirms that apathy has a strong negative impact on stroke recovery, but suggests that it may not be a modifiable construct. The authors argue that the lack of improvement might be the result of a lack of interventions specifically designed to deal with the apathy syndrome in stroke. However, to design an appropriate intervention for apathetic stroke patients, there should be first a robust conceptual model and a good measurement method. To paraphrase Lord Kelvin, we cannot measure and treat what is not fully understood.

Even though there are few studies investigating the impact of pharmacological and non-pharmacological interventions on Apathy, there has been no attempt to systematically compile literature on intervention. Therefore, it is still not clear if apathy can be modifiable when an appropriate intervention is applied.

Summary

Studies on apathy have both theoretical and measurement limitations that can lead to misinterpretation of the results and lack of generalization and comparison across studies. If scores lead to diagnostic decisions, misclassifications can lead to inappropriate treatment decisions. The literature also indicates that apathy is an important construct in rehabilitation, having a negative effect on recovery. However, there has been no attempt to systematically compile literature on intervention. The next chapter presents Manuscript 1 “Can we modify apathy? A structured review”, this structured review is part of the literature review component of the thesis and provides a stronger methodological approach to compile evidence from the available literature and to estimate the effectiveness of an intervention.

Table1.1: Robert et al. (2009)[24] proposed criteria for apathy syndrome

A Loss of diminished motivation in comparison to the patient's previous level of functioning and which is not consistent with his age or culture. These changes in motivation may be reported by the patient himself or by the observations of others.

B Presence of at least one symptom (response to internal stimuli or response to external stimuli) in at least two of the three following domains for a period of at least four weeks and present most of the time.

Domain B1: Loss of, or diminished, goal-directed behaviour as evidenced by at least one of the following:

- Loss of self-initiated behaviour (e.g. starting activities, seeking social activities, expressing choices)
- Loss of environment-stimulated behaviour (e.g. responding to conversations, participating in social activities)

Domain B2: Loss of, or diminished, goal-directed cognitive activity as evidenced by at least one of the following:

- Loss of spontaneous ideas and curiosity (interest) for routine and new events (e.g. challenging tasks, recent news, social opportunity, personal/family and social affairs)
- Loss of environment-stimulated ideas and curiosity (interest) for routine and new events (e.g. persons residence, neighbourhood or community)

Domain B3: Loss of, or diminished, emotion as evidenced by at least one of the following:

- Loss of spontaneous emotions, observed or self-reported (e.g. subjective weak feeling or absent emotion, or observation of others of a blunted affect)
- Loss of emotional responsiveness to positive or negative stimuli or event (e.g. observed-reports of unchanging affect, or of little emotional reaction to exciting events, personal loss, serious illness, emotional-laden news)

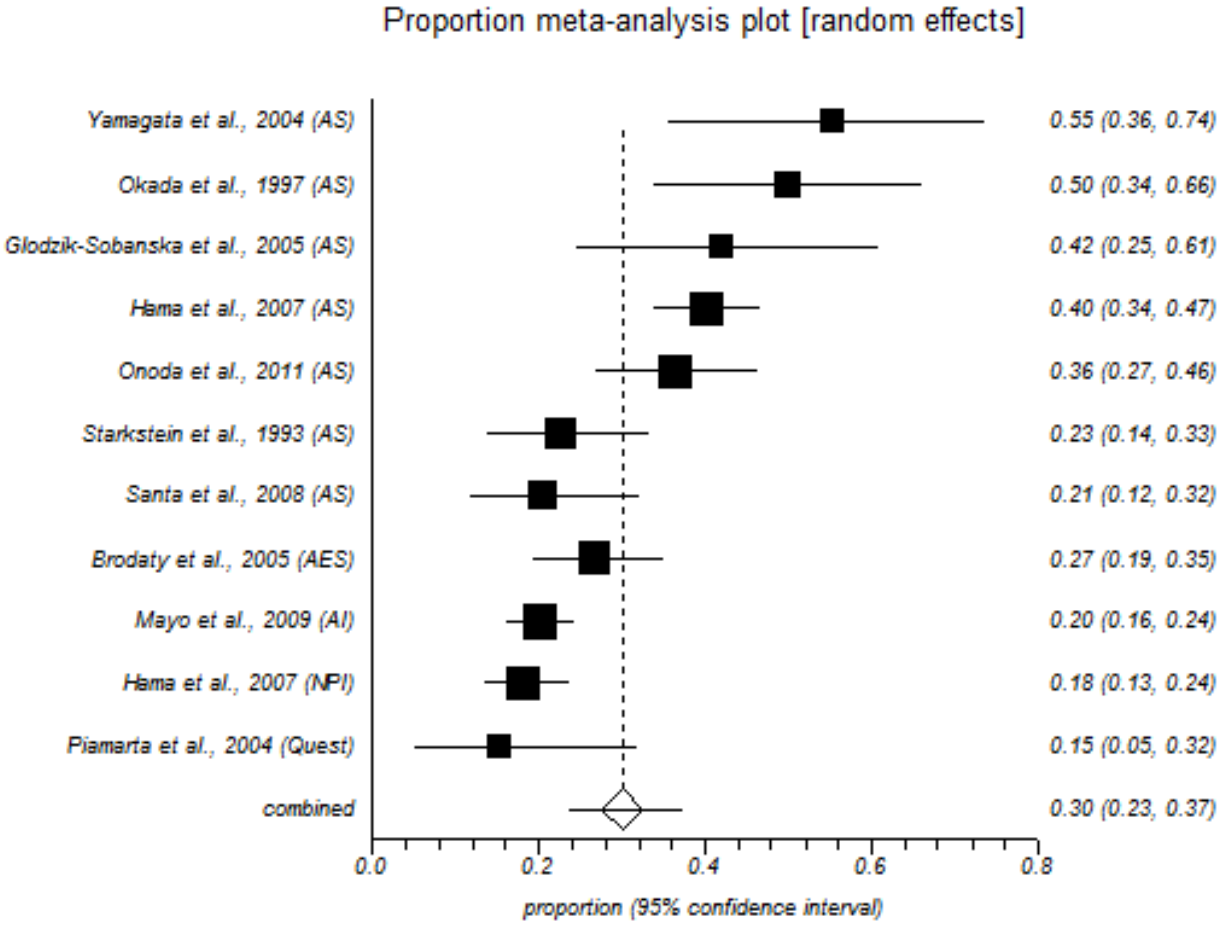
C These symptoms (A-B) cause clinically significant impairment in personal, social, occupational, or other important areas of functioning.

D The symptoms (A-B) are not exclusively explained or due to physical disabilities (i.e. blindness and loss of hearing), to motor disabilities, to diminished level of consciousness or to the direct physiological effect of a substance (e.g. drug of abuse, a medication)

Table 1.2: Studies investigating the association between apathy and lesion location

Authors	Lesion location	Image measurement
Starkstein 1993 [30]	Posterior limb of the internal capsule	CT
Okada 1997 [37]	Decreased rCBF both hemisphere	MRI, CT, and SPECT
	Right dorsolateral frontal	
	Anterior temporal	
	Left premotor area	
	Left anterior temporal regions	
Yamagata 2004 [38]	Frontal lobe	MRI
Brodaty 2005 [68]	Right hemisphere	CT, and MRI
	Right frontal-subcortical circuit	
Glodzik-Sobanska 2005 [39]	Lower NAA/Cr in right hemisphere	MRI
Hama 2007 [69]	Bilateral lesions of basal ganglia	CT
Santa 2008 [70]	Left basal ganglia	MRI
Onada 2011 [41]	Left basal ganglia	MRI and SPECT
	Bilateral basal ganglia	

Figure 1.1: Forest plot of the proportion of apathy in Stroke population across eleven studies



CHAPTER2 -MANUSCRIPT1:Can We Modify Apathy? A Structured Review

For submission to: Systematic Reviews

Authors: Christiane B. Lourenço, M.Sc.[1],Nancy E. Mayo, Ph.D.[1,2,]

[1] School of Rehabilitation Sciences, Faculty of Medicine, McGill University, 3654 Prom Sir William Osler, Montreal, Quebec, H3G 1Y5, Canada.

[2] Division of Clinical Epidemiology, Royal Victoria Hospital, McGill University Health Center, Montréal, Québec.

Communication address to: Christiane Lourenco, Royal Victoria Hospital, Division of clinical epidemiology, 687 Pine Avenue West, Ross 4.27, Montreal, Quebec, H3A 1A1 Canada. Telephone: (514) 934-1934 ext 37066, Fax: (514) 843-1493, Email: christiane.lourenco@mail.mcgill.ca

Key Words: apathy, rehabilitation, physical therapy, cognitive therapy, behavioural therapy

ABSTRACT

Objectives: To assess the extent to which Apathy can be modified by pharmacological and non-pharmacological interventions in people with neurologic disorder.

Methods: Four electronic databases were searched: Medline (1950-Present), Embase (1980-Present), PsyncINFO (1987-Present) and CINAHL (1950-Present); using a combination of Medical Subject Heading (Mesh) and keywords on apathy, neurological disorders, and interventions. Search results, data extraction, and quality were assessed by one author. Subgroup analyses were performed according to type of intervention. Results were presented as standardized mean difference (SMD) and 95% confidence interval, using either a fixed or random effect. Heterogeneity was assessed using visual inspection of forest plot and Higgins I^2.

Results: From the 457 unique citations identified in the literature search, 9 articles met the eligibility criteria (220 patients). Five studies investigated pharmacological intervention and four studies investigated non-pharmacological interventions on apathy. Of the nine included studies, four did not provide sufficient data for calculation of effect size and 95% CI. No overall effect was estimated due to substantial heterogeneity between studies. The effect size of the studies ranged from -1.42 to -0.06.

Conclusions: The result suggests that apathy could be decreased with the proper intervention but the heterogeneity and poor methodological qualities of the included studies do not allow for solid conclusion.

INTRODUCTION

Humans are moved by goals and interests which change in response to personal growth, environmental, social interests, and life experience. Intrinsic and extrinsic motivation is what accounts for these changes; it directly affects the initiation, intensity and persistence of behaviours [71]. So what happens if an individual has no motivation? Most likely he or she will show decreased goal-oriented initiative, productivity, and emotional responses towards his needs and goals [21]. This lack of motivation is defined as apathy. Apathy has been operationalized as loss of emotion, concern and self-initiated action towards one's own health, activities and important life events [21].

The rate of apathy provided in the literature varies considerably across neurological disorders, ranging from 20% to 25% in stroke [72], 33% to 45% in Parkinson's disease (PD) [63, 73], 60% in Alzheimer's disease (AD) [66], and 61% in Traumatic Brain Injury (TBI) [66]. The presence of apathy in neurological disorders has been implicated in poor functional recovery, and prognosis in general, long hospital stay, less adherence to therapy, and lack of engagement in rehabilitation interventions [63, 68, 74]. Clinicians are always challenged to identify motivating and effective interventions for the rehabilitation of neurological patients that will result in functional gains in real-world activities [75]. It is well known that in order to fully benefit from rehabilitation, patients need to participate actively during therapy; otherwise the recovery potential is jeopardized. For patients with apathy, it is even more challenging to adapt rehabilitation to these patient's needs.

Therefore, apathy needs to be properly assessed in clinical settings to provide appropriate interventions and help to achieve optimal functioning [19].

To date, interventions for apathy have ranged from pharmacological to cognitive behavioural therapy. Pharmacological interventions for apathy include Psychostimulants (i.e., Methylphenidate and Dextroomphetamine), Dopaminergic drugs (i.e., Amphetamines, Amadantine and Bupropion), Antidepressants (i.e., Bupropion, Methylphenidate and Dextroomphetamine) and Cholinergic Therapy (i.e., Tacrine and Donepazil). Non-pharmacological therapies are mainly stimulation therapies targeting the apathetic individual needs and can be delivered in different forms like cognitive therapy, behavioural therapy and goal-setting therapy.

As apathy is such a common sequelae of stroke and other neurological disorders, and can interfere with the course of recovery and rehabilitation, it is surprising that it has not been the focus of more research. There are few studies about apathy in the neurological population and the majority tend to focus on the characteristics (i.e., frequency, nature and severity) of apathy rather than on its course over time and on the effect of interventions. Most of the available studies have assessed apathy in just one point in time. Mayo et al.[61] investigated in a longitudinal study, the extent to which apathy changed over time in the stroke population, and found improvement in only 7% of the participants. According to the authors, there may be two main explanations for

this finding: nothing was offered to the patient to change apathy or apathy is not amenable to change.

This study addresses what is currently known in the literature about the extent to which apathy is a modifiable construct.

OBJECTIVES

To assess the extent to which Apathy can be modified by pharmacological and non-pharmacological interventions in people with neurological disorders.

METHODS

Criteria for considering studies for this review

Types of studies

All study designs were considered: Experimental studies (Randomized controlled trials (RCT), quasi-randomized, case controlled trials (CCT), cross-over trials, single subject designs, case report/series and observational studies), observing/analysing the effect of interventions on apathy.

Trials comparing two or more interventions and a control group were also included. For cross-over trials, data were included if there was a washout period and authors reported appropriate data.

No publication date or publication status restriction was imposed. Trials were included where participants were randomized to the following:

- Intervention group: a pharmacological intervention or rehabilitation therapy, versus
- Control: placebo or conventional therapy

Types of participants

Adults diagnosed with one of the four neurological conditions: AD, PD, Stroke and TBI. Studies have shown high prevalence of apathy for all these neurological disorder [63, 66, 72, 73], therefore, it is important to investigate if apathy can be modified independently of the neurological disorder. Participant's characteristic of interest included age, gender, type of neurologic disorder, and apathy status at entry.

Types of interventions

Any pharmacological and non-pharmacological (e.g., physical therapy, cognitive therapy, behavioural therapy) interventions that were intended to modify apathy were

included. There was no restriction on duration or frequency of intervention. These interventions were compared with the control group receiving a different medication dose, placebo, or conventional therapy.

Types of Outcome

Apathy is the primary outcome measure. Studies that measured apathy on at least two time points, using the total score of one of the following four scales: Apathy Evaluation Scale (AES) [31], Apathy Scale (AS) [29], Apathy Inventory (AI) [32], and/or apathy component of the Neuropsychiatric Inventory (NPI) [33] were included.

The AES was developed by Marin [31] to assess apathy in adolescent and adult population in a variety of disorders through multiple rater sources and versions: clinician (AES-C), informant (AES-I), and self-rated (AES-S). The aim of the AES scale is to assess the behavioural, cognitive, and emotional domains of goal direction, and has been used and validated for different clinical disorders, including AD, PD, and TBI. It comprises 18 items scored on a 4-point scale with scores ranging from 0 to 72. Higher scores indicate more severity, and a value of ≥37/38 is used to classify people as apathetic.

The NPI [33] was primarily designed for use in a dementia population, but has also shown applicability for other neurological disorders. Apathy is one of ten psychopathologic disorders examined (e.g., delusion, agitation, dysphoria, anxiety, euphoria, disinhibition,

irritability, and aberrant motor activity). The NPI-apathy has four screening questions that assess changes in the patient's interest, engagement in activities and responsiveness since the disease onset. The screening questions are directed to the informant (caregiver, family). The scores range from 1 to 12 in each domain. According to the apathy operational definition, score of at least four (frequency score =4; severity score 1-3) would indicate presence of apathy, however, there is no evidence in the literature for this information. Higher scores indicate more frequency and severity.

The AI was designed by Robert et al. [32] to assess apathy in people with brain disorders. It is based on the model of the NPI scale in combination with the three dimensions suggested by Marin's criteria (e.g., goal-directed behaviour, goal-directed cognition, and emotional responses). There are three versions of the scale: caregiver, patient and clinician.

Finally, the AS was developed by Starkstein et al. [29] as an edited version of Marin's AES. This scale was first designed to measure apathy in PD but it has also been applied in AD, Huntington's disease, and stroke. This scale contains 14 items adapted from the AES's 18 items. The questions are read by the examiner to the patient who should choose one of four response options in a four point Likert-type scale: not at all, slightly, some and a lot. The scores range from 0 to 42 (cut-off ≥ 13/14), where higher scores indicate more severity.

These scales were chosen due to their frequently use by studies investigating apathy in neurological populations, and have shown validity and reliability for measuring the apathy construct in all the four pathologies included in the present study [29, 30, 45, 54, 58, 60, 76]. There was no restriction for the follow-up period. If the study did not investigate apathy construct or apathy was not measured by one of the four measurement scales, it was exclude.

Search methods for identification of studies

The current research was conducted in July 2011, with the last search run in July 18, 2011. Can you update this and add a note Relevant studies reported in English, Portuguese and Spanish languages were identified from the following four electronic databases: Medline (1950-Present), Embase (1980-Present), PsyncINFO (1987-Present) and CINAHL (1950-Present). The search strategies used a combination of Medical Subject Heading (Mesh) and keywords, and are represented in Table 2.1. Other resources searched for studies that met inclusion criteria were reference lists of all included trials and key publications.

Data collection and analysis

Selection of studies

Results of database searches were imported into EndNote reference manager. One author assessed eligibility by screening all titles and abstracts of all citations identified in the search as potential studies. After potential studies have been selected, full reports were obtained to assess the inclusion criteria using the pre-designed selection criteria. All papers were screened for multiple publications, and studies were excluded if they did not meet the criteria. All reasons for exclusion were documented.

Data extraction and Analysis

One author extracted data from each report consisting of the following: study author, type of study, intervention applied (dose, frequency of administration and duration of treatment), population (which of the four neurological conditions), sample size, age, sex, primary outcome relevant to measurement of apathy (Table 2.2), and results of statistical analyses for apathy component (mean and SD of each group; and/or mean difference and SD of baseline). If data was not available the responsible author was contacted for further clarification. If data was not retrievable, the trial was excluded.

Data was analysed using Stata 11.0 [77]. Results were presented as standardized mean difference (SMD) and 95% confidence interval, using either a fixed or random effect model, depending on the statistical heterogeneity between studies. Heterogeneity was assessed using visual inspection of forest plot and Higgins I^2 which describes the percentage of total variation across studies that are due to heterogeneity rather than

chance. Mild heterogeneity accounts for less than 30% of the variation and $I^2 \geq 50\%$ is considered to be indicator of a substantial level of heterogeneity. [78] Meta-analysis for assessing the extent to which apathy can be modified by therapy was performed only if the results did not show significant heterogeneity.

To calculate the standardized mean difference (SMD), we subtracted the control mean difference (change from baseline) from the intervention mean difference and divided by the pooled standard deviation. Pooled standard deviation was calculated based on the baseline SD of each group, according to the formula described in Cochrane's handbook (see Table 7.7a in Cochrane's handbook) [79].

In case of multiple intervention groups, the data of each intervention were analysed to decide the best approach. If data from different intervention groups were homogenous and confirmed by the study results and p-values, than the intervention groups were combined to create a single pair-wise comparison, according to the formula described in the Cochrane's handbook [79]. If both interventions contributed differently to the analysis, then we would select the most appropriate pair of interventions, according to the study objective, and exclude the non-relevant (Cochrane Hand-Book [79], section 16.5.5).

Publication bias was assessed by examining funnel plots for asymmetry and the Egger method [80]. P-values less than 0.05 were considered significant.

Assessment of risk bias in individual studies

The validity of eligible studies was assessed by one reviewer. Included studies were first categorized by methodological design, and specific quality assessment approach was selected for different designs. Methodological qualities of RCTs, quasi-RCT, CCT and case series were assessed using PEDro rating scale [81]. It includes 11 items, with a total score of 10 (2-11):

- Participants eligibility criteria and source specified;
- Random allocation of participants to interventions; allocation concealed;
- Intervention groups similar at baseline regarding key outcome measures and important prognostic indicators;
- Blinding of all subjects;
- Blinding of all therapists administering the intervention;
- Blinding of all assessors who measure at least one key outcome;
- Drop-outs (attrition bias);
- Intention-to-treat analysis;
- Reporting between group statistical comparisons;
- Reporting of measures of variability;

Single-subject designs and case reports were assessed using SCED scale [82], which includes 11 items, with a total score of 10 (2-11):

- Participants clinical history specified;
- Precise and repeatable target behaviour operationally defined;
- Three phases or multiple baseline design
- Sufficient baseline sampling
- Sufficient treatment phase sampling;
- Data points reported;
- Inter-rater reliability established for at least one measure of target behaviour;
- Independence of assessor;
- Statistical analysis conducted
- Replication of results;
- Evidence of generalization.

Assessment of heterogeneity

Heterogeneity between studies was tested using visual inspection of the forest plot and a standard chi^2 considered statistically significant at $p<0.1$, after consideration of the value I^2 statistic, a value greater than 50% may indicate substantial heterogeneity.

Data synthesis

Statistical analyses were performed using Stata 11.0 [77], Quantitative data for the outcomes listed in the inclusion criteria are presented in the Analysis (Table 2.5). If appropriate, results of comparable groups of studies were pooled using a random-effects model and 95% confidence intervals calculated.

Regardless of possible heterogeneity of the included studies, separate analyses were conducted by type of intervention used (pharmacological or non-pharmacological).

It was anticipated that sensitivity analyses would be undertaken, when indicated, to investigate the effects of methodological differences and quality on the overall effect.

RESULTS

Description of studies

Figure 2.1 shows the flow diagram of the search results [83]. Of 457 unique citations identified in the literature search, 9 articles met the eligibility criteria. The studies characteristics are presented in Table 2.2.

Participants

A total of 220 individuals participated in the 9 studies included [84-92]. The largest study contained 51 participants [84] and 2 studies were single subject experiments [91, 92]. All participants were adults with age ranging from 31 to 84 years old. Neurological population varied across studies: three studies investigated apathy in Alzheimer's disease; three studies in stroke, two studies involved the traumatic brain injury population (TBI), and one a Parkinson's disease population. From the 220 individuals, 117 were males and 103 were females; and disease onset varied from acute (few days) [89-91] to chronic (up to 15 years)[84, 85, 90, 92].

Interventions

There were 5 pharmacological studies investigating the effect of different types of drugs on apathy: dopaminergic therapy using L-dopa [84] and Selegiline [90]; Psycho stimulants and antidepressants using methylphenidate [86]; Cholinergic therapy (acetylcholinesterase inhibitors) using donepezil or galantamine [89]; and nefiracetam [87].

The other four studies were non-pharmacological studies in which cognitive intervention was the most common type of delivered intervention, with three trials. One study investigated the effect of motivational interviewing and external compensation on apathy in a single TBI subject.

Settings

Trials were conducted in hospitals, health clinics, rehabilitation centers and elder centers from different countries: four in the United States and one from each of Canada, New Zealand, Australia, China, and France.

Outcomes

All the included studies assessed apathy on at least two time points. Five [86, 89-92] out of the nine included studies applied the Apathy Evaluation Scale (AES) by Marin [31]. Two studies [84, 87] assessed apathy with Starkstein's Apathy Scale (AS) [29] and the remaining two [85, 88] studies assessed using the Neuropsychiatric Inventory (NPI) apathy component.

Methodological quality of studies

The detailed results of the risk of bias assessments for each included study are summarised in Table 2.3 and 2.4. Many included trials were that not of high methodological quality and were at high risk of bias for at least one of the eleven methodological criteria (more specifically relating to lack of allocation concealment and blinding).

Five of the 7 studies assessed by the PEDro scale used randomization to assign subjects to groups, although concealment of randomization was not discussed. These five studies were included as RCTs [84-88]. Not included for quantitative analysis were one case controlled trial (CCT) [89] and one case series [90]. Allocation concealment was inadequate for all studies.

Binding of behavioural interventions is difficult, especially for participant and therapist; nevertheless, the assessor is expected to be blinded. Only one of the 2 non-pharmacological trials, Niu et al [88], stated that the evaluator was blinded to the intervention. Only two of the five pharmacological studies [86, 93] mentioned blinding in the design; these were labelled as double-blind but were not specific as to who was blind.

Overall, quality scores ranged from poor to high. One study [94] showed poor methodology (2/10), three were moderate (ranged from 4 to 6) [84, 86, 89] and only three qualified as high quality studies in which one had a score of 7/10 [93] and two of 8/10 [85, 88]. Overall, the seven studies assessed by the PEDro rating scale showed moderate methodological quality, with an average score of 5.7 (SD 2.2) out of 10. The quality of the case report study [91] and the single-subject design [92] study were assessed by the SCED rating scale. Both studies demonstrated moderate quality, with scores of 4 and 6 out of 10 respectively.

Effect of Intervention

Of the nine included studies, two [89, 90] did not provide sufficient data for calculation of effect size and 95% CI, and two studies only examined one subject (case report and single-subject design). Therefore, these four studies were excluded from the quantitative analysis. We categorized interventions into two categories. Three studies evaluated the effect of pharmacological intervention and two assessed the effect of non-pharmacological intervention on apathy.

An overall effect was not estimated due to substantial heterogeneity between studies. The non-pharmacological subgroup meta-analysis showed significant heterogeneity with I^2 statistic of 70.8%. Pharmacological studies showed lower variability, with I^2 statistic of 19.8% (Figure 2.2). Even though the I^2 statistics demonstrated mild heterogeneity, the three studies included in the pharmacological subgroup had relatively different study population criteria (PD, AD, and Stroke). Therefore, we considered the meta-analysis inappropriate for the present study. Table 2.5 presents the summary data for all included studies and Figure 2.2 shows the forest plot by type of intervention, with all five studies included in the quantitative analysis.

Niu and colleagues [88] investigated the efficacy of cognitive stimulation therapy (CST) in the treatment of neuropsychiatric symptoms, including apathy, in thirty-two Alzheimer's participants. Participants were either assigned to CST or control group, and received treatment twice a week, for 10 weeks. CST consisted of a set of tasks requiring executive functions and working memory (the reality orientation task, the fluency task,

the overlapping figure task and the photo-story learning task). The control intervention consisted of a relaxed communication exercise (psychological support) administered individually and focused on different activities (i.e. discussing recent topics and important life events, learning about progress in Alzheimer's disease research and external memory aids). The groups that received CST showed significant reduction in the NPI apathy score in comparison to control. The effect size for intervention was -1.03 (95% CI -1.8; -0.25).

In Politis et al.[85], thirty six Alzheimer's participants with apathy were randomly assigned to either 'the geriatrics network kit' (experimental intervention) or the 'one-on-one' (control). The intervention's objective was to decrease apathy through mental stimulation with activities of the participant's choice, for 30 min three times a week, over 4 weeks. Control therapy consisted of an individual, relaxing time, between therapist and patient where they could choose from a variety of activities like talk, play games, do artwork, and reading. The session lasted for 30 minutes. Both groups showed a significant reduction on NPI apathy score across time (z=-1.9, p=0.05; z=-2.7, p=0.007) but there were no significant between group differences (z= -0.52, p=0.6), therefore the effect size for the intervention was small -0.06.

Robinson and colleagues [87] investigated the effect of nefiracetam on forty-eight apathetic stroke patients that were randomly assigned to one of three interventions groups, dose of 600 mg, dose of 900 mg, or identical placebo; for twice daily during 12

weeks. As there was no difference at baseline for both intervention groups, and no difference in apathy change scores between the two dose groups (F=1.45, df=3,65, p=0.23), these two groups were combined, treated as one group, and compared to the placebo group. With these two groups combined, there was a significant effect of Nefiracetam on apathy scores, with change scores of -5.3 ($p<0.05$) with intervention compared to change of -2.0 in placebo group. The effect size for the intervention was -0.98 (95% CI-1.58;-0.39).

In Herrmann et al.[86], thirteen apathetic participants with Alzheimer's disease received treatment with methylphenidate in a cross-over trial for two weeks, with a one week placebo wash out period between each phase to avoid a carryover effect. The baseline scores on all outcomes, at each treatment phase, were similar (Wilcoxon z=-0.99, p=0.33 for AES total) confirming the absence of carryover effect. The dose started at 5 mg twice a day for 3 days and increased to 10 mg twice a day for 11 days. Apathy scores decreased -2.31 ($p <0.05$) from baseline to post-treatment compared to an increase of 0.5 in the placebo group. The effect size for the intervention was -0.62 (95% CI -1.42; 0.19).

Finally, Czernecki et al.[84], compared the influence of dopaminergic therapy on motivation of Parkinson's disease subjects using L-dopa. Twenty-three participants with apathy and twenty-eight healthy controls were assessed to evaluate the effect of medication on On-state and Off-state phases. Since there was no difference for the

order of medication, the intervention groups were combined and compared to control. Apathy scores (AS scales) decreased 2.9 points in the intervention group and 0.6 in the control group. The effect size for the intervention was -1.42 (95% CI -2.0; -0.8).

The remaining 4 studies that were excluded from quantitative analyses also showed results favourable to change in the apathy construct. Whyte et al.[89] investigated, in 26 stroke participants, the effect of acetylcholinesterase inhibitors on cognitive impairments and apathy. Participants received either galantamine (maximum dose 24mg/day) or donepezil (maximum dose 10 mg/day) for 12 weeks. The authors describe a trend favouring change in the AES apathy score for both groups (p=0.21).

Newburn and Newburn [90], in a case series, investigated the use of Dopaminergic therapy (selegiline) in four apathetic patients. The doses varied from 5mg to a maximum of 20 mg. For the majority of the participants, the symptoms decreased with the increase of the dose, and showed none or small side effects compared to other drugs like methylphenidate. AES scores at baseline (maximum of 72 indicating high apathy) were 50, 70, 46 and 72. All subjects improved in the apathy score, with post treatment scores of 27, 24, 26 and 31 respectively. The authors suggest that selegiline might be a potential apathy treatment which is well tolerated.

Skidmore et al.[91], in a case report study, investigated the effect of meta-cognitive training on cognitive impairment and apathy after stroke. A 31-year old male participated in a Cognitive Orientation to daily Occupational Performance (CO-OP).

Each session lasted 45 minutes, 5 days per week for 14 days. Level of apathy increased from 26 to 28 in the AES scale.

Lane-Brown and Tate. [92], evaluated the effect of motivational interviewing and external compensation on apathy in a 32 year old male with TBI. The participant identified three important goals (organize and maintain bedroom, increase fitness and maintain through exercise, and improve social conversation) and planned with the therapist the best approach to reach each goal. Meetings with the therapist occurred weekly, 1 hour per session over 28 weeks. According to the authors, a clinical significant change was observed on the self-rated and informant- rated versions of the AES scale, and these changes were maintained at follow up.

DISCUSSION

Summary of main findings

A structured review of the literature was performed to assess if apathy can be modified by different interventions in people with neurological disorders. Nine studies were included, five studies investigated pharmacological interventions and four were of non-pharmacological interventions. The result suggests that apathy could be decreased with the proper intervention but the heterogeneity and poor methodological quality of the

included studies do not allow for solid conclusions. A meta-analysis was not indicated owing the degree of heterogeneity.

Applicability of evidence

Before any conclusion can be reached on whether apathy can be modified by interventions, a careful look at each included study is necessary. The nine studies described in the present review compared different intervention techniques, with different study populations and controls, different inclusion and exclusion criteria and a variety of measurement scales as well. For example, not all studies had apathy as an inclusion criterion.

Czernick et al.[84], investigated the effect of dopaminergic therapy on motivation of Parkinson's patients and used the AS scale to measure motivation. Although some of the patients had AS scores >14 (AS cut-off score), the majority of the participants had scores <14, and therefore, the AS average at baseline was low for both intervention and control group (table 2.2), suggesting that the participants might not have lack of motivation.

Niu et al. [88], also did not have apathy as specific inclusion criterion; probably because the objective of the study was to investigate the effect of cognitive stimulation therapy on overall neuropsychiatric symptom and apathy was not the primary outcome. Similarly to Czernick et al., the baseline score of the NPI-apathy was also low. The fact that both

Niu et al. [88], and Czernick et al. [84], had mild apathy population might explain the largest and more significant effect size found in these studies. It might be more difficult to modify a more severe apathy.

Another important factor to take into account for the included population is the comparison group. In all studies participants with neurological disorder were stratified by one or the other group, except for Czernick et al. [84], that included healthy participants in the control group. The variability of the included population criteria makes it inappropriate to pool an overall effect.

The strongest evidence for change was observed for pharmacological studies with higher SMD compared to non-pharmacological studies; in particular, dopaminergic therapy seems to be a more effective therapy [84] . Dopamine neurotransmitter is known to modulate motivation, arousal, motor response, and sensorimotor integration [63]. People with PD have decreased levels of dopamine which might also explain the better effect of Dopaminergic therapy on apathetic PD patients rather than Psychostimulants [86], cognitive stimulants [93], or cognitive therapy [85] which have some evidence for stroke and AD participants.

The non-pharmacological study, Politis et al. [85], showed a small SMD. This might be explained by the fact that both interventions were effective in decreasing apathy and there was no additional effect of the cognitive therapy over the ‘one-on-one’ therapy sessions. In fact, the participants in the control group did slightly better, with modestly

improvement in quality of life and reduced need for a push in activities of daily living. Maybe providing individual time and attention in addition to stimulating activities can help motivate apathetic patients as well as provide a more structured cognitive therapy; but might not be suitable for implementation in clinical practice.

The moderate SMD presented by Herrmann et al. [86], is probably a consequence of the short treatment duration and small sample size. Even though, cross-over designs can be a good alternative to demonstrate precise results in a much smaller sample size, it would be interesting to investigate the effect of methylphenidate on apathetic patients in a larger trial. The results would probably be more similar to the larger trials included in the analysis.

The single subject-designs [92] and case series [90] showed promising results, with both studies reporting successful treatments. While the cholinergic therapy by Whyte et al. [89], demonstrated a trend towards apathy change over time, Skidmore et al [91] did not report any change in apathy score, maybe because apathy was not the target behaviour, the purpose of the intervention was to evaluate the tolerability of meta-cognitive strategy in cognitive impaired subject. It would be useful to investigate the success of these techniques in modifying apathy in neurological populations with more rigorous methodological design rather than a single component and with interventions specifically designed for apathy.

Quality of Studies and Potential Bias

Meta-analysis combines data across studies in order to estimate change with more precision than is possible with a single study. The main limitation to a meta-analysis in this review was the heterogeneity of the patient population (e.g. stroke, AD, and PD), types of interventions and outcome measures across studies. Also, the quality of the studies varied considerably. The majority of the studies were of moderate quality and one was of poor quality, as they did not clearly reported allocation sequence generation and concealment. Blinding was also an issue, indicating high risk of bias for the included studies. Blinding of the participant and therapist in non-pharmacological studies may not be possible, but blinding the outcome assessor can minimize bias. This could be verified for the two RCT non-pharmacological studies that presented the highest methodological quality (8/10) [85, 88]. On the other hand, all pharmacological studies were unclear on blinding, but for this type of studies it is feasible to blind all people involved in the study. Two studies [86, 93] stated that they were double-blind but still didn't specify who was blind. And the other three did not report about blinding.

For the studies assessed by PEDro scale, random allocation might be another limitation of the included studies, only two [85, 88] out of the seven studies assessed, reported the appropriate method of randomization. On the two [91, 92] studies assessed by SCED, replication of results and generalisation were not reported, which might be the potential limitations of the studies. Single-subject designs, if well developed, can be very valuable

to guide treatment, especially when high quality RCTs are not available. There were no other major methodological biases in the included studies; however, since the methodological quality of the studies was not optimal, the positive results towards apathy should be interpreted with caution.

Publication bias should be taken into consideration as well. In general, smaller trials are analysed with less methodological rigidity than larger studies, and have less chance to be published. Another important consideration is the fact that studies and reports from grey literature or that have not been published are not usually included in systematic reviews[95]. In the funnel plot analysed (Figure 2.3), we would expect the studies with large effect size to be on top, close to the SMD and the small effect studies on the bottom and spread around the mean. This is a symmetrical model. However, in the present study this is not observed; there is an asymmetry on the distribution of, suggesting presence of publication bias. In addition, there is a gap of studies on the right, suggesting that non-significant studies were likely not published [95]. We also only included studies in English, Portuguese, and Spanish languages, which might also lead to publication bias. One paper, found by the data base search, was excluded due to language restrictions (German). Therefore, publication bias should be considered prior to a conclusion.

Finally, selection criteria might be another limitation of the present study. Three studies were excluded for using other assessment criteria rather than the four measurement

scales described in inclusion criteria. More studies using standardized measures of apathy should be developed and a minimum clinically important difference should be defined for the available apathy scales in order to provide evidence-based guidelines for apathy.

There were very few studies evaluating behavioural interventions for apathy and none in a stroke population. Given the importance of motivation for engaging in a stroke rehabilitation program, this would be a valuable field for more research including developing measures that reflect the apathy situation for people with stroke, and developing interventions that are grounded in neurobiology.

CONCLUSION

In conclusion, evidence suggests that apathy can be modified by different types of interventions, but the heterogeneity of the studies and poor methodological quality do not allow for solid conclusions. There is a need for more high quality, methodologically rigorous studies of both pharmacological and non-pharmacological interventions to confirm these findings. More research is needed on dose, timing, duration, and intensity of therapeutic approaches for apathy, either as single or complex interventions. Maybe associating of non-pharmacological therapy with the effect of medication can lead to a more significant change in apathy and increase the efficacy of both treatments. Probably a multidisciplinary and multi-type approach will be the most effective way for

dealing with apathy for people with neurological disorders as motivation and engagement with rehabilitation are key to achieve function optimization.

Table 2.1: Search strategies for databases

Search set	Medline	Embase	PsycINFO	CINAHL
Date of last search	12/07/2011	12/07/2011	12/07/2011	18/07/2011
1	exp Cerebrovascular Disorder*/	exp cerebrovascular disease/	exp Cerebrovascular Disorders/	Cerebrovascular disorders (explode)
2	exp Parkinson Disease/ or Parkinson* disease.mp.	exp Parkinson disease/	exp Parkinson's Disease/ or Parkinson* disease.mp.	Parkinson disease, Alzheimer's Disease (explode)
3	exp Alzheimer Disease/ or Alzheimer* disease.mp.	Parkinson* disease.mp.	exp Alzheimer's Disease/ or Alzheimer* disease.mp.	Alzheimer* disease (keyword)
4	exp Brain Injuries/	exp Alzheimer disease/	exp Traumatic Brain Injury/	brain injury (explode)
5	tbi.mp.	Alzheimer* disease.mp.	TBI.mp.	TBI (keyword)
6	1 or 2 or 3 or 4 or 5	exp brain injury/	exp Physical Therapy/ or physical therapy*.mp.	1 or 2 or 3 or 4 or 5
7	exp Physical Therapy Modalities/ or physical therapy*.mp.	TBI.mp.	exp Rehabilitation/	rehabilitation (explode)
8	exp Rehabilitation/	1 or 2 or 3 or 4 or 5 or 6 or 7	exp Cognitive Rehabilitation/	Nursing Practice
9	exp "Recovery of Function"/	Physiotherapy*.mp.	exp Cerebrovascular Accidents/	Evidence-Based (explode)
10	exp Behaviour Therapy/	exp physiotherapy/	1 or 2 or 3 or 4 or 5 or 9	Professional Practice Research-Based (explode)
11	Cognitive Therapy/	Physical therapy*.mp.	exp Occupational Therapy/	recovery
12	goal setting*.mp.	exp rehabilitation/	exp "recovery (disorders)"/	recovery of function (keyword)
13	electrotherapy.mp. or	exp physiotherapy	exp Behaviour	behaviour

	exp Electric Stimulation Therapy/	practice/	Therapy/ or behaviour* therapy.mp.	therapy (explode)
14	exp Motivation/ or motivation* therapy.mp.	recovery of function.mp.	exp Cognitive Therapy/	goal setting* (keyword)
15	7 or 8 or 9 or 10 or 11 or 12 or 13 or 14	exp "physical activity, capacity and performance"/	exp Goal Setting/ or exp Motivation/ or goal setting*.mp.	Decision Making (explode)
16	apathy.mp. or Apathy/	exp behaviour therapy/	exp Electrical Brain Stimulation/	electrotherapy (keyword)
17	6 and 15 and 16	behaviour* therapy.mp.	electrotherapy.mp.	motivation* therapy (keyword)
18		exp cognitive therapy/	motivation* therapy.mp.	7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17
19		exp motivation/ or goal setting*.mp.	6 or 7 or 8 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18	Apathy (mesh and keyword)
20		electrotherapy.mp. or exp electrostimulation therapy/	apathy.mp. or exp Apathy/	6 and 18 and 19
21		motivation* therapy.mp.	10 and 19 and 20	
22		9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21		
23		apathy.mp. or apathy/		
24		8 and 22 and 23		

Table 2.2: Characteristics of Included Studies

Studies	Type of Study	Type of Intervention	Population	Apathy Scale	Age	Inclusion Criteria	Apathy at baseline		Control	Sample size
							I	C		
Czernecki et al., 2002[84]	RCT	Pharmacological (Dopaminergic Therapy)	Parkinson's disease	AS	57.9 (1.9)	Apathy NOT a criteria	12.8*[1] (2.0)/ 14.8*[2] (2.1)	8.2 (0.6)	Healthy	33
Politis et al., 2004[85]	RCT	Non-Pharmacological (Cognitive therapy)	Alzheimer's disease	NPI	84.0 (4.5)	Presence of apathy (staff judgment)	5.4 (4.8)	6.2 (4.3)	Patients	36
Herrmann et al., 2008[86]	RCT (Cross-over)	Pharmacological (Methylphenidate)	Alzheimer' disease	AES	77.9 (7.8)	Presence of apathy (NPI>1)	5.9 *[3] (3.0) 48.3*[4] (11.0)	same	Patients	13
Robinson et al., 2009[87]	RCT	Pharmacological (Nefiracetam)	Stroke	AS	66.3 (11.8)	Presence of Apathy (Score of 2 or 3 in specific AS items)	20.3 *[5] (4.8) 21.2 *[6] (5.7)	19.3 (4.8)	Healthy	70
Niu et al., 2010[88]	RCT	Non-Pharmacological (Cognitive Stimulation)	Parkinson's disease	NPI	79.8 (4.3)	Apathy NOT a criteria	3.5 (1.86)	3.25 (1.61)	Patients	32

*[1] Baseline value for on-first group; *[2] baseline value for off-first group; *[3] NPI baseline value used for inclusion criteria; *[4]AES baseline value used for analysis; *[5] baseline value for Nefiracetam 600mg group;*[6] baseline value for Nefiracetam 900mg group

Table 2.3: PEDro rating of methodological quality of RCT, CCT and case series (CS)

Criterion	Czernecki et al., 2002[84]	Politis et al., 2004[85]	Herrmann et al., 2008[86]	Robinson et al., 2009[87]	Niu et al., 2010[88]	Whyte et al., 2008[89]	Newburn & Newburn 2005[90]
Study Design	RCT	RCT	RCT	RCT	RCT	CCT	Case series
1. Eligibility Criteria were specified	Yes	Yes	Yes	Yes	Yes	Yes	No
2. Participants randomly allocated to intervention	Yes	Yes	No	Yes	Yes	No	No
3. Allocation was concealed	No	No	No	No	No	No	No
4. Intervention group was similar at baseline regarding key outcome measure(s) and most important prognostic indicator	yes	Yes	Yes	Yes	Yes	Yes	No
5. There was blinding of all participants	No	No	No	No	No	No	No
6. There was blinding of all therapists	No	No	No	No	No	No	No
7. There was blinding of all assessors who measured at least one key outcome	No	Yes	No	No	Yes	No	No
8. Measure of at least one key outcome were obtained from more than 85% of the participants initially allocated to groups	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Criterion	Czernecki et al., 2002[84]	Politis et al., 2004[85]	Herrmann et al., 2008[86]	Robinson et al., 2009[87]	Niu et al., 2010[88]	Whyte et al., 2008[89]	Newburn & Newburn 2005[90]
8. All participants for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by "intention to treat"	No	Yes	Yes	Yes	Yes	No	Yes
9. The results of between intervention group statistical comparisons are reported for at least one key outcome	Yes	Yes	Yes	Yes	Yes	Yes	No
10. The study provides both point measures and measures of variability for at least one key outcome	Yes	Yes	Yes	Yes	Yes	No	No
Total (item 2-11)	6/10	8/10	6/10	7/10	8/10	4/10	2/10

Table 2.4: SCED rating of methodological quality of single-subject design and case report (CR)

Criterion	Skidmore et al., 2011[91]	Lane- Brown and Tate, 2010[92]
Study Design	Case Report	Single-subject
1. Clinical history was specified	Yes	Yes
2. Target behaviours. Precise and repeatable measures that are operationally defined and specified	Yes	Yes
3. There are three phases of the study design, either ABA or multiple baseline	No	Yes
4. Sufficient baseline sampling was conducted	Yes	Yes
5. Sufficient treatment phase sampling was conducted	No	Yes
6. Raw data points were reported	Yes	Yes
7. Inter-rater reliability was established for at least one measure of target behaviour	No	No
8. There was independence of assessors	No	No
9. Statistical analysis was undertaken	No	Yes
10. There was replication of results either across subjects, therapist or setting	No	No
11. There was evidence for generalisation	No	No
Total (2-11)		

Table 2.5: Summary data for each included study

Studies	Type of Study	Sample Size			Apathy Scale	Apathy at baseline		Apathy Post		Change (post-base)		MD	pSD	SE (MD)	Cr	SMD (95% CI)
		T	I	C		I	C	I	C	I	C					
Czernecki et al., 2002	RCT	51	23	28	AS	13.9 (2.3)	8.2 (0.6)	11.0 (2.5)	7.6 (0.6)	-2.87 (2.3)	-0.6 (0.6)	- 2.87	2.02	-	-	-1.42 (-2.0;-0.8)
Politis et al., 2004[85]	RCT	36	18	18	NPI	5.4 (4.8)	6.2 (4.3)	1.2 (2.3)	2.3 (3.8)	-4.2 (4.8)	-3.9 (4.3)	-0.3	4.56	-	-	-0.06 (-0.7; 0.6)
Herrmann et al., 2008	RCT Cross-over)	13	13	12	AES	48.3 (11.0)	-	-	-	-2.31 (5.1)	0.5 (3.9)	- 2.81	-4.7	1.56	1.03	-0.62 (-1.4; 0.2)
Robinson et al., 2009	RCT	70	48	22	AS	20.7 (5.3)	19.3 (4.8)	-	-	-5.3 (7.7)	-2.0 (7.0)	-3.3	3.36	-	-	-0.98 (-1.6; -0.4)
Niu et al., 2010	RCT	32	14	15	NPI	3.5 (1.9)	3.25 (1.6)	-	-	-1.06 (0.8)	-0.31 (0.6)	- 0.75	0.73	-	-	-1.03 (-1.8; -0.3)

T = total; I= intervention; **C**= control; **MD**= mean difference; **pSD**= pooled standard deviation; **SE (MD)** = standard error of mean difference; **Cr**=correlation; **SMD**= standardized mean difference; **CI**= confidence interval

Figure 2.1: Systematic Review Flow Chart

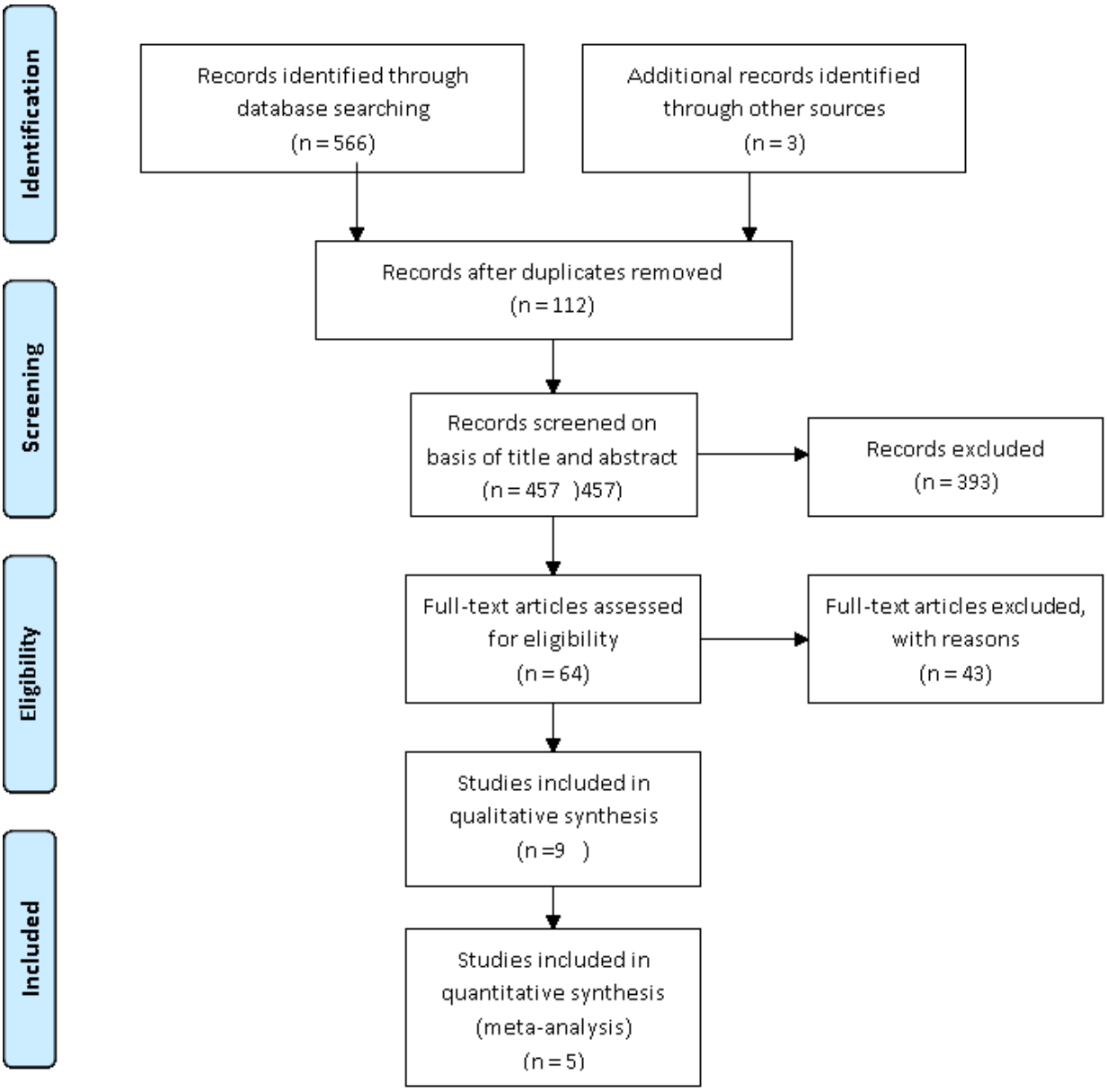


Figure 2.2: Effect size of intervention on Apathy

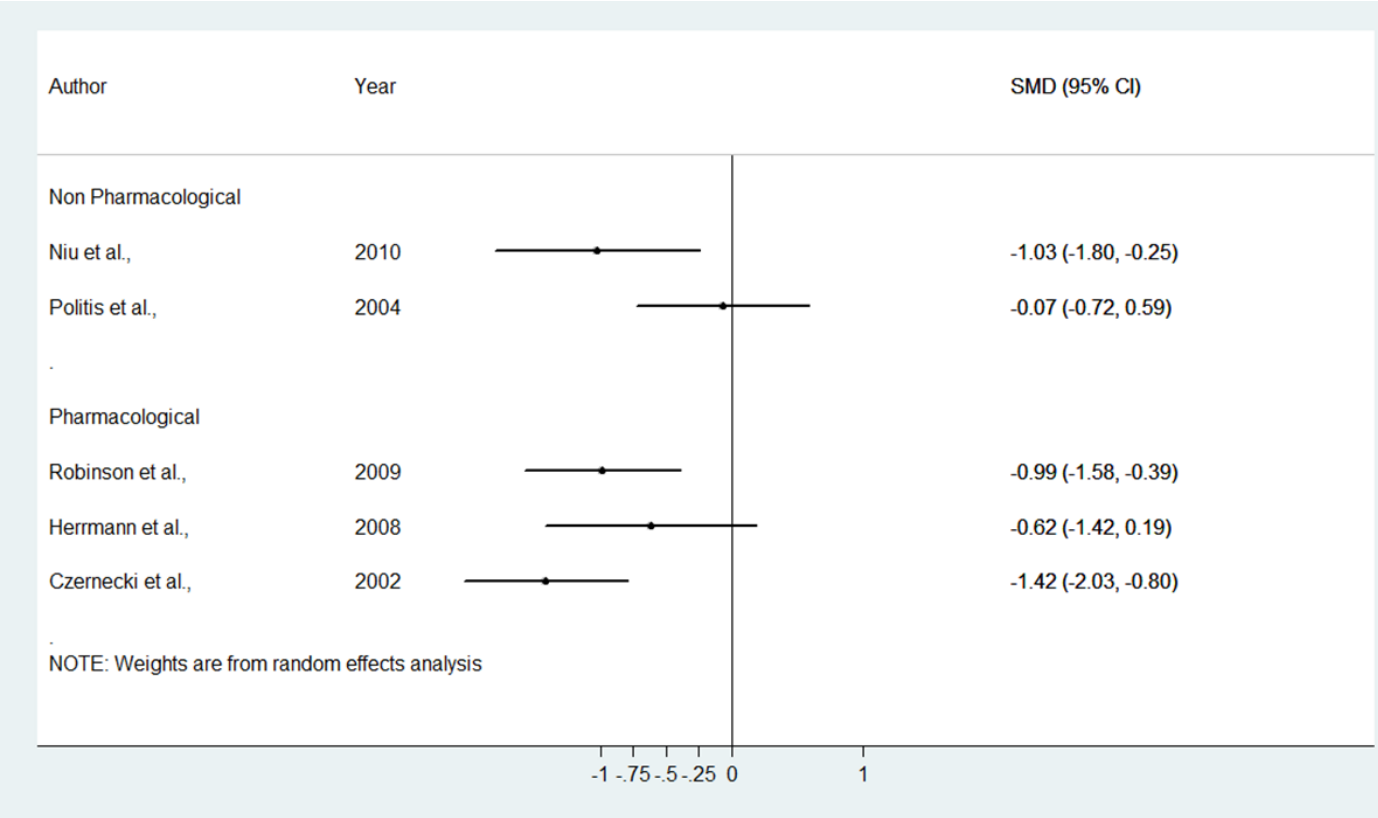
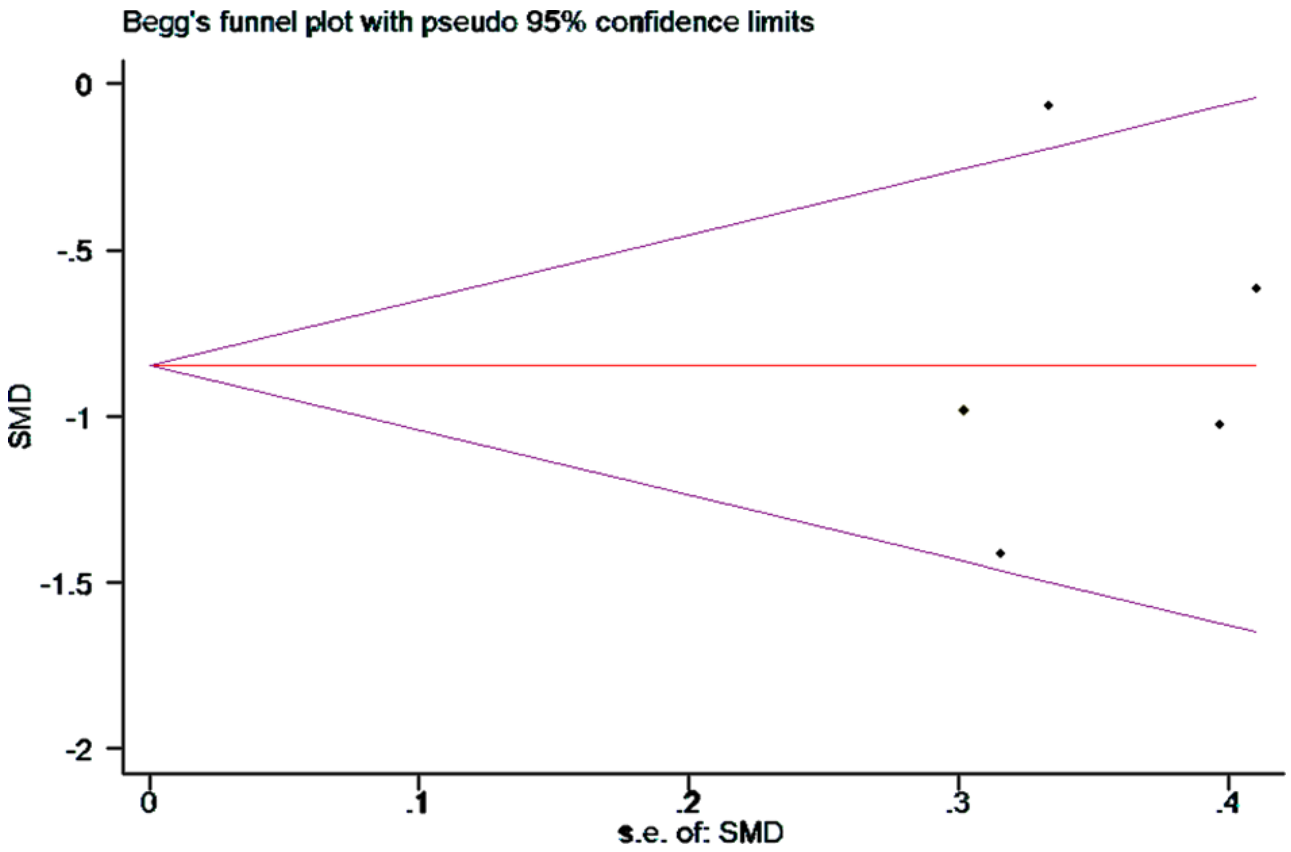


Figure 2.3: Funnel Plot for assessment of bias



Bias indicators

Begg-Mazumdar: Kendall's tau = 0.4 P = 0.33 (low power)

Egger: bias = 2.52 (95% CI = -18.00 to 23.05) P = 0.72

CHAPTER 3 - RESEARCH OBJECTIVES

Global:

The overall objective of this thesis is to contribute to the understanding of the role of apathy in stroke rehabilitation by taking a longitudinal view at the apathy-motivation continuum. The work will have three distinct components: (i) construct conceptualization; (ii) construct measurement; and (iii) construct impact.

Specific:

I. The extent to which the items in a widely used apathy index can be linked to the ICF and fit a unidimensional linear continuum;

II. To identify if items from closely related constructs fit onto the linear continuum;

III. To identify the extent to which the location along the apathy-motivation continuum as defined using Rasch analysis is associated with a lesion location;

IV. To estimate the extent to which change along the apathy-motivation continuum occurs during the first year post stroke;

V. To estimate the extent to which location along the apathy-motivation continuum impacts functional recovery after stroke.

CHAPTER4-MANUSCRIPT2: Improving the measurement properties of the AS using Rasch Analysis revealed a measure of motivation

PREFACE TO MANUSCRIPT 2

The concept of Apathy was analysed based on literature review [26] which provided evidence that apathy is a complex and abstract construct that can be difficult to identify and assess. Before the diagnostic criteria of Robert et al. [24], at least 3 other conceptualizations of the apathy syndrome had been identified, however, they lacked clear definition and agreement. Now that the Apathy Syndrome has well established characteristics, restrictions, preconditions, and outcomes [24], the accuracy with which apathy is identified in a clinical setting may improve.

However, the nomenclature used to characterize the Apathy Syndrome are best fit for the medical and psychiatric field and may be difficult to operationalize in the context of rehabilitation. As rehabilitation professionals most often have the difficult task of trying to motivate and engage patients in therapy during stroke recovery, this group of health professionals needs to be able to understand and measure motivation/apathy in order to adapt therapy and enhance motivational techniques.

Using a rehabilitation framework to situate apathy within the disability construct would be of great value to bring this field into the forefront of rehabilitation sciences. The

predominant framework for disability is the World Health Organizations' (WHO) International Classification of Functioning, Disability, and Health (ICF) [96] [97].

Summarizing the four Robert et al. criteria[24], the Apathy Syndrome is defined by: (i) manifestation of symptoms; (ii) comparison with previous level of motivational function; (iii) impact on activities and participation; and (iv) not be caused by diminished level of consciousness, emotional distress, or dementia. Observed symptoms of lack of motivation (e.g. interest, curiosity, action initiation, effort) and of emotional response, are the measurable components of the syndrome, and both need to be present for a person to be classified as having the Apathy Syndrome (Figure 4.1).

We have chosen to use the term "motivation continuum" to describe impairment of "pure" motivation; and the term "apathy-motivation continuum" to describe the construct when the emotional component is included in the measurement model.

Within the ICF framework, the symptoms of the apathy-motivation continuum fall within the Mental Function chapter of the Body Function domain. The specific categories reflecting the motivation continuum are: "Openness to experience", "Energy and drive functions", and "Organization and planning". The additional required component to form the apathy-motivation continuum reflects the category of "Emotional function", which involves the appropriate range and regulation of emotion. Figure 4.1 illustrates this conceptualization.

The practice of linking a complex construct to the ICF, in order to validate its relevance to rehabilitation and the scope of its content, is well developed. Dr Alaracos Cieza, a leader in this field, has over 200 publications linking diverse constructs to the ICF for the express purpose of understanding and measuring it in the context of rehabilitation [98-105]. She has also developed a set of linking rules for this purpose [101]. Moriello et al.[106] have developed a specific protocol for mapping items of a measure to the ICF to support content validity. Many others have also contributed to this field [107-115].

While a number of measures have been specifically developed to assess the apathy-motivation continuum such as the Apathy Evaluation Scale (AES) [31], the Apathy Scale (AS) [29], and the Apathy Inventory (AI). There is still no agreement as to which is the best measure of the apathy-motivation continuum. If Apathy is to be considered a relevant construct in rehabilitation, the content of these measures needs to fit within the ICF framework.

The AS is the most widely used measure of apathy in stroke. Its development, validity, and reliability in different populations has been demonstrated, however, the methods are not fully described and results are not fully presented [29]. To our knowledge, there has been no study to date that evaluated the psychometric properties of the AS, in particular, assessing how comprehensively it reflects the construct; its unidimensionality; whether the response categories are functioning properly; and whether there are differential item function by personal factors.

The most frequently employed methods to develop and examine the psychometric properties of a measure are 'traditional psychometric methods', or Classical Test theory (CTT) [43, 116],which are based on correlational or descriptive analysis to assess scaling assumptions, reliability and validity [42]. The methods include descriptive statistics, assessing missing data pattern, correlations, item redundancy, endorsement frequencies, factor analysis, Cronbach's alpha per scale, among others. Although widely used, this psychometric approach has some important limitations that should be taken into consideration. Cano et al.[43] has outlined four main limitations: (i) the metric originates from ordinal data that are summed to a total score which remains ordinal; (ii) the score for the person and the sample under study dependent on what measure was used; (iii) reliability and validity are sample dependent; and (iv) data is suitable for group use and not individual patient assessment.

A more modern psychometric method such as Rasch analysis can address some of these limitations. Recent studies comparing both methods have shown that internal consistency, discrimination ability, and unidimensionality were better assessed by Rasch analysis than by CTT methods [42, 116-118].

Rasch analysis is a method based on a mathematical model (the Rasch model), which tests the extent to which the observed data fit this mathematical model. It focuses on the probability that a person will respond to a certain item; that is if the observed score matches the prediction from the Rasch model (expected score). This is different from

the CTT that focuses on the person score in relation to the total score in the measure [43, 119].

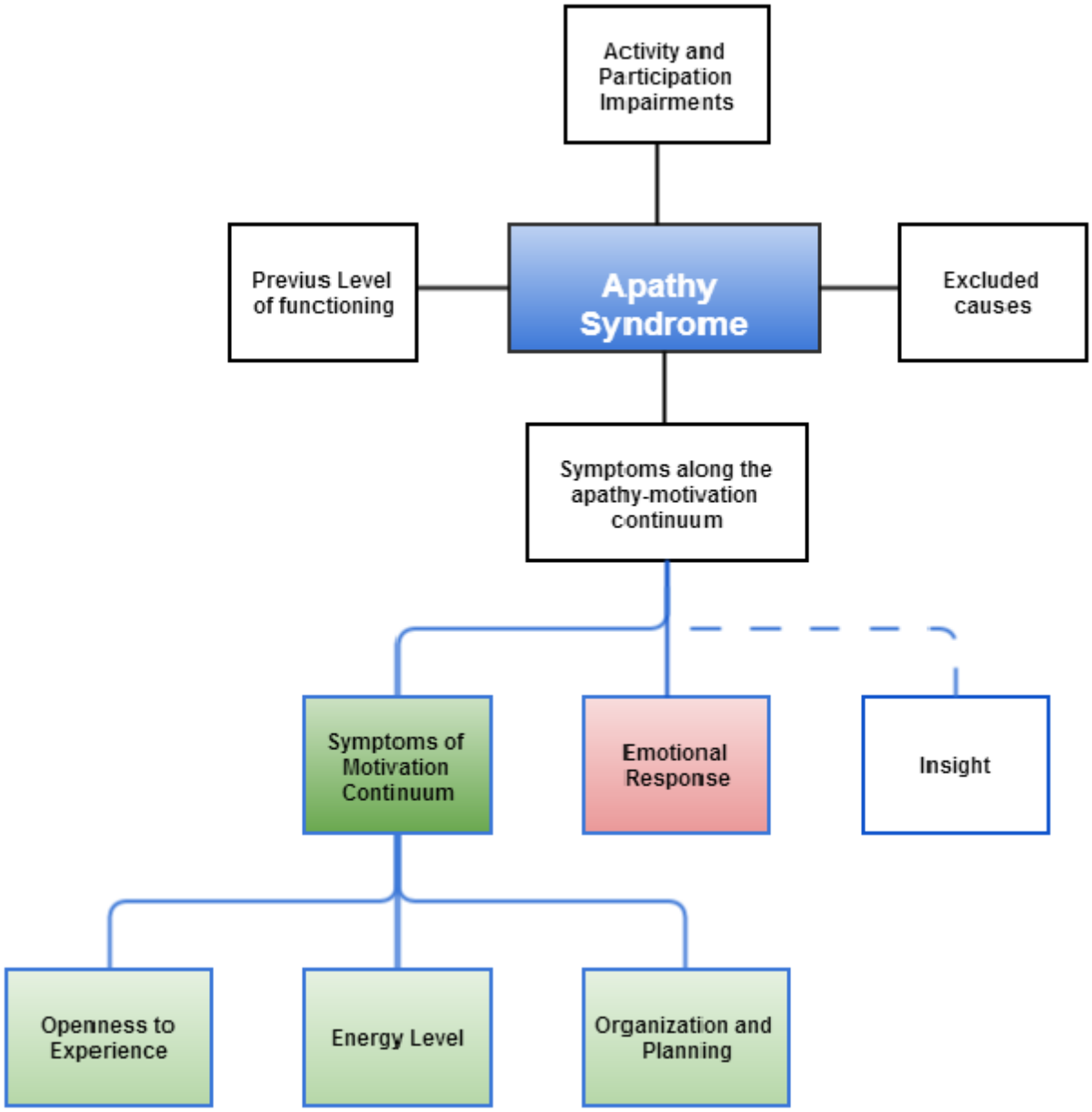
It is well known that summing ordinal categories violates the assumption of additivity as the units being added are not mathematical quantities, but rather ranks, and each item may not contribute equally to the quantity being measured. This is the main limitation of CTT, as the units lack mathematical properties. Also under CTT the assumption is that the "true score" is a function of the total score (derived from this ordinal units) plus error, the error is also assumed to be the same across items and across people [43], which is unlikely but cannot be tested.

On the other hand, when data fit the Rasch model, the scores are transformed into interval-like scores, providing a valid scoring system that can be summed into a total score providing valid comparisons across objects to be measure [117]. The Danish mathematician, George Rasch [120] demonstrated that responses to questions could be aligned onto a linear continuum using the logistic function with the probability of responding correctly or endorsing a particular level of ability as a metric for item difficulty. Each item has its own difficulty level and associated variability. Independent of the item difficulty, he also showed that the ability of persons can also be quantified with a degree of error that is specific to the person. With linearized units, change in person ability can be more accurately quantified.

In Manuscript 2, the psychometric properties of the apathy scale will be examined using Rasch analysis. The use of a modern technique such as Rasch analysis to assess the psychometric properties of a questionnaire has increased in rehabilitation research [121-127].

The approach taken in the current study is to estimate the extent to which the items on the Apathy Scale, administered in people recovering from acute stroke, fit the Rasch model using Rasch analysis. To create a valid metric to understand how the Apathy Scale characterizes behaviour in people with stroke, post-hoc adjustments of responses can be made to improve the fit of the data to the Rasch model. We believe that measuring the apathy-motivation continuum with linearized units will improve the measurement properties of the AS which will permit other analysis of this construct to be carried out in the context of stroke recovery. This exploratory approach will also provide evidence of what needs to be improved and guide future steps to develop a better measure [119].

Figure 4.1: Apathy syndrome Conceptual Map



TITLE PAGE

Improving the measurement properties of the Apathy Scale using Rasch Analysis revealed a measure of motivation

For submission to: Neurology

Authors: Christiane B. Lourenço, M.Sc.,[1]Nancy E. Mayo, Ph.D.[1,2,]

[1] School of Rehabilitation Sciences, Faculty of Medicine, McGill University, 3654 Prom Sir William Osler, Montreal, Quebec, H3G 1Y5, Canada.

[2] Division of Clinical Epidemiology, Royal Victoria Hospital, McGill University Health Center, Montréal, Québec.

Communication address to: Christiane Lourenco, Royal Victoria Hospital, Division of clinical epidemiology, 687 Pine Avenue West, Ross 4.27, Montreal, Quebec, H3A 1A1 Canada. Telephone: (514) 934-1934 ext 37066, Fax: (514) 843-1493, Email: christiane.lourenco@mail.mcgill.ca

Key Words: apathy, motivation, Rasch measurement theory, rehabilitation

ABSTRACT

Objectives: The objective of the study was to estimate the extent to which the 14 items of the Apathy Scale map to the ICF and fit the Rasch model

Methods: A secondary data analysis of a prospective longitudinal study on stroke depression was carried out. ICF mapping was used to assess the content validity of the AS. Rasch analysis was conducted in a step-wise manner, to estimate the extent to which the 14 AS items fit the underlying theoretical Rasch model.

Results: The results of this study indicate that the AS, in its original form, has a number of psychometric problems at the item level. The mapping exercise showed that all 14 items could be readily mapped to the ICF. Eleven items were mapped to the component of motivation (openness to experience, energy and drive function, and organization and planning); one to the component of emotional function; two items were mapped to other categories. Rasch Analysis results show that all items had psychometric limitations. Four Items did not fit conceptually and were deleted. The steps taken to improve the measurement properties of the AS resulted in a new measure comprising 10 of the original 14 items reflecting the essential motivation component of the apathy syndrome. As such, it is best described as a measure of motivation symptoms.

Conclusions: The results of this manuscript contributed to the understanding of the conceptual and psychometric properties of the most used measure of the apathy syndrome in Stroke, the Apathy Scale (AS) [29]. The limitations identified here reinforced

the importance of having a strong conceptual model prior to developing a good measure.

INTRODUCTION

Apathy syndrome has been shown to be associated with neurological disorders, including stroke. The frequency of the apathy syndrome in people post stroke has varied across studies, with an average of 33% (15%-50%)[128]. Apathetic stroke patients tend to have poor functional outcomes, more severe depression, and score lower on cognitive testing [70, 129-131].

Different clinical definitions of the apathy syndrome have been proposed in the literature [24]. There is agreement that the apathy construct includes impairments of motivation (e.g. interest, action initiation), lack of emotional response, and possibly an introspective dimension related to self-awareness (insight) [21, 24].The fact that these symptoms are represented in the World Health Organizations' (WHO) International Classification of Function Disability and Health (ICF) solidifies situating the apathy syndrome within functionand disability. The ICF categories that are related to the measurable component of the Apathy Syndrome are: openness to experience (b1264), energy level (b1300), motivation (b1301), organization and planning (164), and emotional function (b152).

To date none of the existing ways of measuring or assessing apathy would meet the psychometric criteria for a gold standard [32]. The most widely used measure in stroke is the Apathy Scale (AS) proposed by Starkstein [29, 30]. Other measures that have been used include the Apathy Evaluation Scale (AES) [31], Apathy Inventory (AI) [32], and the Neuropsychiatric Inventory (NPI) [33]. A systematic review of 5 RCT and 2 observational

studies, evaluating the effect of pharmacological and non-pharmacological interventions on apathy in people with neurological disorders, concluded that the heterogeneity in apathy assessment limited the strength of conclusions about the effect of interventions on apathy. The use of different measures also yielded different estimates of prevalence post stroke, ranging from15%to 50% (average of 33%) (see Manuscript 1, Chapter 2) [128].

The importance of properly identifying and quantifying apathy in stroke is evident by its high occurrence and impact on recovery. Apathetic stroke patients pose a rehabilitation challenge as they do not adequately engage in the process [132]. On the other hand, the rehabilitation process may not be tailored to meet the needs of apathetic stroke patients.

Without a strong measurement model for the apathy syndrome, effective intervention is not possible. To quote Lord Kelvin: “If you cannot measure, you cannot improve it”. Rasch Measurement Theory (RMT) [119, 133] is an experimental paradigm based on strong measurement theory providing an evidence base for the extent to which a set of items form a real measure. Rasch analysis is a mathematical methodology to identify whether a total score can be justified from adding the item scores. However, just because the “data” (item scores) fit the Rasch model, a hierarchical model with linearized units of measurement, only suggests that “something” is being measured properly but cannot specify what is actually being measured. To establish a strong theoretical base, mapping items to a reference standard and examining the items within Rasch analysis

would satisfy criteria. Thus, the objective of the study was to estimate the extent to which the 14 items of the Apathy Scale map the ICF and fit the Rasch model.

METHODS

Participants

A secondary data analysis was conducted from a prospective longitudinal study on stroke depression. Briefly, all stroke patients admitted at any of the three sites of the McGill University Health Centre, and discharged back into the community or to an inpatient rehabilitation centre, were included in the study. Exclusion criteria were: (i) presence of a co-morbid condition (organ failure or malignancy) that could lead to death in the first year following stroke; (ii) residence located more than 50 km from the hospital; and (iii) unable to provide informed consent. Of the 724 patients admitted, 120 were eligible and consented to participate. The Apathy Scale [29] was not part of the original measurement battery but was added after the study commenced and hence only 82 of the 121 (68%) completed this measure. In this study this group is referred to as the Apathy assessed cohort.

Procedures

Participants were evaluated at 10 days post-stroke and subsequently at 3, 6, 9, and 12 months by trained health care professionals who assessed level of activity and

participation, as well as administered questionnaires on mood and emotional function. The AS was not administered at the 10-day evaluation so four time points (3, 6, 9, 12 months) were available. The study had ethical approval from the McGill University Institutional Review Board.

Measures

The Apathy Scale (AS) is a 14-item rating scale (14 items) with four ordered response categories (0, 1, 2, or 3) which are added over items to form a total score ranging from 0 to 42. Items 1 to 8 are positively worded and scored either 3 (‘not at all’) 2 (‘slightly’), 1 (‘some’), or 0 (‘a lot’). Items 9 to 14 are negatively worded and scored either 0 (‘not at all’) 1 (‘slightly’), 2(some), or 3 (‘a lot’) (Appendix A4).

There is a conceptual challenge in measuring negative constructs as the absence of the negative does not mean the presence of the positive. It is more meaningful to think of a person possessing a quantity of a positive construct. For the apathy continuum, it is more sensible to think of the construct of motivation at the higher end, and of apathy at the lower end of the spectrum. Therefore, for purposes of considering this construct as a quantity, the original scoring was reversed, so a high score here will indicate motivated, a low score apathetic.

ICF mapping

In order to assess the content validity of the AS, we conducted a structured ICF mapping protocol developed by Moriello [106] applying the rules developed by Cieza [101], and incorporated a Delphi technique, to arrive at a consensus on the best code for each item. Eight health care professionals, from different areas of expertise including occupational (n=1) and physical therapists (n=5), and kinesiologists (n=2), participated independently by assigning alphanumeric codes to the items of the AS.

The percentage of agreement was calculated for all suggested codes. Agreement greater than or equal to 70% was selected as an *a priori* threshold indicating endorsement. If a code was endorsed at a 4-digit level, then the 3-digit root of that code was also endorsed. For items with less than 70%, the Delphi technique was used to arrive at a consensus, This process was repeated until the 70% agreement threshold was reached for an item, or if it was determined that agreement would not be reached.

Rasch Analysis

Data from 82 participants' ratings on the 14 items of the AS over four time points were available for analysis (n=232). Rasch analysis was conducted in a step-wise manner [134], to estimate the extent to which each item fits the underlying theoretical Rasch model. All analyses were performed using the Rasch Unidimensional Measurement Model programme, RUMM 2030 [135]. Figure 4.2 displays the Graphical Rasch model [136, 137] for the Apathy Scale in which the latent variable, apathy, is connected to the categories that represent the construct; and disconnected from the personal factors. This graphical

model represents a unidimensional measure (i.e., items reflecting different aspects of one construct), with locally independent items (i.e., items are disconnected from each other) without DIF (i.e., disconnected from personal factors) [138] .

Overall model fit

Misfit between the data and the model was verified by the analysis of Chi-square probability (<0.05 with a Bonferroni correction), and inspection of items and person fit residuals values. Fit residual for both items and people were expected to have mean location of 0 and SD of 1. A high mean residual and SD (>1.5) indicates a large deviation from the expected value and suggests that items and/or persons are the cause of misfit.

The Person separation index (PSI) is an indicator of how strong is the measure to discriminate between respondents with different levels of the trait being measured (e.g. apathy). A value of 0.7 is considered the minimal acceptable value to discriminate between two groups; a value of 0.8 can differentiate 3 ability groups; and a value ≥ 0.9 can differentiate 4 or more groups [139].

Ordering of Response Categories

For ‘n’ response categories there are n-1 thresholds. Statistical and graphical inspection identified whether the ordering of thresholds was as expected. For the present study, the expectation was that individuals with high levels of apathy would endorse items that

even people with low motivation would find easy to endorse. The expectation for people without apathy was that they would endorse items that only people with high motivation would endorse. If this threshold ordering was not met, a total score would not be meaningful. In the presence of disordered thresholds, response categories were collapsed.

Individual person fit and item fit

Item and person fit residuals were expected to range between ± 2.5 and have non-significant chi-square probability (>0.05). Graphical inspection of item characteristic curves (ICC) was used to verify if the individuals with similar level of ability (class interval) were located on the curve linking probability of response to the expected value on the latent trait. Misfit items were removed iteratively, starting with those with the highest fit residual statistics, until no more improvement was achieved.

Response Dependency

Item dependency was assessed by the magnitude of the residual correlation between items. The average of all item residual correlations was -0.07. Therefore, correlations above 0.1 indicated local dependencies. In the presence of local dependency, items were combined as subtests to form a “super-item”.

Unidimensionality

Principal component analysis (PCA) of the residuals was used to detect any meaningful patterns in the residuals that would indicate multidimensionality. This procedure creates two subtests of items (the positively and negatively correlated items) which are used to derive two estimates of person-location for each person. And independent t-test is used to compare these estimates for each person; the percentage of such test which fall outside the range -1.96 to 1.96 should not exceed 5%. A binomial test of proportions was calculated for the observed number of significant tests, and this value should overlap the 5% expected value for the scale to be unidimensional [138].

Invariance across Sample

Once global fit to the model was achieved, we verified whether items were functioning differently in relation to six personal factors: time of assessment (3, 6, 9 and 12 months), age (<60, 60-70, 70-80, >80), gender (male or female), education (less than high school, graduated high school, graduated college, university plus), language first spoken (English, French, or others), and language of the test (English or French). Analysis of variance, (ANOVA) was conducted to compare the standardised residuals across groups by personal factors, across class intervals, and the interaction of both. The Bonferroni adjustment of significance was applied for a Type-I-error level of 0.01. Given that there are fourteen items, four class intervals, two main effects (a class interval effect and an personal factors effect), and a personal factor by class interval interaction, the criterion level of misfit for each statistic was taken as 0.01/56= 0.00018.

If uniform DIF was detected, items were split by the specific personal factor; however, non-uniform DIF required the removal of the item from the scale.

Targeting

Rasch analysis was used to verify if the items were targeted for the stroke sample being measured. A well-targeted measure would have a person mean location matching the item standardized location of 0, as the hierarchy was standardized to a normal distribution with a mean of zero and a standard deviation of 1. Assessment of targeting was achieved by an examination of the spread of person and item locations on the Person-Item Distribution graph and summary statistics. The full range of measurement should cover ± 4 standard deviate (logits), assuming a standard normal distribution of a mean of 0 and SD of 1.

Statistical analysis

In order to compare the performance of the original AS and the new version (R-AS) that emerged from Rasch analysis, the following analyses were carried out: i) Pearson coefficient of correlation was calculated comparing the original AS, the R-AS, MHI, MOCA, 2 minute walk test (2MWT), and SIS participation; ii) Bland-Altman plot between AS and R-AS to identify measurement bias; iii) McNemars' chi-square was used to compare the percentage of people classified as apathetic on the two forms of the measure; iv) Repeated measures analysis of variance was used to compare the values

on the two measures across levels of depression, where the cut point was greater or less than 60 on the MHI of the Rand-36.; and v) Kendall's W test of concordance of ranks [140] was carried out to test stability of the item hierarchy across time points..

RESULTS

Participant's characteristics

Table 4.1 compares the characteristics of the Apathy assessed cohort to those excluded from the cohort at three months post stroke. Although there were some differences between those who administered the apathy scale and not (age, severity of stroke, side of lesion) these occurred by chance; the missing data was a consequence of design.

ICF Mapping

Figure 4.1 shows the conceptual model of apathy derived from the literature review (see chapter 1). The measurable components of the apathy syndrome were identified in five categories of the ICF: "openness to experience", "energy level", "organization and planning", and "emotional function". The expectation for the AS was that each domains of the apathy construct would be reflected in one or more items.

Table 4.2 presents the results of the mapping of the AS items to the ICF. The items are ordered by ICF domains. Four items (1, 2, 10, and 11) mapped to the 4-digit code for

“openness to experience” (B1264) with more than 71% to 86% of ratters agreeing on the mapping. Six items mapped to the 3-digit code for “energy and drive” function category (B130), and four items mapped to two more granular level: three for “motivation” (B1301) and one for “energy level” (B1300). Only one item, item 6, mapped to the “organization and planning”; and one item to the “emotional function” category (item 13). The remaining two items (3 and 9) each mapped to a different ICF category outside of the apathy construct. Also shown in Table 4.2 is the degree of which each rater agreed in the mapping.

First overview of key psychometric properties of AS using Rasch Analysis

The first analysis of the 14-item AS identified some important measurement limitations, particularly that the AS in its original format does not fit the Rasch measurement model. Summary statistics for the original AS and subsequent model resulted from alterations are displayed in Table 4.3. For the model of the original AS, the total chi-square was very high (211.8) and associated with a very low p-value (0.0000) indicating deviance from the expected model. The item-person interaction shows that person measures of ability are not centred (at 0 logits) but are at the higher end of the spectrum (0.70) indicating that individual’s levels of ability in this sample are higher than the items’ level of difficulties. The item-person threshold map for the original AS is shown in Appendix A5.The high standard deviation for item residuals (± 2.64) suggests that the cause of model misfit may be on the item level. The Person Separation Index (PSI) meets the

minimal internal consistency reliability requirement (PSI 0.714, and Cronbach's alpha 0.77).

Table 4.4 displays the results of the Rasch item analysis for the original AS. Eight out of the 14 items showed disordered thresholds; for five items some thresholds were non-discriminatory (Appendix A6). Items 3 and 6 did not fit the underlying linear model (Rasch model).The high positive residual of item 3 suggests under discrimination, i.e., that this item may not be measuring the same thing as the other items. Conversely, the high negative residual of item 6 suggests over discrimination and redundancy; Items 1, 4, 6, 7 and 9 also showed evidence for misfit with p-values<0.05. In addition majority of the subjects found items 13 and 14 to be confusing and needed further explanations.

Inspection of residual correlations matrices identified local dependency for several item-pairs (1 and 2; 3 and 4; 12 and 13; and 10 with 11, 12, 13, and 14). PCA of residuals identified two main subset loadings: the positively correlated items (9,10,11,12,13,14) and the negatively items correlated (1, 2, 3, 4, 5, 6, 7, 8). A paired t-test confirmed that the two subtests were estimating different levels of ability for more than 5% of the sample (23.8 %; CI 7.4, 30.5). This raises concern about the validity of the scale for use at the individual person level, such as adopting a particular cut-off score to classify individuals as apathetic or not.

Finally, nine items showed DIF in relation to five of the six personal factors (Table 4.4, Appendix A7); there was no DIF for time. The longitudinal comparison between item

ranks showed that the item hierarchy was very similar across time. In Table 4.5, it can be seen that item 9, the most difficult item in the all-time-points model, was also the most difficult item in the 3, 6, and 9 month models, and the third most difficult in the 12 month model. Despite small variation in ranking, there was convincing evidence of concordance (chi-square 105.6; DF 9; $p<0.01$). The null hypothesis of no systematic ranking was rejected.

Rasch Analysis to Refine the AS

Table 4.6 summarize all the steps taken to improve the measurement model. The first step was to delete items that did not fit the apathy conceptual model. According to the mapping exercise, item 9 is measuring the activity domain which is a consequence of apathy, rather than a mental function contributing to apathy. Item 3 is also measuring a different domain (insight) confirmed by the high positive fit residual identified in the item fit statistics and by the analysis of the ICC showed in Appendix A10.Therefore both items were excluded from the model. Items 13 and 14 were also deleted because these items were difficult for the respondent to understand.

The second step towards improving the measure was collapsing the thresholds. Even though there were only 8 items with disordered thresholds, there was an attempt to rescore all the items to the same response pattern to facilitate clinical use. All the items thresholds were modified according to the Category probability curve (CPC) and

Threshold probability curve (TPC); nine items to 3 response options: 0 ‘not at all’, 1 ‘some’, and 2 ‘a lot’, and one item (item 7) to 2 response option 0 ‘no’ and 1 ‘yes’.

After collapsing, there were no more misfitting items but multidimensionality was present and local dependency was observed between items 1 and 2, 10 and 11, and 6 and 7. As the pattern of response dependency was consistent with the results of the mapping exercise (e.g. items 1 and 2 measured the same domain, openness to experience), three ‘super-items’ (subtests) were created to represent the major categories of the apathy construct and to retain as many of the original items as possible. Therefore, subtest 1 was denominated ‘openness to experience’ with items 1, 2, 10 and 11; subtest 2 was ‘energy and drive function’ formed by items 4, 5, 8 and 12; and subtest 3 ‘motivation and plan’ formed by items 6 and 7.

The subtesting procedure improved considerably the fit of the statistical model (chi-square 5.7, p=0.76) and solved local dependency and multidimensionality as now the t-test identified that only 3.76 % (CI 0.8%, 6.7%) of the sample to have different estimates.

After subtesting, two out of the three subtests (openness to experience and motivation and plan) had DIF in relation to language first spoken. Thus, the subtest with the lowest probability value, openness to experience (p=0.003), was split by first language spoken. As language had 3 factors (French, English and any other language),the post-hoc analysis revealed a significant difference between French and other (absolute difference

0.39, D min 0.30) but no difference between English and other, suggesting that splitting subtest 1 by French and not French would be the most appropriated approach. Analysis of the ICC (Appendix A11) supported this choice.

The measurement model now satisfied all statistical requirements: unidimensionality, local independency, and lack of DIF. The summary statistics for the final model are shown in Table 4.3 (chi-square 6.4, p=0.89).

This new measure will be labelled Rasch version of the AS (R-AS).Figure 4.3 shows the targeting of the final items in relation to this sample population and it is clear that the sample is located at a higher level of motivation than the items average level of difficulty. The internal consistency of the measure satisfies the minimum value required for group use (PSI=0.730).

The hierarchy of the individual items is displayed in Table 4.7. The item threshold difficulty ranges from -2.76 logits to 2.48 logits. As the items only fit the Rasch model through three subtests, only three separate scores are calculated, but they can be summed into a total score. To make a direct comparison between the AS and the R-AS the score of the new measure was rescored to have a maximum value of 42, in which higher levels indicate more motivated and lower scores more apathetic.

The scoring algorithm for the R-AS is shown in Table 4.8. The logit score produced by Rasch Analysis was converted back into the score range of the original scale using the following formula: y= m + (s * logit score), in which “m” is the wanted minimum minus

the current minimum times “s”; and “s” is the wanted range divided by the current logit range [141]. This transformation has an aesthetic purpose to facilitate interpretation. The new score is derived directly from the logit score and has the same value and meaning as the original score [134, 142].

The R-AS has 3 super-items with each super-item having n*(k-1) thresholds, where n is the number of items and k is the number of response option. The first super-item openness to experience has four items and three scoring options and so the number of thresholds is 8 (12-4=8). As each threshold lines up hierarchically the highest value achieved is the score for that super-item. As all super-items fit the Rasch model, a total score can be derived by summing the value on each super-item.

Because two of the super items, openness to experience and energy and drive, had an almost equivalent threshold locations at the highest level of the continuum (+2.74 and +2.86 logits, respectively) and had the same number of thresholds, their maximum score would be equal. The third super-item achieved a maximum of 1.30 logits which is less than half of the other two super-item maximums. Because we wanted a total score of 42, to be comparable to the original measure scoring, a distribution across the three super-items was created yielding 18+18+6=42. The only other scoring options that would fit this data structure was 16+16+8=42, but as the standard error for the third super-item was larger (mean location 0.296; SE 0.09) than the other two super-items (mean location -0.30; SE 0.062, openness to experience and mean location 0.009; SE

0.06, energy and drive) we chose to give it less weight. However, as any scoring option is derived directly from the original logit scale using a linear transformations (with rounding), the absolute value chosen does not affect the estimate of the person's ability with respect to any other person in the sample. The non-uniform scoring across response options and items indicates the different location of each threshold on the linear continuum.

Comparisons between original and Rasch versions of the AS

Figure 4.4 present the Bland-Altman plot comparing the two measures, AS and R-AS. The y-axis is the magnitude of the difference between the AS and the R-AS (AS minus R-AS). The x-axis is the average of the two measures. The dotted line shows the average mean difference between the measure (2.5 ±5.2) and 95% CI (-7.5, 18).The solid line at 0 indicates no difference in score between the two measures. The dots represent individual subjects. As there are many dots above the zero line it is evident that AS gives people higher scores (more motivation) than R-AS. And in fact four subjects fall below the 95% confidence band for measurement error.

Table 4.9 compares AS and R-AS on indicators of constructs validity. The average values differed by 3 points, and AS classified fewer people as falling in the apathetic range(42.7%) in comparison to the R-AS (72%; $p<0.001$). Also shown are the average scores for people classified as depressed or not depressed using the MHI of the RAND-36 and 60 as the cut-point. The difference in the score between those not depressed

and those depressed in the original AS was 2.6 yielding a discrimination effect size of 0.39. In contrast the discrimination effect size for R-AS was 0.80 showing a significant difference in discrimination between subjects level of depression ($p<0.05$).

The correlations between the two versions of the measure with depression, mobility and participation were somewhat similar. However, the correlation between the original AS and cognition (0.21, CI -0.03 to 0.41) was greater than the R-AS and cognition (0.03, CI 0.18 to 0.24). The correlation between the original AS and the R-AS is illustrated in Figure 4.5, showing a strong positive linear correlation ($r=0.68$).

DISCUSSION

The results of this study indicate that the AS in its original form, although widely used in the Apathy literature, has a number of psychometric problems at the item level and consequently a total score cannot validly represent the construct. In fact, not one item (see table 4.4) was free of psychometric limitations. However, by applying Rasch analysis to identify anomalies with the items and guide appropriate modifications to items and response options, a valid measure was retained from the original 14 items. The four items that were deleted were those that did not fit conceptually or were too difficult to understand by patients (see table 4.4). The remaining items all had to have response options collapsed and even then only fit the Rasch Model if they were combined into 3 “super-items”.

During the application of Rasch Analysis, a considerable amount of DIF across items was found. Common sources of DIF in questionnaire items are gender, language, culture, and time [143]. These factors contribute to differences in interpretation of the items with respect to their location on the latent trait. Ignoring DIF can result in measures whose values do not signify the same thing across different groups of people. This is particularly important in multicultural countries like Canada in which there are two official languages, English and French, and a large immigrant population who speak a myriad of other languages. DIF was found for language spoken and language of test administration, which would require creating two versions of the measure for clinical use, French and English, with scoring systems particular to each version of the measure. This finding also emphasize the importance of proper translation and cultural adaptation. One of the challenges with translation is being able to translate the meaning of the construct rather than solely translating the words.

Rasch analysis is a latent trait methodology and as such defines the latent trait by the items. The actions towards improving the AS resulted in a final measure with three "super-items" covering the motivation component of the apathy syndrome. These modifications improved the fit of the items to the model; however, the latent trait was changed. The items remaining tap "openness to experience", "energy and drive", and "organization and planning" functions which represent the mental function that produces the interest, energy, and determination to act towards a specific need and goal in a persistent manner [71]. This is the definition of motivation and hence the new measure is

reflecting the motivation continuum rather than the apathy continuum. This was further confirmed by the strong correlation (0.68) between the two versions which indicates that the two measures are highly associated but are not measuring the exact same construct which would be indicated with correlations higher than 0.8.

Comparisons between the original AS and the R-AS identified some areas where the new measure has advantages. The proportion of potentially apathetic people was higher (see Table 4.8) and this would indicate that motivation needs to be a main stream target of intervention in people with stroke. An additional advantage of R-AS was that it showed better discrimination between people who were depressed and not depressed supporting construct validity. Also shown in Table 4.9 is the higher correlation of the original AS with cognitive status in comparison to the R-AS and cognitive status. We considered this association as an undesirable feature because we would not want a value on a measure to depend on person's ability to understand the items. A point of note is that when administering the original AS, it was clear that people did not understand item 14 asking if they considered themselves to be apathetic. This item was deleted.

Originally, eight of the fourteen items were negatively worded and these items had a deleterious impact on fit and dimensionality. The R-AS comprises only 3 negatively worded items and these fit the model only if they were embedded into subtests (super-item). Switching between positively and negatively worded items may have been particularly difficult for the stroke population and rendered the responses inconsistent. In

fact, negatively worded items have been shown to reflect different constructs than positively worded items, as patients interpret them as different concept [144].

Motivation has been identified by rehabilitation professionals as one of the most important factors influencing stroke rehabilitation [18] . But to date, there is no standard measure being used by clinicians, and people are labelled as motivated or not based on the clinician's judgement of the patient's behaviour and participation towards therapy. Having a measure of motivation would provide a more accurate method for identifying and targeting people with low motivation. The R-AS could be a possible solution to measure the motivation continuum. Further study is necessary to validate the measure with different clinical sample and to obtain patient's and clinician's perspective. It may also prove with further work to be screening tool. For this purpose, potentially only one item from each subtest would serve, the item that best spread across the respective domain. For example, in Table 4.8, item 4 ("interested in learning new things") spans the entire spectrum of the domain "openness to experience" with possible scores of 0, 6, or 18. Similarly item 4, for "energy and drive", and item 2 for "motivation and planning", were identified as the best items for the respective domains.

This study has a number of design limitations and also limitations that point out the need for future research. This was a study of an inception cohort with the primary aim of understanding the emergence, maintenance, and resolution of depressive symptoms over time. The registration of the inception cohort strengthens the generalizability of the results as all people with stroke were registered for the study at stroke onset. However,

as this was a research project, informed consent had to be obtained and, as is typical, many ill and cognitively impaired persons cannot be recruited; others shy away from committing to research in an uncertain and vulnerable state. Others are not eligible because of language barriers or distance from the study site. Until measures of motivation become part of routine care, all research on the topic will be restricted to populations who are capable of consent and wish to do so. While some of the reasons for exclusion do not result in bias, language and distance, it is difficult to know about the motivation/apathy state of those too ill or unwilling to consent. A comparison of refusers (n=67) to consenters (n=121) revealed that refusers, while of the same gender and age distribution, had slightly milder strokes and were more functional at discharge [145]. However, degree of apathy should not affect response to questions on apathy and the sample recruited and assessed in this study demonstrated a range on this construct.

The distribution in the item-person threshold distribution map (Figure 4.3) is normal but person level of ability is higher than item level of difficulty. There are more items covering the lower end of the continuum but there are more people at the higher end. This indicates, that if we want to measure change of motivation, more items at the higher end of the continuum should be added. If we want to detect low motivation, we can do that. The original version of the AS was not well targeted and person ability was higher than item difficulty (Appendix A5). The changes proposed improved the spread of person ability and item difficulty despite many fewer thresholds used for measurement. This indicates that there was considerable misclassification of the original scoring

structure. Therefore, it would be important to test this in other samples before firm conclusions can be made.

Another limitation is that the apathy questionnaire was not included in the study from the outset owing to concern about respondent burden and the awkwardness of the available questionnaires. However, as the study unfolded, the interviewers identified that motivation/apathy emerged and the protocol was amended to include the AS. This type of missing data is termed missing completely at random (MCAR) and does not introduce bias but reduces power [146]. Despite the small sample size, n=82, we were able to take advantage of multiple time points to provide enough power for the Rasch analysis as the item calibration was done on 232 observations.

Rasch analysis has a great potential to identify the strength and weaknesses of measures that can serve as a base to guide further improvement and development of the measure. Future research is needed to solidify the conceptual framework of motivation and include input from patients of different ages, with different health conditions, neurological and non-neurological, and caregivers, clinicians, neuroscientists, and representatives from diverse fields such as education and business.

At this stage, the new version of the AS could be used to further explore the motivation construct before a new measure with better psychometric properties can be developed. Future research also needs to delve further into the conceptualization of the apathy syndrome construct and estimate the extent to which the emotional component is an

essential part of the syndrome. The refinement of the original AS has resulted in the R-AS which is really a “Motivation Ladder”. As there are no items capturing the category of emotional function, the construct represents the motivation continuum. To be a “true” measure of the apathy syndrome, items reflecting both the motivation and emotional response component need to be included.

CONCLUSION

The results of this manuscript contributed to the understanding of the conceptual and psychometric properties of the most used measure of the apathy syndrome in stroke, the Apathy Scale (AS) [29]. The limitations identified here reinforced the importance of having a strong conceptual model prior to developing a good measure.

Table 4.1: Comparison of the Apathy assessment cohort and excluded subjects on key characteristics at 3 months post stroke

Characteristics	Apathy cohort (n=82) N (%)	Excluded (n=36) N (%)
Age (years) at stroke (Mean ± SD)	68 ± 11	77.4 ± 14.6*
<60	21 (25.6%)	3 (10.3%)
60-80	44 (53.7 %)	11 (35.9%)
>80	17 (20.7 %)	17 (53.8%)
Men	48 (58.5%)	17 (53.8%)
Education		
Less than High School	20 (24.4%)	9 (28.2%)
Completed High School	18 (22.0%)	9 (28.2%)
Post High School education	36 (54.6%)	18 (56.4%)
Language Spoken at birth		
English	34 (41.5%)	17 (53.1%)
French	28 (34.1%)	15 (46.8%)
Neither	20 (24.4%)	-
First stroke	72 (87.8%)	28 (76.9%)
Ischemic/Haemorrhagic	75 (91.5%) / 7 (8.5%)	33 (92.3%)/ 3 (7.7)%
Side of lesion (%)		
Left	35 (42.7%)	8 (25.2%)
Right	46 (56.1%)	22 (68.8%)
Bilateral lesion	1 (1.2%)	2 (6.0%)
Stroke Severity (Mean ± SD)	8.2 ± 2.4	6.9 ± 2.9 *
Severe (0-5)	11 (13.4%)	11 (30.8%)
Moderate-high (5.5-9)	41 (50.0%)	13 (35.9%)
Moderate-low (9.5-10.5)	17 (20.7%)	8 (23.1%)
Mild (11-11.5)	13 (15.9%)	4 (10.3%)
Barthel Index at discharge means SD (0-100)	**52.6 ± 27.3**	**41.7 ± 28.3**

*p<0.05; Canadian Neurology Scale (CNS) 0-11.5;**

Table 4.2: Items of Apathy Scale and corresponding codes

#	Items Description	3-digit level (%agreement)	4-digit level (% agreement)
	Openness to experience		
1	Are you interested in learning new things	B126	**B1264 (71%)**
2	Does anything interests you	B126	**B1264 (71%)**
10	Are you indifferent to things	B126	**B1264 (86%)**
11	Are you unconcerned with many things	B126	**B1264 (86%)**
	Energy and drive function		
14	Would you consider yourself apathetic	**B130 (71%)**	B1301 (43%)
4	Do you put much effort into things	**B130 (86%)**	B1301 (57%)
12	Do you need a push to get started on things	**B130 (86%)**	B1301 (71%)
	Motivation		
5	Are you always looking for something to do	B130	**B1301 (100%)**
7	Do you have motivation	B130	**B1301 (100%)**
	Energy level		
8	Do you have energy for daily activities	B130	**B1300 (100%)**
	Emotional function		
13	Are you neither happy nor sad, just in between	B152	**B1522 (86%)**
	Other categories		
9	Does someone have to tell you what to do each day	**D177 (71%)**	-
3	Are you concerned about your condition	B164	**B1644 (83%)**
6	Do you have plans and goals for the future	**B164 (86%)**	B1641 (71%)

B126- Temperament and personality function; **B1264** - Openness to experience; **B130** – Energy and Drive Function; **B1301** – Motivation; **B1300** – Energy Level; **B152 –** Emotional function; **B1522**- Range of Emotion; **B1644** - Insight; **B164** - Higher-Level cognitive Function; **B1641** – Organization and planning; **D177** – Making Decision

Table 4.3: Summary of Global Fit Statistics for the Rasch Model of the original AS and the final modified version

	Original 14 item model (n=232)	Final model (n=229)
Item-Trait Interaction		
Total item Chi-Square	211.8	6.4
Total Deg of Freedom	42	12
Total Chi-Square Probability	0.00000	0.89
Item-Person Interaction		
Item		
Difficulty	0.00 ± 0.42	0.00 ±0.37
Fit Residual	0.38 ± 2.64	0.16 ± 0.72
Person		
Measure of ability	0.70 ± 0.66	1.13 ± 1.23
Fit Residual	-0.22 ± 1.29	-0.46 ± 1.28
Reliability		
PSI	0.714	0.730

Table 4.4: Results of item analysis of the original AS

Item #	N° of Threshold	Item Fit	Local dependency	PCA	DIF
1	3	p <0.05	Item 2	-	
2	**			-	
3	**	+ 7.2*	Item 4	-	
4	4	p <0.05		-	
5	**			-	Age, Education
6	**	- 2.6*		-	Education, Language, Test language
7	**	p <0.05		-	Language, Test language
8	3			-	Gender, Language, Test language
9	**	p <0.05		+	Gender, Education
10	3		Items 11,12,13,14	+	Language, Test language
11	3			+	Language
12	**		Item 13	+	Language, Test language
13	3			+	Language
14	**			+	

** Disordered thresholds; * fit residuals value; PCA – Principal Component Analysis; DIF – Differential Item Functioning; Language – language first spoken at birth (English, French or Neither); Test Language – Language of the measure (English or French)

Table 4.5: Item location ranking overall time points and for each time point separately

Item #	All time points	3 months	6 months	9 months	12 months	Sum of ranks across 4 time points
9	1	1	1	1	3	6
14	2	3	2	4	1	10
4	3	2	5	3	2	12
10	4	5	3	5	5	18
12	5	7	4	2	4	17
11	6	4	7	6	6	23
2	7	6	8	7	7	28
8	8	9	6	8	9	32
13	9	8	10	9	8	35
1	10	10	9	11	10	40
6	11	11	11	10	11	43
3	12	12	13	14	12	51
5	13	13	12	12	13	50
7	14	14	14	13	14	55

The expected sum of ranks calculated as m(n+1)/2 where m is the number of time points and n is the number of items; here this value is 4(14)/2 or 28. The sum across all items of squared deviances from 28 (S) is 3486. Kendall's coefficient of concordance (W) is calculated as: $W = 12S / m (n^2-1)$ and yields a Chi-square value of 105.6 with 9 degrees of freedom ($p<0.01$) providing convincing evidence of concordance in item ranking across time points.

Table 4.6: Steps taken to refine AS

	Step	Specifics
1	Deleted items	3,6,9 and 14
		Item 6 could not be deleted
2	Collapsing of Thresholds	3 response options for 8 items ('Not at all', 'Some' and " A lot')
		1 binary 'yes' or 'no', item 7
3	Investigated residual local dependency	Items 1 and 2; 10 and 11; and 6 and 7
4	Created "super items" to deal with local dependency and DIF	Openness to experience (items 1, 2, 10 and 11)
		Energy and Drive Function (items 4, 5, 8 and 12)
		Motivation and Planning (items 6,7)
5	Ruled out multidimensionality	Significant t-test for < 5% of the sample
6	Verified item DIF	Openness to experience had DIF for language spoken and was split between English and French

Table 4.7: The hierarchy of the individual items retained into the Rasch version of AS (R-AS)

			Response options		
Item #	**Item Mean Location**	**Item:**	**0**	**1**	**2**
4	-0.79	Do you put much effort into things		1 (-2.76)	13 (1.16)
11	-0.23	Are you unconcerned with many things		2 (-1.68)	14 (1.22)
12	-0.46	Do you need a push to get started on things		3 (-1.66)	12 (0.73)
8	0.20	Do you have energy for daily activities		4 (-1.65)	16 (1.50)
10	-0.48	Would you consider yourself apathetic		5 (-1.61)	11 (0.64)
2	-0.02	Does anything interest you		6 (-1.57)	15 (1.53)
1	0.54	Are you interested in learning new things		7 (-0.66)	17 (1.75)
7	-0.62	Do you have motivation		8 (-0.62)	18 (2.06)
6	0.61	Do you have plans and goals for the future		9 (-0.51)	16 (1.74)
5	1.27	Are you always looking for something to do		10 (0.06)	19 (2.48)

Table 4.8: Rasch version of AS (R-AS) with Rasch scoring algorithm

R-AS			
1. Openness to experience	**total score = 18/**		
1.1. Are you indifferent to things?	**not at all** 13	**some** 1	**a lot** 0
1.2. Are you unconcerned with many things?	**not at all** 15	**some** 1	**a lot** 0
1.3. Does anything interest you?	**not at all** 0	**some** 2	**a lot** 17
1.4. Are you interested in learning new things?	**not at all** **0**	**some** **6**	**a lot** **18**
2. Energy and Drive	**total score = 18/**		
2.1. Do you put much effort into things?	**Not at all** 0	**some** 1	**a lot** 14
2.2. Do you need a push to get started on things?	**Not at all** 12	**some** 5	**a lot** 0
2.3. Do you have the energy for daily activities?	**Not at all** 0	**some** 5	**a lot** 17
2.4. Are you always looking for something to do?	**Not at all** **0**	**some** **10**	**a lot** **18**
3.Motivation and Plan	**Total score = 6/**		
3.1. Do you have motivation?	no 0	yes 1	
3.2. Do you have plans and goals for the future?	**not at all** **0**	**some** **1**	**a lot** **6**
Total Score 42/			

The items in bold (1.4, 2.4, and 3.2) represent the item with the best spread across the respective domain

Table 4.9: Comparison of AS and R-AS on variables indicating construct validity

	AS	R-AS
	N=82	N=82
Total score (0-42)[1]		
Mean ± SD	28.6 ± 6.7	25.9 ± 6.4
Proportion in apathy range		
N (%)	35 (42.7%)	60 (72.8%)*
95% CI	(31.9%, 54.0%)	(60.7%, 81.0%)
Depression** (Mean ± SD)***		
Not depressed (n=66)	29.1 ± 6.5	26.8 ± 5.7
Depressed (n=15)	26.5 ± 7.3	21.9 ± 7.6
Effect size of depression (Cohen's d)	0.38	-0.73*
Correlation with key measures	r (95%CI)	r (95%CI)
MHI	0.35 (0.14, 0.53)	0.46 (0.27, 0.61)
MOCA	0.21 (-0.03, 0.41)	0.03 (-0.18, 0.24)
2 min walk test	0.20 (-0.01, 0.40)	0.20 (-0.02, 0.40)
SIS participation	0.35 (0.14, 0.52)	0.45 (0.26, 0.61)

[1] higher score is better; * McNemars' chi-square p<0.001; ** Cut-point for depression <60 in the MHI RAND-36; *** p<0.05 for between subject, within subject and interaction in the repeated measure ANOVA;, MHI – Mental health index measure; MOCA – Montreal cognitive assessment; SIS – Stroke impact scale.

Figure 4.2: Graphical Rasch Model

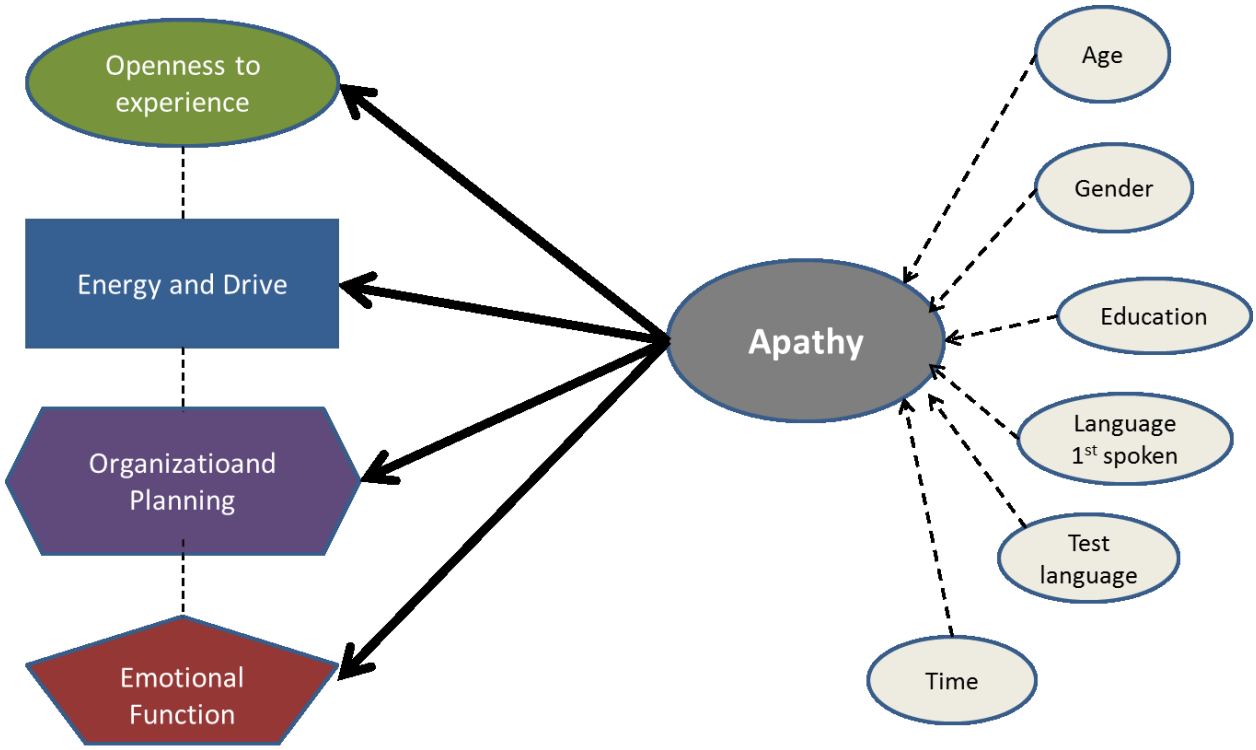
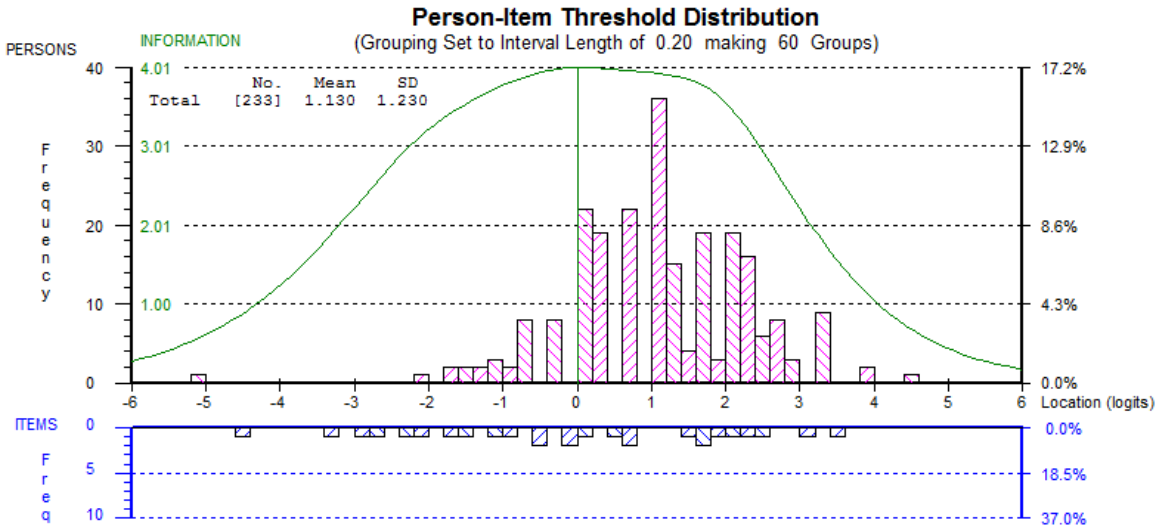


Figure 4.3: The item-person threshold distribution and test information function for the R-AS.



The horizontal axis, scaled in logits, denotes motivation symptoms, from least at the left to most at the right. The vertical axis denotes the frequency. The bars represent the distributions of subjects and items at each location. The line in the top of the figure represents the Test Information Function (TIF). An items' information function is the inverse of the item standard error squared; a TIF is the sum of the item information function.

Figure 4.4: Bland-Altman plot of the difference between AS and R-AS estimations

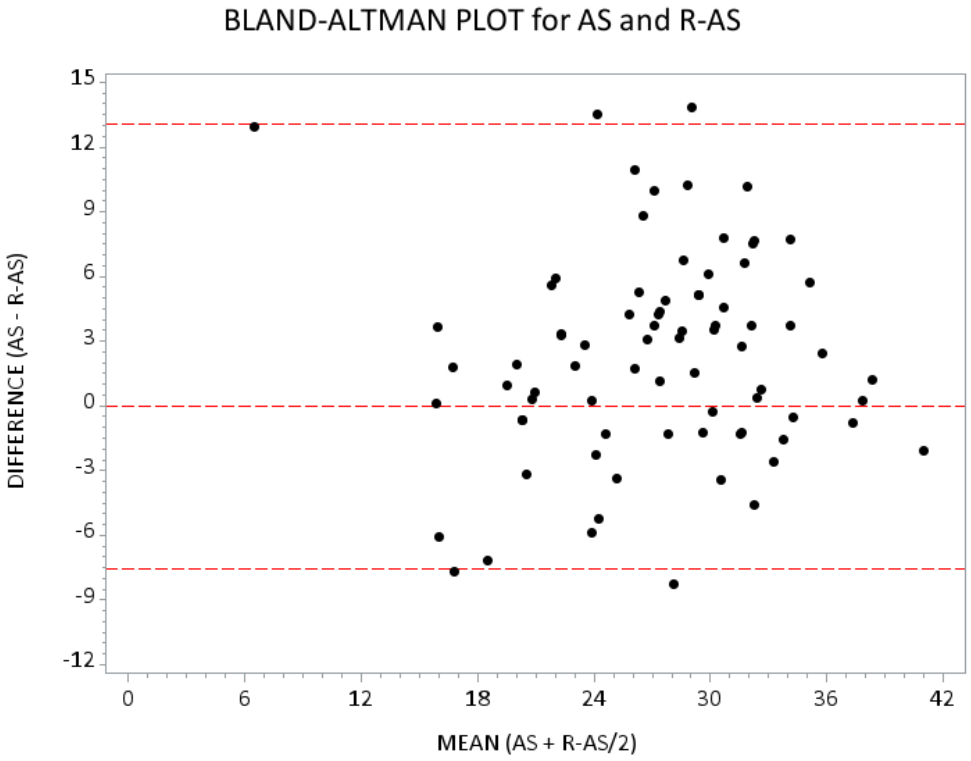
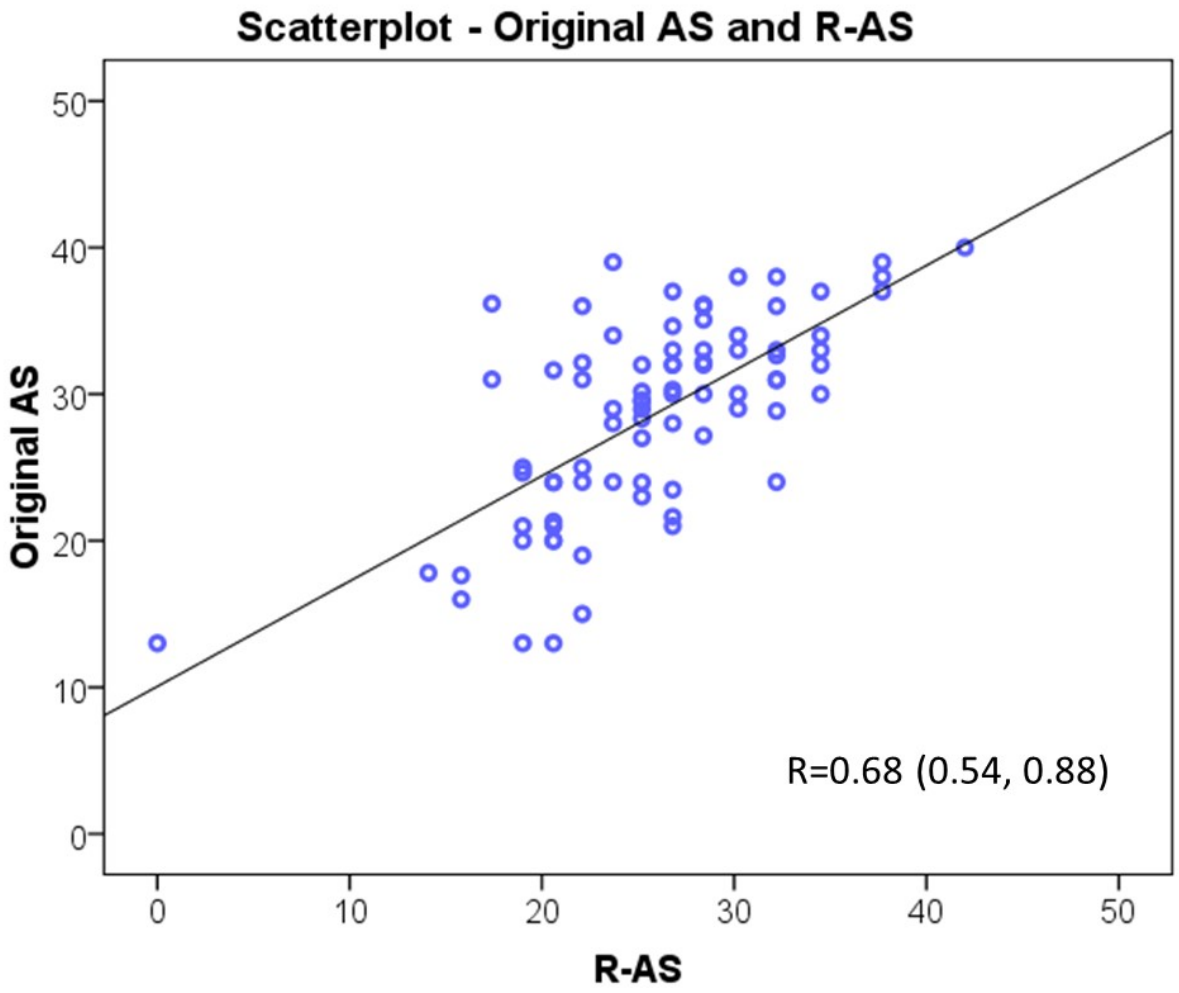


Figure 4.5: Scatterplot and correlation for Original AS and Rasch version of AS (R-AS)



CHAPTER 5- MANUSCRIPT3: Improving the Measurement of Apathy Using Rasch Analysis: An Example from Stroke

PREFACE TO MANUSCRIPT 3

Manuscript 2 identified major limitations in the AS that would compromise measurement of apathy. One important limitation identified, by ICF mapping and Rasch analyses, was that the items from the AS do not cover the whole spectrum of the apathy-motivation continuum. The AS items that fit the model and formed the R-AS, reflect only the motivation continuum: openness to experience, energy and drive function, and organization and planning; which are mental functions necessary to start, energize, sustain and direct a behaviour.

Table 5.0 summarizes the conceptual limitations of the original AS and the R-AS. For example, the AS has four items for each of “openness to experience” and “energy and drive”, but just one for “emotional function”; three items covered other domains not part of the theoretical framework. In contrast, the R-AS retained no items for emotional function.

A number of statistical decisions were made in manuscript two that were repeated in the next manuscript. The decision to identify parameters using an all-time-point model was made based on wanting to have as much data as possible to have stable estimates, no

DIF across time, and subsequent validation of the item hierarchy between the all-time-points model and the four time specific models, see table 4.5. In addition, we demonstrated the known phenomenon that repeated measures treated as independent observations produces an underestimation of the error associated with the parameter of interest. Appendix A8 presents the summary of the global fit statistics for the all-time-points model and each of the time specific models. As expected because of increased power of the all-time-point model, fit was rejected but accepted for the time specific models. All other indices of global fit were similar for all models. Appendix A9 presents item difficulty for the all- time-points model and the time-specific models and, as expected from the larger sample size, the standard errors associated with the item difficulties were smaller. But as presented in Manuscript 2, Table 4.5, the ranking of item hierarchy was concordant across models. Further discussion on analysis of repeated measures is presented in the final chapter.

The objective of Manuscript 3 was to identify if items from closely related constructs, like mood and emotion, reflecting the emotional function component of the apathy-motivation continuum could fit onto the model and form a valid measure.

Table 5.0: Construct roadmap to date

	Number of items per domain						
	Openness to experience	Energy and Drive Function		Emotional Function	Motivation and Planning	Other	Total
		Energy Level	Motivation				
Theory	✓	✓	✓	✓	✓	×	
AS	4	4		1	2	3	14
R-AS	4	4		0	2	0	10

Theory – Apathy construct conceptual theory; AS – original Apathy Scale; R-AS – Rasch Version of the AS; RMA- Rasch Measure of Apathy; *Not a component of apathy construct.

TITLE PAGE

Improving the Measurement of Apathy using Rasch Analysis: An Example from Stroke

For submission to: Neurology

Authors: Christiane B. Lourenço, M.Sc.,[1]Alyssa Herzig, Nancy E. Mayo, Ph.D.[1,2,]

[1] School of Rehabilitation Sciences, Faculty of Medicine, McGill University, 3654 Prom Sir William Osler, Montreal, Quebec, H3G 1Y5, Canada.

[2] Division of Clinical Epidemiology, Royal Victoria Hospital, McGill University Health Center, Montréal, Québec.

Communication address to: Christiane Lourenco, Royal Victoria Hospital, Division of clinical epidemiology, 687 Pine Avenue West, Ross 4.27, Montreal, Quebec, H3A 1A1 Canada. Telephone: (514) 934-1934 ext 37066, Fax: (514) 843-1493, Email: christiane.lourenco@mail.mcgill.ca

Key Words: apathy, motivation, Rasch measurement theory, rehabilitation

ABSTRACT

Objectives: To identify the extent to which enriching the motivation continuum, identified on the Rasch Version of Apathy Scale (R-AS), with items taping the emotional function domain can lead to a way of quantifying the measurable components of the apathy syndrome criteria.

Methods: A secondary data analysis from a prospective longitudinal study on stroke depression was carried out. Items were selected from three measures based on the conceptual model of apathy. A Confirmatory Factor Analysis (CFA) was conducted to verify if the selected items belonged to one of the hypothesized categories (emotional function, openness to experience, and energy and drive function).Rasch analysis was conducted in a step-wise manner, to estimate the extent to which those items fit the underlying theoretical Rasch model of threshold ordering, item fit, invariance, unidimensionality, and targeting.

Results: Twenty-three items from three different measures of mental health and mood were identified to be reflecting the motivation and emotional components of the apathy continuum. Those items and the10 R-AS items were included in Rasch Analysis. Based on Rasch analysis, three emotional items were identified to have the best spread across latent trait and fit. The 13 items fit the Rasch measurement model and demonstrated unidimensionality, local independency, and lack of DIF. This new measure was denominated Rasch Measure of Apathy (RMA)

Conclusions: In conclusion, the RMA might be a potential solution for measuring Apathy Syndrome because it spans the apathy-motivation continuum, and has 10 of the original 14 Apathy Scale items but with much improved psychometric properties. This version could be used in lieu of the Apathy Scale for clinical and in research purposes even though further testing is needed.

INTRODUCTION

Motivation has an important role in our lives; it is responsible for the initiation, direction, energy, and effort indispensable to the achievement of desired goals. Therefore, lack of motivation is expected to have a negative effect in a persons' life, and even more so if this person needs to recover from stroke impairments and disabilities. To fully benefit from stroke rehabilitation, the patient needs to participate actively, understand the rehabilitation process, take initiative for therapeutic exercises, and adhere to the therapist recommendations. Recovering from stroke requires tremendous effort and therefore, stroke patients require an additional level of motivation to overcome the stroke impairments [75].

While motivation is a positive construct, apathy is the term used to describe the negative side of motivation. Inherent to the motivation continuum are measured impairments in the domains of "openness to experience", "energy and drive" and "plans and goals for the future". Blunting of emotional response is a necessary additional component to the motivation-apathy continuum, but it is not sufficient for a person to be considered as having the apathy syndrome as it requires three additional criteria:(i) decreased motivation in relation to previous level of functioning; (ii)significant impairments in activity and participation; (iii) ruled out diminished level of consciousness, emotional distress, and dementia [20, 21, 24].

Our previous work showed that 10 out of the original 14 items of the AS, the most frequently used measure of the apathy syndrome in the stroke population, were able to capture the motivation continuum. However, to create a method to measure the severity of the apathy syndrome, items reflecting emotional response are also required. Thus, the purpose of this study is to identify the extent to which enriching the motivation continuum, identified on the Rasch Version of Apathy Scale (R-AS), with items tapping the emotional function domain can lead to a way of quantifying the measurable components of the apathy syndrome criteria.

METHODS

Participants

A secondary data analysis was carried out on people with stroke recruited into a prospective study on depression. Subjects and procedures have been described previously in Manuscript 2. Briefly, the sample comprised people admitted for acute-stroke to a university teaching hospital in a major Canadian metropolitan area. Participants were evaluated at 10 days post-stroke and subsequently at 3, 6, 9 and 12 months post-stroke by trained health care professionals who assessed participants' levels of activity and participation, as well as administering questionnaires on mood and emotional function.

Measures

The original study protocol included the Apathy Scale (AS) [29], an index comprising 14 self-reported items. However, the analyses described in Manuscript 2 indicated a number of measurement weaknesses in the AS. By applying Rasch Analysis, the measurement properties of the AS were improved and the new format was denoted as the Rasch Measure of Apathy (R-AS). The R-AS comprised 10 of the original AS items but rescored into three multi-item sub-scales reflecting three domains that fit with a motivation construct: openness to experience (4 items); energy and drive function (5 items); and organization and planning (2 items). As data fit the underlying Rasch model, a meaningful total score could be derived by summing the scores on each of these three super-items. To fit with the original AS scoring, the R-AS was transformed to range from 0-42, with high score indicating greater motivation, the Rasch Measurement psychometric criteria were partially supported, PSI was low (0.70). Scale-to-sample targeting of the data was adequate (X^2=6.4, DF=12, p=0.73).

The items in R-AS do not tap the emotional function domain and, hence, cannot be considered a measure of the apathy-motivation continuum. In an attempt to build an emotional function domain, items were selected from other measures included in the assessment battery such as: Stroke Impact Scale (SIS), Geriatric Depression Scale (GDS), and RAND-36 (Mental Health Index).

The SIS [147] is a 59-item scale that is used to assess eight domains: strength, memory, emotion, communication ability, basic instrumental activities of daily living (ADL), mobility, hand function and social participation. For the purpose of this study only the

emotion and participation domains were used. Each item of a domain is measured on a 5-point Likert scale: 5 ‘not difficult at all’, 4 ‘a little difficult’, 3 ‘somewhat difficult’, 2 ‘very difficult’, and 1 ‘extremely difficult’; in which each domain score is transformed on a scale of 0-100. The SIS has evidence for reliability and validity [147].

The Mental Health Index (MHI) is a subscale of the Short Form Survey from the RAND Medical Outcomes (RAND-36) [148] with 5 items referring to feelings in the past four weeks. Items are scored on a 6-point Likert scale ranging from 1 “all the time”, and 6 “None of the time”. The score ranges from 0 to 100, in which a high score indicates better mental health. Reliability, both test-retest and internal consistency, has been extensively demonstrated for the RAND-36, as have content, criterion and construct validity and responsiveness to clinical changes [148].

The Stroke Depression Ladder (GDS) [149] is a stroke-specific version of the original 30-item Geriatric Depression Scale [150] mathematically developed using Rasch analysis. It uses simple phrases to represent the construct of depression with yes (0) or no (1) response options. The items range from “easiest” to “hardest”, and the scores from 0-17; a high score indicates severe depression. However, for the purpose of this study, the scale scoring was reversed so that a higher score would indicate better mental function similarly to the other measures.

Item generation

The three available measures provided 31 items of mood and emotion to be selected for creating the emotional function domain based on the conceptual model of Apathy (Figure 4.1). Any item reflecting the categories of “openness to experience”,” energy and drive” function, “Organization and Planning”, and especially “emotional” function, were selected for exploratory analysis. Where appropriate, item scoring was reversed so that all measures would have high scores representing the positive end of the construct (e.g., good mental health).

A Exploratory Factor Analysis (EFA) using SPSS 17.0 was conducted to verify if the selected items could be grouped together into subscales based on the hypothesized categories of emotional function, openness to experience, energy and drive function, or organization and planning; or if they should be dropped from the instrument entirely [151]. The criterion for considering which items went within a category was based on a cut-off factor loading of 0.4. If an item was hypothesized to belong to a certain category but showed loading inferior to 0.4, the analysis was repeated with items from other categories to determine which category that item showed the highest loading.

Then, a Principal Component Analysis (PCA) was conducted for all items together, except those excluded after EFA, to have a first overview of the measure structure, how the items group together and how many dimensions exists. The criteria for dimensionality at this stage was not strict, just intended to identify a set of possible items to be included or excluded from Rasch Analysis, with the purpose of data reduction and to be subsequent validated with Rasch analysis.

PCA was considered an appropriate method, if the value on the Kaiser-Meyer-Olkin measure of sample adequacy (KMO) was > 0.6 and Bartlett's Test of Sphericity was ≤ 0.05. The number of factors obtained was based on Parallel Analyses (PA) [152] which simulates predicted Eigen values from 100 random data sets. The predicted Eigen values were compared with the actual Eigen values and only factors with Eigen values greater than those derived randomly were retained. Once the numbers of factors were determined, an orthogonal Varimax rotation was conducted to facilitate interpretation.

Statistical Analysis

All Items that fit the conceptual criteria were analysed by RUMM 2030 [135] software to estimate the extent to which each the selected items fit the underlying theoretical Rasch measurement model. All time points were used in the analysis to yield a total sample size of 232 observations.

Threshold ordering, item fit, invariance, unidimensionality, and targeting were assessed (see Manuscript 2 for more details).

The association between the original AS, the R-AS, and the new measure of apathy that emerged from Rasch analysis in this study (RMA) with other measures of key construct (MHI, MOCA, 2 min walk test, and SIS participation) were compared using Pearson Coefficient of Correlation. A Bland-Altman plot was carried to identify measurement bias. McNemars' chi-square was used to compare the percentage of people classified

as apathetic on the three forms of the measure. Repeated measures analysis of variance was used to compare the values on the three measures across levels of depression, where the cut-point was greater or less than 60 on the MHI of the Rand-36.

RESULTS

Participant's characteristics

Table 5.1 compares the characteristics of the Apathy assessed cohort to those excluded from the cohort at 3 months post stroke. Although there were some differences between those administered the apathy scale and those that were not (age, severity of stroke, side of lesion) these occurred by chance. The missing data was a consequence of design, as the AS was introduced later in the study.

Item Generation

Based on the conceptual model, two of the authors identified 26 items out of the 31 items available from the three different measures of mental health and mood, to be reflecting the motivation continuum or the emotional function components of the apathy-motivation continuum. These 26 items and the 10 R-AS items were than classified into 6 conceptual categories (happiness, sadness, anxiety, openness to experience, energy and drive, and motivation and plan).Table 5.2 display the result of the EFA for the 36 items showing the 6 categories and the items with the highest loading for each category

(values of the1st component extraction, Eigen, and percentage of variance).Only three items (GDS 2 – openness to experience; 7 and 14 - happiness) did not load to the expected category, or to any other category, and were dropped from further analysis. The remaining items all loaded strongly with the other items in the expected category which suggested that they were reflecting the hypothesized concept.

PCA analysis of the remaining 33 items revealed that three factors solution was best yielding a pattern of loading, with eigenvalues greater than those derived randomly in parallel analysis. As seen in Table 5.3, Eigen values ranged from 1.85 to 8.77, and explained together 40.6% of the variance. The first component includes mainly anxiety and sadness items. The second component motivation and enjoyment items, and the third component did not represent any specific category. As the R-AS items loaded in all three components, items from all three components were included in Rasch analysis for subsequent analysis.

Rasch Analysis

The 33 items did not fit the Rasch model. There were disordered thresholds for 15 items (Appendix A12), local dependency between 6 pairs of items, DIF by language first spoken and language of the test, and finally PCA loaded all emotion items into one subtest and all R-AS into another subtest (Appendix A13). The two subtests yielded different estimations for more than 21.6% of the sample.

After rescoring the thresholds, and subtesting the items into 4 categories: openness to experience (R-AS 1.1, 1.2, 1.3, and 1.4), energy and drive function (R-AS 2.1, 2.2, 2.3, and 2.4), motivation and plan (R-AS 3.1 and 3.2), and emotional function (remaining 23 items); multidimensionality was still observable in which the two components estimated different ability for 11.4% of the sample (CI 8.4%, 14.3%).

It is clear that too many emotional function items were redundant and changed the latent construct towards depression. Therefore, the 23 emotional function items were trimmed to retain only one item from each sub-category of emotional function (anxiety, sadness, and happiness). The three Items that best spread across latent trait, showed strong loadings in PCA analysis, and were easy to comprehend by patients were chosen to form the emotional function subtest (SIS 3a ‘feel sad’, SIS 3g ‘feel quite nervous’, and MHIh ‘have you been happy’).

The model with 13-items and 4 subtests satisfied the requirements of unidimensionality and local independency. Lack of DIF was achieved only after ‘openness to experience’ subtest was split by gender, and ‘motivation and plan’ subtest by level of education (less than high school and more than high school). Summary statistics for the final model is displayed in Table 5.4 (chi-square 27.8, p=0.06). Figure 5.1 shows the targeting of the final items in relation to this sample population. The level of difficulty of the items shows a good spread across the latent construct, with items measuring high and low levels of apathy. The stroke sample in the present study had lower level of apathy, spreading

between 0 and 4 logits. The internal consistency of the measure (PSI 0.684) is very close to the minimum value required for group use (0.70).

The hierarchy of the 13 individual items is displayed in Table 5.5. The item threshold difficulty ranges from -2.66 logits to +2.32. As there were four subtests, four separate scores were calculated and can be summed to a total score. To make comparisons between the original AS and the R-AS, the raw scores were rescored to have a maximum value of 42, in which higher levels indicate less apathy. The scoring algorithm for the RMA is shown in Table 5.6, in which the highest score for the openness to experience category is 12 by endorsing that a person is always interested in learning something new. The highest score for energy and drive function is 12, for emotional function is 8, and for motivation and plan is 6 adding to maximum score of 42. The non-uniform scoring across response options and items indicates the different location of each threshold on the linear continuum.

Comparisons between original and Rasch versions of the AS

Figure 5.2 presents the Bland-Altman plot comparing the motivation measure (R-AS) and the new apathy measure (RMA). The y-axis is the magnitude of the difference between the AS and the R-AS (R-AS minus RMA). The x-axis is the average of the two measures. The dotted line shows the average mean difference between the measure (0.07 ± 4.6) and 95% CI (-9.0, 9.2). The solid line at 0 indicates no difference between the two measures scores. The dots represent individual subjects, and for the most part,

they are evenly distributed above and below the zero line, indicating similar ability estimates. However, at the high motivate end of the latent construct (formed by the average of the 2 measures) the R-AS gives people higher scores than RMA. Only one subject falls below the 95% confidence band for measurement error.

Table 5.7 compares the new measure of apathy, the measure of motivation developed in Manuscript 2*,* and the original AS on indicators of construct validity. These results support the Bland-Altman plot, as the averages of the RMA and RAS are equal and lower than the original AS (Table 5.7; Appendix A14). The RMA classified the highest number of people falling in the apathetic range (81.7%), and the original AS the lowest (42.7%; $p<0.001$). There was no significant difference between R-AS and RMA in the proportion of apathetic/low motivated subjects. Also shown are the average scores for people classified as depressed or not depressed using the MHI of the RAND-36 (cut-point of 60). The difference in the score between those not depressed and depressed in the original AS was 2.6 yielding a discrimination effect size of 0.38,significantly different from both R-AS (0.72) and RMA (0.88) ($p<0.05$). Overall, RMA is the measure that best discriminates between participant's level of depression. The RMA also shows better correlation with depression than the other two measures. R-AS and RMA show similar correlation with mobility, participation and cognition.

The correlation between the original AS and the RMA and the R-AS and RMA are illustrated in Figure 5.3 and 5.4 respectively, showing a very strong positive linear

correlation (r=0.82) between original AS and RMA and only moderate correlation between R-AS and RMA.

DISCUSSION

The results of this study showed that a measure of apathy-motivation continuum can be constructed from four sub-tests covering each of the essential components of motivation (openness to experience, energy level, organization and planning) and emotional function (anxiety, happiness, and sadness). The 13 items fit the Rasch measurement model and demonstrated unidimensionality, local independency, and lack of DIF; the items targeted the population under study well and covered the apathy continuum. The resulting measure was named the Rasch Measure of Apathy (RMA).

The conceptual and Rasch analysis indicated that the measurable components of the apathy syndrome are very heterogeneous concepts and a valid measure of apathy-motivation continuum was only achieved when the individual items were grouped into conceptual subtests in which the score of each subtest can be added to form a total score. Similar to what was observed in the analysis of the original Apathy Scale (AS) in Manuscript 2. These findings support that the apathy-motivation continuum is very complex and not easy to measure.

In constructing the RMA, our approach was to include items from other scales, and the large number of added items reflecting emotion and mood drove the latent towards a

depression construct., rather than apathy construct. When all 33 items were included, the items of motivation continuum no longer fit, showing that the latent trait was defined by the emotional items. This illustrates that the latent trait is defined by the items and reinforces the need to have a strong conceptualization for an item bank.

To ensure that the apathy-motivation continuum was the one we were measuring, we started with only the motivation continuum items of the R-AS (as described in Manuscript 2), as a measure core set . Then a large number of items were tested to form the essential emotional function sub-test, necessary to fulfil the apathy criteria. Three items, (SIS 3a and 3g, and MHI h) were identified to best fit the construct. The RMA has 13 items with a nice balance across domains.

Apathy syndrome and depression are closely related constructs and sometimes are difficult to distinguish. Studies have established that these are two independent syndromes that can coexist [19, 48, 131]. The main difference is the unique emotional pain (distress) experienced by depressed patients. Apathetic patients tend to be more passive and compliant, and show lack of emotional distress and response [20, 153]. To define the emotional domain we identified items that would indicate the presence or absence of emotional response, like feeling sad, or anxious, or happy to not confound with depression. An apathetic patient would be expected to answer ‘none of the time’ for all those items, or even be oblivious to his feelings.

It was also interesting to find that the RMA was the measure that best correlated (0.53) with the measure of mental health (MHI), this correlation indicates a moderate relationship with depression, but not very strong that would indicate measuring the same construct. The RMA also showed a very strong correlation with the original AS (0.82) in comparison to only strong correlation between AS and R-AS (0.68) suggesting that the emotional items added to the R-AS may have changed the latent trait of motivation, as demonstrated in Manuscript 2, back to apathy syndrome. The correlation between R-AS and RMA was also 0.68, indicating strong correlations but somewhat different construct.

On the other hand, adding emotional items to the R-AS did not make any significant change between R-AS and RMA estimations. This was very clear in the Bland-Altman plot, and further confirmed by descriptive statistics (Table 5.7). The RMA showed only slightly better discrimination between apathetic and not apathetic (81.7%) when compared to R-AS (72.8%); and between those depressed and not depressed (effect size 0.88). The new measure also did not correlate with cognition which is a desirable feature in a measure.

The similarities between R-AS and RMA estimations raise questions whether emotional function, an essential component of the apathy syndrome criteria, is relevant for the definition of apathy. According to Marin [21, 154] the presence, quality, and dynamics of emotional response (e.g. anger and happiness) can provide information on the extent to which having lack of motivation is of significance to the patient. Do patients feel sad

about not being interested? Are they unhappy with the condition? A “true” apathetic patient (i.e. with apathy syndrome), would be indifferent. Therefore, this domain would be essential for differential diagnosis, (e.g. depression) and for estimating the extent of motivation impairment. Nevertheless, this distinction of apathy syndrome and apathy state may be relevant for neurorehabilitation if apathy syndrome is shown to have a greater impact on functional recovery and engagement to therapy than the temporary apathy state.

It would have been desirable to use both measures in a completely new sample of patients and test if they can predict poor functional recovery in a similar way, or if having apathy syndrome, rather than symptoms along the apathy-motivation continuum, cause a more significant impact on function and recovery. Further research is needed to clarify the distinction between these concepts and how they impact on the rehabilitation process.

The results from this study yielded two approaches for measuring the motivation/apathy continuum; the RMA with 13-items spans the latent construct with enough precision at each level of the apathy domain. In addition, we identified one item in each subtest that best represented the spectrum of the intended domain. These four items could be used as a short form to categorize people into having low, moderate, or high apathy, so that further diagnostic criteria investigation could be done.

This is the first time that apathy has been tested applying modern psychometric theory; however, a number of other constructs in the mental health domain have been Rasch analysed, including HADS [155], GDS [149], DASS [156], and Beck Depression inventory II [121]. Item banks for self-reported outcomes on mental, physical and social health components have also been developed by PROMIS [157], based on IRT models, but there is not one for apathy. The results from this study will provide a starting point to understand the characteristics and importance of the apathy syndrome and impaired motivation in stroke rehabilitation.

The study has some important limitations that need to be addressed. Those related to the construction of the cohort and designed in missing data have been discussed at length previously (Mayo et al.[145] and Manuscript 2). The cut-off score adopted in this study was based on Starkstein's criteria, in which uses a scores greater or equal to 14 (on a scale from 0-42) to classify a person as apathetic [29]. For the purposes of this study, the items were reversed scored, so a high score would indicate no apathy. In this case, a cut-off score ≤ 29 was shown by Rasch analysis to be the same as the ≥ 14 in the original measure. This cut-off seems to estimate higher prevalence's (81.7%) of apathy than the original AS (42.7%) and the average identified in studies with stroke population (33%). Further work would need to be conducted to develop valid cut-points for the RMA.

The major limitation in this study is that the measure was developed based on existing items. The best available methodology was used to deal with the existing items,

including the application of a strong conceptual model. However, it would have been optimal to develop items based on solid conceptualization and measurement theory and input from patients, caregiver, and other health professionals.

While the FDA provides a process for building a new measure [158], the work in this paper provides a strong indication of what the content for a good measure should be. First, the content needs to cover the four domains of the apathy construct. Second, the items should reflect the granularity of the construct. For example, four items were reflecting the openness to experience domain, either interest (e.g. “does anything interest you?”), or lack of interest (e.g. “are you indifferent to things?”), and had to be combined into a subtest due to high correlation and redundancy. It would be more appropriate to have items asking if the person has interest in doing specific activities like novel experiences, participating in social events, curiosity and interest for life events. Similarly, the emotional function domain would be better represented by items reflecting the emotional reaction to activities and events. For example, “do you feel sad when you cannot do an activity that you were used to do?”, or “do you feel happy when you go out with friends or family?” The activity and participation category of the ICF would provide appropriate content to be queried.

In addition, based on Rasch Measurement theory an optimal measure of apathy would not have negatively worded items. This measurement approach is intended to detect response agreement pattern. However, patients tend to interpret negatively worded items as a different concept from positively items. As a result, analysis at the item level

detects two different constructs, violating the Rasch Measurement Model of unidimensionality [144].

Finally, the distribution in the item-person threshold distribution map (Figure 5.1) is normal but person level of ability is higher than item level of difficulty. There are more items covering the lower end of the continuum but there are more people at the higher end. This indicates, that if we want to measure change of motivation, more items at the higher end of the continuum should be added. If we want to detect low apathy, we can do that. The original version of the AS was not well targeted and person ability was higher than item difficulty (Appendix A5). The changes proposed improved the spread of person ability and item difficulty despite many fewer thresholds used for measurement. This indicates that there was considerable misclassification of the original scoring structure. Therefore, it would be important to test this in other samples before firm conclusions can be made.

Despite the limitations of the stepwise approach, the RMA may have much better psychometric properties than the available measures of apathy. As the items fit the Rasch model, it delivers unidimensionality, scoring order, additivity and objectivity; which are important attributes of a good measure.

CONCLUSION

In conclusion, we believe the RMA might be a potential solution for measuring Apathy Syndrome because it spans the apathy-motivation continuum, has 10 of the original 14 items from Apathy scale but with much improved psychometric properties. Future development in this area is needed as motivation is a driving force for all human action and the RMA, while better than the AS, is still not optimal. The results from this study should guide the conceptualization and operationalization of items that will best reflect the construct the apathy-motivation continuum and fit a Rasch model, including inputs from patients, caregivers, and experts.

Table 5.1: Comparison of the Apathy assessment cohort and excluded subjects on key characteristics at 3 months post stroke

Characteristics	Apathy cohort (n=82) N (%)	Excluded (n=36) N (%)
Age (years) at stroke (Mean ± SD)	68 ± 11	77.4 ± 14.6*
<60	21 (25.6%)	3 (10.3%)
60-80	44 (53.7 %)	11 (35.9%)
>80	17 (20.7 %)	17 (53.8%)
Men	48 (58.5%)	17 (53.8%)
Education		
Less than High School	20 (24.4%)	9 (28.2%)
Completed High School	18 (22.0%)	9 (28.2%)
Post High School education	36 (54.6%)	18 (56.4%)
Language Spoken at birth		
English	34 (41.5%)	17 (53.1%)
French	28 (34.1%)	15 (46.8%)
Neither	20 (24.4%)	-
First stroke	72 (87.8%)	28 (76.9%)
Ischemic/Haemorrhagic	75 (91.5%) / 7 (8.5%)	33 (92.3%)/ 3 (7.7)%
Side of lesion (%)		
Left	35 (42.7%)	8 (25.2%)
Right	46 (56.1%)	22 (68.8%)
Bilateral lesion	1 (1.2%)	2 (6.0%)
Stroke Severity (Mean ± SD)	8.2 ± 2.4	6.9 ± 2.9 *
Severe (0-5)	11 (13.4%)	11 (30.8%)
Moderate-high (5.5-9)	41 (50.0%)	13 (35.9%)
Moderate-low (9.5-10.5)	17 (20.7%)	8 (23.1%)
Mild (11-11.5)	13 (15.9%)	4 (10.3%)
Barthel Index at discharge means SD (0-100)	52.6 ± 27.3	41.7 ± 28.3

*p<0.05; Canadian Neurology Scale (CNS) 0-11.5;**

Table 5.2: Items selection for developing the Rasch measure of Apathy based on the ICF framework and exploratory factor analysis.

Index	Item #	Item Description	1st component extraction
Anxiety Category			
SIS	3g	Feel quite nervous?	0.79
MHI	B	Have you been a very nervous person?	0.77
GDS	6	Do you often get restless and fidgety?	0.68
MHI	D	Have you felt calm and peaceful?	0.66
GDS	4	Do you frequently get upset over little things?	0.62
GDS	8	Are you bothered by thought you can't get out of your head?	0.57
GDS	11	Are you afraid something bad is going to happen?	0.50
GDS	15	Do you worry a lot about the past?	0.42
		Eigenvalue	**3.2**
		% of variance	**40.7**
Happiness Category			
SIS	3h	Feel that life is worth living?	0.75
MHI	H	Have you been a happy person?	0.71
SIS	3f	Enjoy things as much as ever?	0.66
SIS	3i	Smile and laugh at least once a day?	0.64
GDS	16	Are you in good spirits most of the time	0.57
GDS	14	Do you think is wonderful to be alive now?	**
GDS	7	Are you basically satisfied with your life?	**
		Eigenvalue	**2.8**
		% of variance	**40.0**
Sadness Category			
SIS	3d	Feel that you have nothing to look forward to?	0.80
MHI	c	Have you felt so down in the dumps that nothing could cheer you up?	0.78
SIS	3a	Feel sad?	0.75
SIS	3b	Feel that there is nobody you are close to?	0.72
MHI	f	Have you felt downhearted and blue?	0.74
GDS	10	Do you feel that your life is empty?	0.69
GDS	13	Do you often feel helpless?	0.57
GDS	9	Do you frequently feel like crying?	0.56

		Eigenvalue	**3.4**
		% of variance	**49.6**
Openness to experience			
AS	2	Does anything interest you?	0.70
AS	1	Are you interested in learning new things?	0.69
AS	10	Are you indifferent to things?	0.64
AS	11	Are you unconcerned with many things?	0.60
GDS	2	Do you often get bored?	**
		Eigenvalue	**2.0**
		% of variance	**34.9**
Energy level			
GDS	1	Do you feel full of energy?	0.72
AS	8	Do you have energy for daily activities?	0.67
GDS	12	Do you enjoy getting up in the morning?	0.64
AS	12	Do you need a push to get started on things?	0.50
AS	4	Do you put much effort into things?	0.52
		Eigenvalue	**1.9**
		% of variance	**38.0**
Motivation and Plan			
AS	6	Do you have plans and goals for the future?	0.84
AS	7	Do you have motivation?	0.83
AS	5	Are you always looking for something to do?	0.66
		Eigenvalue	**1.8**
		% of variance	**61.0**

Table 5.3: Factor Analysis showing three major factor loadings for the 33 items identified in the construct conceptualization.

Item		Components		
#	Description	1	2	3
SIS-3g	Feel quite nervous	.730		
SIS-3a	Feel sad	.728		
MHI-c	Felt down in the dumps	.695		.308
MHI-f	Felt downhearted and blue	.688		
MHI-b	Have been a very nervous person	.664		
SIS-3b	Feel that there is nobody you are close	.646		
SIS-3d	Have nothing to look forward to	.639		
GDS-9	Feel like crying	.625		
GDS-10	Feel that your life is empty	.610		
GDS-4	Frequently get upset over little things	.543		
GDS-6	Often get restless and fidgety	.520		
MHI-d	Felt calm and peaceful	.447		.398
GDS-8	Bothered by thought in your head	.416		
GDS-11	Afraid something bad is going to happen	.406		
SIS-3h	Feel that life is worth living	.380	.368	.327
SIS 3-i	Smile and laugh at least once a day	.371	.362	
GDS-15	Worry a lot about the past	.309		
AS-6	Have plans and goals for the future		.808	
AS-7	Have motivation		.770	
AS-1	Have interest in learning		.713	
AS-2	Have interest in anything		.630	
AS-4	Put much effort into things		.628	
AS-8	Have energy for daily activities	.308	.572	
AS-5	Always looking for something to do		.502	
GDS12	Enjoy getting up in the morning		.324	
SIS-3f	Enjoy things as much as ever		.49	

AS-10	Indifferent to things			.708
AS-12	Need a push to get started			.667
AS-11	Unconcerned with many things			.645
GDS-13	Often feel hopeless			.539
GDS-16	Are you in good spirits			.474
MHI-h	Have you been a happy person	.452		.467
GDS-1	Feel full of energy			.466
SIS-3f	Enjoy things as much as ever		.63	.432
	Eigenvalue	**8.77**	**3.10**	**1.85**
	% of variance explained	**25.6%**	**9.4%**	**5.6%**

Table 5.4: First and final model summary of Rasch Model global fit statistics for the 13-items measure of apathy

	Rasch Measure of Apathy (13 items)
	Final Model N=232
Item-Trait Interaction	
Total item Chi-Square	27.8
Total Deg of Freedom	18
Total Chi-Square Probability	0.06
Item-Person Interaction	
Item	
Difficulty	0.00 ±0.42
Fit Residual	-0.02 ± 1.16
Person	
Measure of ability	0.69 ± 0.95
Fit Residual	-0.32 ± 1.05
Reliability	
PSI	0.684

Table 5.5: The hierarchy of the individual items retained into the Rasch Measure of Apathy (RMA)

			Response options		
Item #	**Item Mean Location**	**Item:**	**0**	**1**	**2**
R_AS 1.1	-0.76	Do you put much effort into things		1 (-2.66)	19 (1.13)
MHIh	0.33	Have you been happy		2 (-2.28)	25 (2.94)
R_AS 2.2	-0.47	Do you need a push to get started on things		3 (-1.65)	16 (0.70)
R_AS 1.1	-0.49	Would you consider yourself apathetic		4 (-1.59)	15 (0.60)
R_AS 2.3	0.23	Do you have energy for daily activities		5 (-1.59)	23 (2.05)
R_AS 1.3	-0.02	Does anything interest you		6 (-1.56)	20 (1.50)
R_AS 1.2	-0.23	Are you unconcerned with many things		7 (-1.52)	17 (1.05)
SIS 3a	-0.51	Feel sad		8 (-1.27)	14 (0.24)
SIS 3g	0.17	Feel quite nervous		9 (-0.74)	18 (1.09)
R_AS 3.1	-0.58	Do you have motivation		10 (-0.58)	18 (2.06)
R_AS 1.4	0.52	Are you interested in learning new things		11 (-0.56)	21 (1.61)
R_AS 3.2	0.60	Do you have plans and goals for the future		12 (-0.47)	22 (1.69)
R_AS 2.4	1.22	Are you always looking for something to do		13 (0.12)	24 (2.32)

Table 5.6: Rasch measure of apathy (RMA) with Rasch scoring algorithm

Rasch Measure of Apathy			
1. Openness to experience	**total score = 12/**		
1.1. Are you indifferent to things?	not at all 9	some 1	a lot 0
1.2. Are you unconcerned with many things?	not at all 10	some 1	a lot 0
1.3. Does anything interest you?	not at all 0	some 1	a lot 12
1.4. Are you interested in learning new things?	not at all 0	some 5	a lot 12
2.Energy and Drive	**total score = 12/**		
2.1. Do you put much effort into things?	Not at all 0	some 1	a lot 9
2.2. Do you need a push to get started on things?	Not at all 12	some 3	a lot 8
2.3. Do you have the energy for daily activities?	Not at all 0	some 3	a lot 11
2.4. Are you always looking for something to do?	Not at all 0	some 7	a lot 12
3.Emotional Function	**Total score = 8/**		
3.1 Are you happy?	Not at all 0	some 1	a lot 8
3.2 Are you sad?	Not at all 0	some 2	a lot 4
3.2 Do you feel quite nervous?	Not at all 0	some 3	a lot 6
4.Motivation and Plan	**Total score = 6/**		
4.1. Do you have motivation?	no 0	yes 6	
4.2. Do you have plans and goals for the future?	not at all 0	some 1	a lot 6
Total Score 42/			

The items in bold (1.4, 2.4, and 4.2) represent the best spread across the respective domain

Table 5.7: Comparison of AS, R-AS and RMA on variables indicating construct validity

	AS 14 items N=82	R-AS 10 items N=82	RMA 13 items N=82
Mean ± SD	28.6 ± 6.7	25.9 ± 6.4	25.6 ± 5.0
Proportion in apathy range [2]			
N (%)	35 (42.7%)	60 (72.8%)*	67 (81.7%)
95% CI	(31.9%, 54.0%)	(60.7%, 81.0%)	(72%, 88.6%)
Depression** (Mean ± SD)***			
Not depressed (n=66)	29.1 ± 6.5	26.8 ± 5.7	26.4 ± 4.8
Depressed (n=15)	26.5 ± 7.3	21.9 ± 7.6	22.3 ± 4.5
Effect size of depression (Cohen's d)	0.38	-0.72*	0.88
Correlation with key measures	r (95%CI)	r (95%CI)	r (95%CI)
R-AS	0.68 (0.54, 0.78)		
RMA-13	0.82 (0.74, 0.88)	0.68 (0.55, 0.78)	
MHI	0.35 (0.14, 0.53)	0.46 (0.27, 0.61)	0.53 (0.35, 0.66)
MOCA	0.21 (-0.03, 0.41)	0.03 (-0.18, 0.24)	0.06 (-0.15, 0.27)
2 min walk test	0.20 (-0.01, 0.40)	0.20 (-0.02, 0.40)	0.26 (0.04, 0.45)
SIS participation	0.35 (0.14, 0.52)	0.45 (0.26, 0.61)	0.46 (0.28, 0.62)

AS – original Apathy Scale, R-AS – Rasch Version of AS, RMA- Rasch Measure of Apathy;[1] higher score is better; [2] Cut-point for apathy range <29 in the AS,R-AS, and RMA; * McNemars' chi-square p<0.001; ** Cut-point for depression <60 in the MHI RAND-36; *** p<0.05 for between subject, within subject and interaction in the repeated measure ANOVA;, MHI – Mental health index measure; MOCA – Montreal cognitive assessment; SIS – Stroke impact scale.

Figure 5.1: The item-person threshold distribution and test information function for RMA

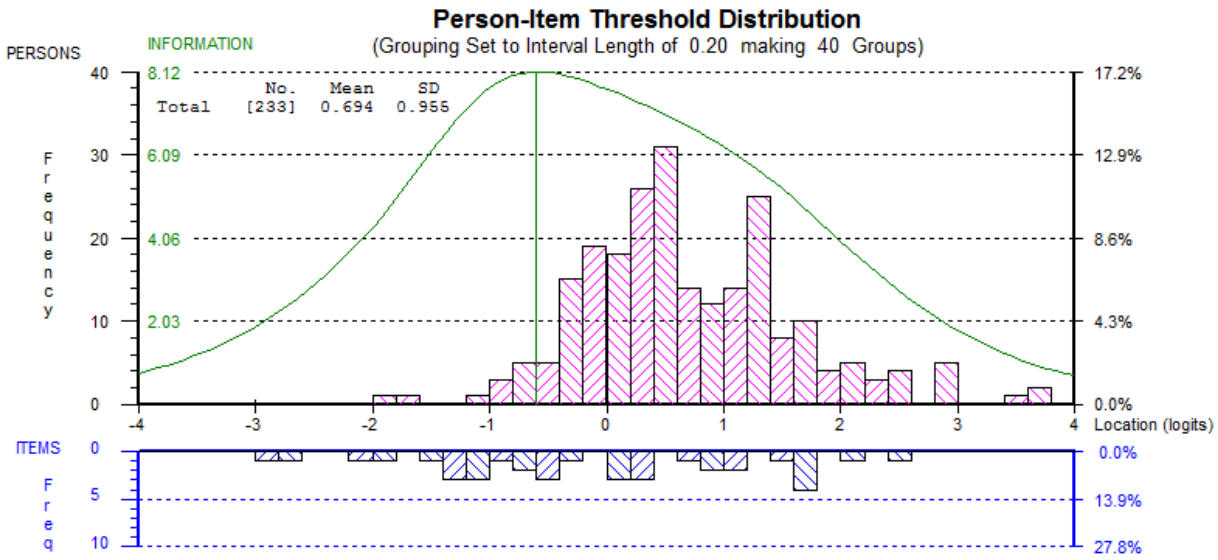


Figure 5.2: Bland-Altman plot of the difference between R-AS and RMA estimations

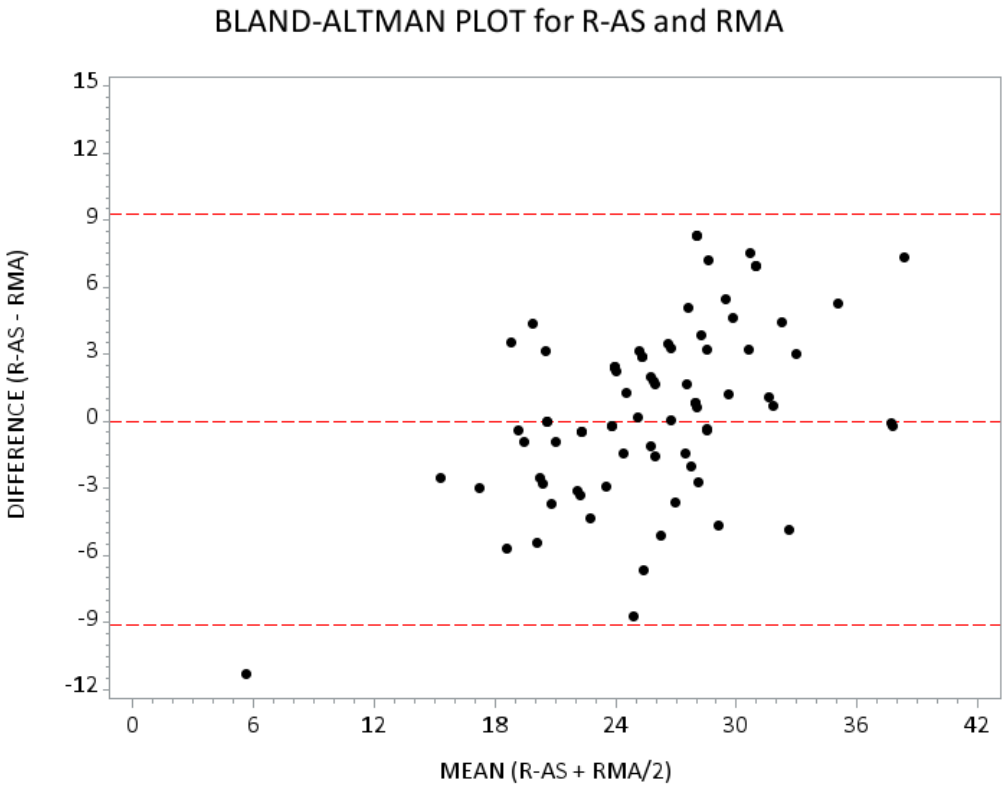


Figure 5.3: Scatterplot and correlation for Original AS and Rasch Measure of Apathy (RMA)

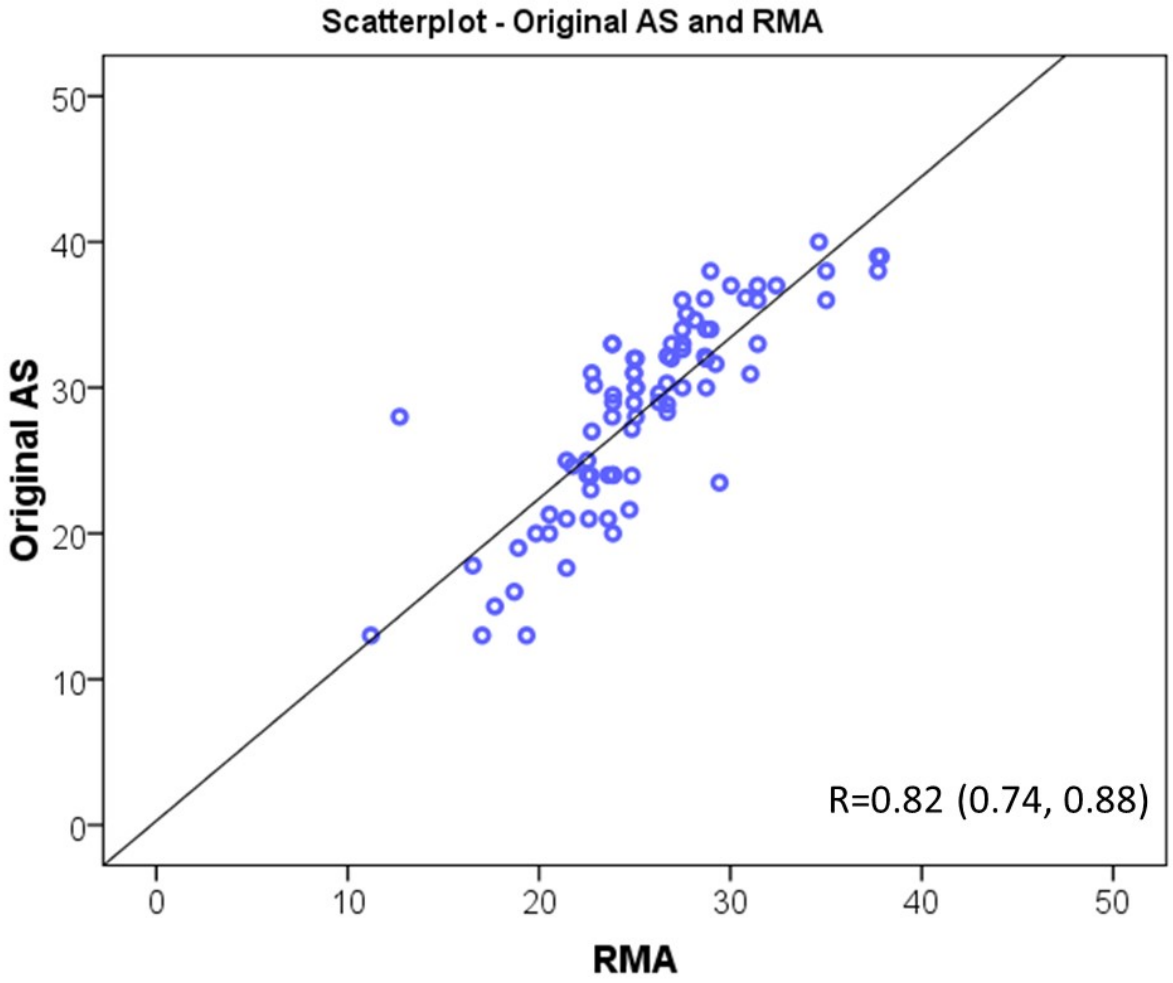
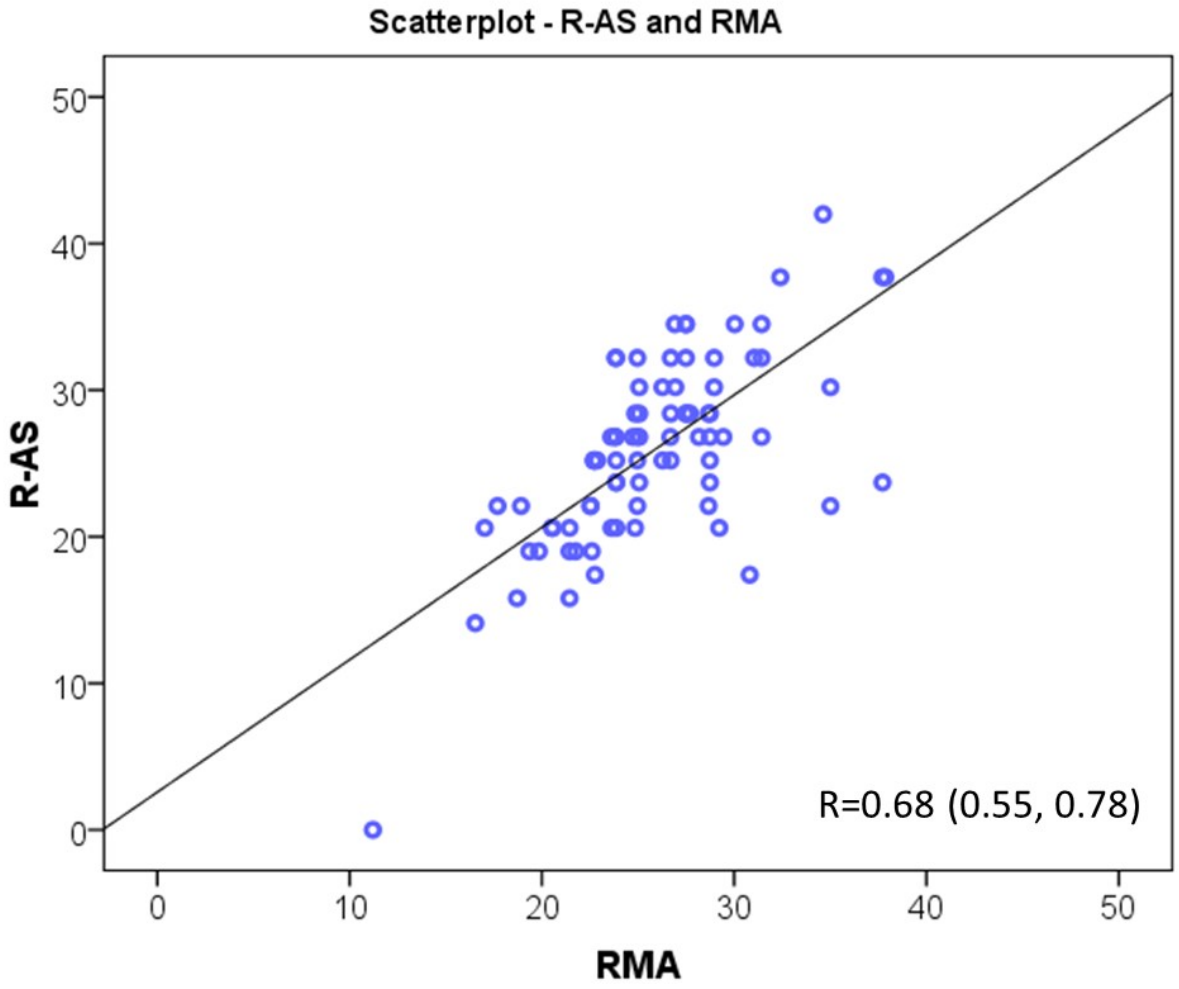


Figure 5.4: Scatterplot and correlation for Rasch version of AS (R-AS) and Rasch Measure of Apathy (RMA)



CHAPTER 6 -MANUSCRIPT4–Identifying and Characterizing Trajectories of Apathy in Stroke and Impact on Functional Recovery

PREFACE TO MANUSCRIPT 4

At this stage, two versions of the AS original measure were proposed, one for the motivation continuum comprising 10 of the original 14 AS items (with rescoring); and one for the apathy-motivation continuum adding 3 additional items to reflect the emotional component necessary for apathy. The links between the measures under study, the original AS, the R-AS, and RMA, to theoretical concept as derived in the literature is outlined in Table 6.0. It would be ideal to administer the R-AS and the RMA on a different sample of stroke participants to test the measures and to investigate further the relationship of the two constructs. However, this was not viable in the context of the doctoral thesis.

The availability of longitudinal data allowed to “pre-test” longitudinal validity by providing evidence on how these two measures behave over time in this stroke sample and if the two measures differ in the impact of functional recovery after stroke. There are a number of statistical methods available for the analysis of correlated data. The repeated measures structure of longitudinal data essentially creates multilevel data in which one level reflects the multiple data points per person (termed person-level) and the other

reflects the group, i.e. the person. The choice of statistical model depends on whether the dependency in the data is a nuisance or is a phenomenon of interest [159].

In the context of this study, the repeated measure is time and it is the phenomenon of interest. The two principal methods for analysing longitudinal data are mixed models and latent class models. The terminology used to describe these longitudinal statistical procedures is inconsistent. Mixed models have many terms that are used fairly synonymously: random effect models, hierarchical models, multilevel models and growth curves (the other terms are synonymous; growth curves are a subset of the others). Latent class models include: mixture models, latent curve analysis, trajectory analysis, and group based trajectory analysis (GBTA).

The following manuscript used a form of the latent class model, GBTA, because it assumes that the population under study is made up of a mix of people with different developmental trajectories rather than assuming an average developmental trajectory with individual variability around the mean [160] as is the assumption underlying growth curve analysis.

This GBTA model has been a more commonly used in social sciences and has only recently been applied in health field[131, 161-166]. Xie et al. [162] contrasted GBTA and growth curve analysis, in a sample of people assessed for cognitive decline over time, and showed that GBTA identified 5 different trajectories of decline. None of these trajectories were similar to the average decline identified from the growth curve

analysis. Thus GBTA provided a more accurate representation of longitudinal change for this sample of people.

Another advantage of GBTA is that both the outcome variable (dependent variable) and exposure variable (independent variable) can be longitudinal and their joint change can be related, this is called joint trajectory analysis. Mayo et al. [145], investigated necessary and sufficient causes of participation in stroke using this approach to infer causal association between walking status, mood, social support, and participation following stroke in the same data set used in the current study. Therefore Manuscript 4 applies the same methodology to relate change in motivation/apathy over time to change on participation over time. The results from GBTA are reported as trajectories and the term is used in its statistical sense rather than in a descriptive sense.

Thus, the objective of this study was to contribute to the understanding of apathy and motivation continuum in the first year of stroke. The specific objectives are (i) to compare the longitudinal behaviour of stroke patients measured by the R-AS and the RMA; (ii) to estimate the extent to which impaired motivation and apathy syndrome impact functional recovery after stroke. We hypothesized that apathy and motivation will vary both cross-sectional and longitudinal; and that the apathy syndrome will have a stronger relationship with poor participation.

Table 6.0: Updated roadmap to date

	Number of items per domain						
	Openness to experience	Energy and Drive Function		Emotional Function	Motivation and Planning	Other	Total
		Energy Level	Motivation				
Theory	✓	✓	✓	✓	✓	✖	
AS	4	4		1	2	3	14
R-AS	4	4		0	2	0	10
RMA	4	4		3	2	0	13

Theory – Apathy construct conceptual theory; AS – original Apathy Scale; R-AS – Rasch Version of the AS; RMA- Rasch Measure of Apathy; *Not a component of apathy construct.

TITLE PAGE

Identifying and Characterizing Trajectories of Apathy in Stroke: Impact on Functional Recovery

For submission to: Stroke

Authors: Christiane B. Lourenço, M.Sc.,[1] Nancy E. Mayo, Ph.D.[1,2,]

[1] School of Rehabilitation Sciences, Faculty of Medicine, McGill University, 3654 Prom Sir William Osler, Montreal, Quebec, H3G 1Y5, Canada.

[2] Division of Clinical Epidemiology, Royal Victoria Hospital, McGill University Health Center, Montréal, Québec.

Communication address to: Christiane Lourenco, Royal Victoria Hospital, Division of clinical epidemiology, 687 Pine Avenue West, Ross 4.27, Montreal, Quebec, H3A 1A1 Canada. Telephone: (514) 934-1934 ext 37066, Fax: (514) 843-1493, Email: christiane.lourenco@mail.mcgill.ca

Key Words: apathy, motivation, rehabilitation, recovery, trajectory

ABSTRACT

Objectives: To identify the extent to which longitudinal change in symptoms of apathy differ according to the measure and to estimate the extent to which location along the apathy-motivation continuum impacts functional recovery after stroke.

Methods: The data for this study comes from a prospective study of depression post-stroke. Group-based trajectory modelling (GBTM) was used to identify distinctive groups of individuals with similar trajectories. Dual trajectories were used to estimate concordance between trajectories of apathy-motivation continuum and participation.

Results: Group-based trajectory analysis revealed five trajectories of motivation and four trajectories if symptoms of apathy were measured. The majority of participants had scores in the mid-range of the apathy-motivation continuum which remained stable over time. However, even this moderate level of apathy had an important impact on participation.

Conclusions: The relative stability of the motivation-apathy continuum could indicate that this is resistant to change or that it hasn't been adequately addressed in rehabilitation. Clearly, better measurement of this important construct is needed. The R-AS, which measures only the motivation continuum, may be the most feasible clinical tool for rehabilitation purposes.

INTRODUCTION

The current theory underlying apathy is that it impacts on function [21, 68, 70, 129-131] and recovery in neurological conditions [68]. However, a systematic review and meta-analysis investigating the effect of apathy on the severity of clinical outcomes (e.g. Functional Independence Measure, Barthel, and Modified Rank Scale) in people with stroke did not support a difference in function between apathetic and non-apathetic people [167]. The heterogeneity of measures of apathy and the nature of the clinical outcomes limited the strength of the conclusions [167].

There is also limited evidence on the effectiveness of treatment for apathy. The systematic review in Manuscript 1 identified only 9 studies that investigated the effect of pharmacological and non-pharmacological interventions for apathy. Although the heterogeneity and poor methodological qualities of the studies did not allow for strong conclusions, there is evidence that apathy might be a modifiable construct if appropriate interventions (e.g. dopaminergic therapy and cognitive behaviour therapy) are implemented.

Mayo et al.[131] was the first study to estimate the behaviour of apathy over the first year of stroke and showed that 50% of the stroke participants (n=402) had some level of apathy, ranging from minor to high, and groups with higher apathy demonstrated poorer recovery in domains related to activity and participation. In addition, only 14% of the sample showed signs of change, either improving or decreasing; the remaining subjects

showed stable levels of apathy across time. However, the measure of apathy used in this study had not been validated as an independent measure of apathy; it was derived from an index of behavioural manifestation as observed by caregivers.

The lack of gold standard measure is currently limiting the study of apathy in stroke. During the course of this thesis, the complexity of the apathy construct and its inherent measurement challenges were demonstrated. Manuscript 2 and 3 provided two new measures, with stronger psychometric and conceptual properties. It would be of interest to see how these two measures behave over time and if they are more suitable to detect change and impact on apathy after stroke than what was currently available.

Therefore, the specific objectives are to identify the extent to which longitudinal trajectories of symptoms of apathy-motivation continuum differs according to the measure used and to estimate the extent to which location along the apathy-motivation continuum impacts functional recovery after stroke. We hypothesized that motivation will vary both cross-sectionally and longitudinally; and that location along the apathy-motivation continuum will have a strong relationship with poor participation

METHODS

Source of Data

The data for this study comes from a prospective study on depression post-stroke. The methods of this study have been previously described [145]. Briefly, all people admitted to

McGill University Health Centre adult sites with acute stroke were assessed for eligibility. Participants were evaluated at 10 days post-stroke and subsequently at 3, 6, 9, and 12 months post-stroke by trained health care professionals who assessed subjects' level of activity and participation, as well as administered questionnaires on mood and emotional function.

Measures

The Rasch version of AS (R-AS) was developed in Manuscript 2, and comprises 10 of the original AS items but rescored into three multi-item sub-tests in order that the data fit the Rasch Model. The subtests reflect three domains that fit with a motivation construct: openness to experience (4 items); energy and drive function (5 items); and organization and planning (2 items). As data fit the underlying Rasch model, a meaningful total score can be derived by summing the score on each of the three super-items. To fit with the original AS scoring, the R-AS was transformed to range from 0-42, with high score indicating greater motivation. The Rasch measurement psychometric criteria were supported with person separation index 0.70 and adequate scale-to-sample targeting of the data to the Rasch model (X^2=6.4, DF=12, p=0.73).

A new measure to span the apathy-motivation continuum named the Rasch Measure of Apathy (RMA) was developed in Manuscript 3. The RMA builds from the R-AS by adding an important fourth dimension for emotional function. The RMA is comprises 13 items reflecting all four domains that fit with the apathy construct: openness to

experience (4 items); energy and drive function (5 items); organization and planning (2 items); and emotional function (3 items). As data fit the underlying Rasch model, a meaningful total score can be derived by summing the score on each of these three super-items. To fit with the original AS scoring, the RMA was transformed to range from 0-42, with high score indicating lower apathy. The Rasch measurement psychometric criteria were supported with person separation index 0.68 and adequate scale-to-sample targeting of the data to the Rasch model (X^2=27.8, DF=18, p=0.06).

Participation was measured using the participation subscale of the Stroke Impact Scale (SIS) [147], a 8-item scale that assesses the ability to participate in activities related to work, recreation, religion, family and society. Each item is measured on a 5-point Likert scale; ranging from 0-100. Higher scores indicate a greater degree of participation. Reliability estimates and content and construct validity have been demonstrated [147]. A 10-15 change score in a domain represents a clinically significant change. The SIS has been Rasch analysed by Duncan et al., that showed that the scale contain domains which are unidimensional, items that have excellent range difficulty, and a domain of scores that can differentiate patients into different strata [124]. Therefore, it is seen as a gold standard to monitor recovery of physical function in stroke survivor since the scale is supported by very strong psychometric properties.

Statistical Analysis

Group-based trajectory modeling [127] was used to identify unique trajectories of motivation/apathy over the first year of stroke as measured by three different measures: the AS, the R-AS, and the RMA. Trajectories were estimated using the TRAJ CNORM procedure, a semi parametric group-based modeling strategy appropriate for data that are approximately normally distributed with or without censoring, with SAS version 9.3.

Model selection was based on the iterative estimation of (1) the number of trajectory groups and (2) the shape/order of each trajectory group using both statistical and non-statistical considerations. Fit statistics (Akaike's information criterion (AIC), the Bayesian information criteria (BIC), and sample size adjusted BIC (ssBIC) and posterior probabilities of group membership were compared across models. The choice of the best-fitting model took into consideration the following criteria: having AIC and BIC values closest to 0; high mean posterior probabilities specific to each group; and similar theoretical and assigned proportions. To relate longitudinal change in motivation to longitudinal change in participation, joint group-based trajectory modelling [168] was used.

RESULTS

Eighty-two individuals were initially analysed in the present study, however, one subject with extreme values was negatively interfering on the trajectories model and was excluded from the analysis. Of the 81 participants, there were 52 at time 1 (3 months post stroke), 57 at time 2 (6 months post stroke), 54 at time 3 (9 months post stroke)

and 68 at time 4 (12 months post stroke). The mean and standard deviation for all the outcomes, at four time points, are displayed in Table 6.1. All three apathy/motivation measures range from 0-42, and SIS participation from 0-100. Higher scores indicate more motivated. A score of 100 on the SIS-participation indicate excellent participation. The average score for each outcome measure, across the four time points, suggested that the latent trait was not changing over time, except for participation that showed almost 10 point increase on the SIS-participation from 3 months to 1 year.

Table 6.2 shows the values of fit indices (AIC, BIC, and ssBIC) used to define the best number of trajectory for each measure. The best trajectory model fit for each outcome is highlighted in bold. A 3-group model best represents the AS, a 5-group model the R-AS, and a 4-group model the RMA and SIS-participation. Figures 6.1 to 6.4 illustrate the group-based trajectory for each outcome investigated, original AS, R-AS, and RMA. The scale score is represented in the y-axis and the 4 time points on the x-axis. Each trajectory line represents one distinct group of participants score over time. The proportion of participants assigned to each group is given on the bottom of the graph. For example, Figure 6.1 illustrates participants' level of apathy measured by the original AS over time. The original scale scoring ranges from 0-42, but participants' estimates in this sample ranged from 19 to less than 36. Trajectory number one (in red) represents the apathy behaviour of 19.9% over 4 time points; trajectory number 2 represents 36.9% of the sample and trajectory number 3, 43.1%.

Trajectory parameters are given in Table 6.3. The largest group for the AS is the "Motivated" group (35.0%) with an average score of 33.9 (out of 42) at 3 months that remained stable, with no change over time. The two other groups "Apathetic" and "Moderate Apathy", showed an average score of 20.7 and 28.2 respectively that also remained stable over time.

The largest group of the 5-group model R-AS is the "Moderate Motivation" group (41.5%) with an average score of 29.3 (± 5.6) out of 42 at 3 months that also remained stable over time. The lack of change was observed for two other groups, the "Moderate Apathy" and the "Apathetic", with R-AS scores at 3 months of 19.0 and 24.5 respectively. Nevertheless, the R-AS was able to discriminate two additional groups, the "Improving" and the "Highly Motivated". The "Improving" groups was formed by 5% of the sample with an average R-AS score of 21.0 at three months that increased 5.3 points per time point, reaching a total score of 36.5 at 12 months. Another 4% of the sample was highly motivated at 3 months (R-AS of 41.7) but decreased the motivation over the next 6 months (R-AS 36) and then improved again at 12 months (R-AS 39.5) (see Appendix A15).

Four group trajectories were identified for the Rasch Measure of Apathy (RMA). The largest group is the "Moderate Apathy" group (46.3%) with an average score of 25.8 (± 0.5) at baseline that remained stable, with no change over time. Only one group showed improvement over time, the "Highly motivated and improving" group (6.8%) which had

an average score of 32.3 (±14.8)at 3 months and increased a total of 6.2 points at end of the first year.

Four trajectories best fit participation over time. The majority of participants showed "fair" participation (39.9%) with mean score of 51.2 (±4.3) that increased 4.7 points per time point. Only 6.6% of the participants reached "excellent" participation, starting from a score of 89.2 on the SIS-participation at 3 months and reached 100 at 12 months (6.0 points per time point). The "poor" participation group (16.3%) had a mean score of 32.9 that remained constant over time. The "very good" group showed small improvement (3 points) per time point, starting at 79.2 and reaching 87.3 at 12 months.

The prevalence of motivation/apathy groups trajectories conditional on the trajectories of participation are shown in Table 6.4. Across at the top of the table are the proportions of each participation trajectory. The second column represents the proportions within the trajectories of each apathy/motivation outcome. Within the body of the table are the conditional probabilities: the distribution across trajectories of subjects post stroke measured by one of the three measures conditional on participation.

The results displayed in Table 6.4 indicate that a stroke patient must be in top end of the apathy-motivation continuum (RMA≥ 35) in order to achieve "excellent" participation; a "moderate" level of motivation can only achieve "very good" participation. On the other hand, few subjects with only "moderate" motivation (R-AS 28 to 31) were able to achieve excellent participation. Overall, subjects with score ≤29 on all three measures

were assigned to apathetic or moderate apathy groups and the majority of these subjects only showed only fair participation, especially subjects with scores ≤20.

Table 6.5 shows how subjects assigned to trajectories by the AS fared when assigned to trajectories based on the other two measures. There were three groups in the original AS: “Apathetic” (n=16), “Moderate Apathy” (n=30), and “Motivated” (n=35). The results displayed in the table suggest that RMA classified subjects as more severe apathy symptoms than the AS. From the 30 subjects considered “Moderate Apathy” in the AS, 6 were assigned as “Apathetic”, and 24 continued as “Moderate Apathy”. From the 35 subjects classified as “Motivated”, only 5 remained in the highly motivated category and 12 classified as moderately motivated.

On the other hand, of the 30 moderate apathy form the AS, 23 remained classified as moderately apathetic on the R-AS. Of the 35 motivated subjects according to AS, only 5 were so classified on the R-AS.

DISCUSSION

This study showed that the diversity and shape of longitudinal change in the apathy-motivation continuum depends on how the construct was measured. The original AS had the fewest trajectories, and they were all flat, which is not surprising given that it had a number of psychometric limitations. The types of limitations introduce misclassification which tends to reduce variability in the measured construct. The R-AS,

tapping only the motivational continuum, had the most trajectories (n=5), the majority of which were flat (90% of subjects did not change over time). However, a small group of subjects who improved were identified. The RMA, which included an additional component of emotional function, showed three trajectories largely flat, and one improving.

Mayo et al.[131] found in a larger clinical sample, using an entirely different approach based on caregiver assessment of patients behaviour, that 86% of the sample had flat trajectories and only 6% of the sample improved; 7% deteriorated. There has been no other longitudinal study on apathy in stroke.

The distinction between the three measures which is shown in Table 6.5 confirms that using better measures can help discriminate at the group level, as above, but also at the individual level. At the individual level, the AS identified 16 of the 82 people as being apathetic, 30 as having moderate apathy, and 35 as being motivated. The other measures were able to discriminate another group of highly motivated subjects that were not captured by the AS. Of those people classified in the AS as being moderately apathetic, the vast majority of them were reclassified in the other measures to either moderate motivation of or even apathetic. This misclassification would have an impact as to who would be referred for further diagnosis and who would require a different rehabilitation approach.

The link between apathy and participation was also demonstrated. When apathy was measured by the AS, being motivated was necessary for excellent participation because nobody with excellent participation had moderate or high apathy. However, it was not sufficient, because, motivated subjects had participation that ranged from poor to excellent, see table 6.4

For R-AS, tapping the motivation continuum, having high motivation was sufficient for having very good participation or higher. For the RMA, which included emotional function items, the motivated and improving group was highly associated with very good and excellent participation, but subjects in the apathy range only showed poor and fair participation. Overall, to reach very good and excellent participation, motivation was required; conversely scoring in the apathetic range produced poor and fair participation. Similarly, Mayo et al.[131] found a linear dose response relationship between higher motivation and better participation. Motivation is required for people with stroke to recover beyond basic capacity to function to regain meaningful roles in life (participation).

It was also interesting to note that subjects who scored <29 did not reach excellent participation, in fact less than 50% reached even very good participation. In Manuscript 2 and 3, we used 29 out of 42 as the cut-off to detect apathy. This score was based on Starkstein's original AS cut-off score, and it yielded a higher proportion of apathy with the revised measures (73% R-AS; 82% RMA) compared with 36% reported in the

literature using existing measures [167]. These findings seem to contribute to the validity of the cut-point, as this classification was able to predict poor functional recovery.

CONCLUSIONS

The relative stability of the motivation-apathy continuum could indicate that this is resistant to change or that it hasn't been adequately addressed in rehabilitation. Clearly, better measurement of this important construct is needed. Given that the R-AS is a shortened and improved version of the original AS, it may be a feasible clinical tool to be used for measuring the level of motivation for rehabilitation purposes.

Table 6.1: Mean and SD of participants in the four outcomes under investigation at 4 time points

	Total N= 81			
	3 months	6 months	9 months	12 months
	N= 52	N= 57	N= 54	N= 68
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Apathy Scale (0-42)	29.2 (6.2)	29.3 (5.8)	29.7 (5.9)	29.3 (6.0)
Motivation (0-42)	26.2 (5.5)	26.7 (5.5)	25.8 (5.4)	26.9 (5.6)
Rasch Measure of Apathy (0-42)	25.8 (4.3)	25.8 (4.8)	26.7 (4.4)	26.4 (5.4)
SIS participation (0-100)	63.2 (23.4)	67.1 (23.1)	68.0 (24.3)	72.2 (23.0)

Table 6.2: Fit indices for two to six groups' growth models for original Apathy Scale, Motivation, Rasch Measure of Apathy, and Participation.

	Growth mixture model				
	2 groups	3 groups	4 groups	5 groups	6 groups
AS					
AIC	-706.9	**-695.7**	-695.2	-697.1	-700.2
BIC	-713.8	**-706.0**	-709.0	-714.3	-722.6
ssBIC	-711.7	**-702.9**	-704.8	-709.1	-715.8
R-AS					
AIC	-649.3	-640.1	-631.5	**-628.4**	-637.1
BIC	-666.1	-665.3	-651.7	**-650.2**	-679.1
ssBIC	-660.4	-656.8	-644.8	**-642.8**	-664.8
RMA					
AIC	-611.4	-596.8	**-589.6**	-590.1	-598.0
BIC	-628.2	-621.9	**-608.5**	-620.4	-638.1
ssBIC	-622.5	-613.5	**-601.8**	-610.1	-632.3
SIS-Participation					
BIC	-1440.1	-1423.4	**-1414.7**	-1432.0	-1431.5
ssBIC	-1436.0	-1416.3	**-1408.2**	-1421.1	-1418.4
AIC	-1426.8	-1400.5	**-1393.7**	-1396.3	-1389.5

AS – Apathy Scale; R-AS – Rasch Version of AS; RMA – Rasch Measure of Apathy; AIC– Akaike's Information Criterion; BIC – Bayesian Information Criterion; ssBIC –Sample size Bayesian Information Criterion

Table 6.3: Summary of group-based trajectory analysis for original Apathy Scale, Motivation, Rasch Measure of Apathy, and Participation.

	Group Membership N (%)	Intercept β (SE)	Slope β (SE)	Slope2 β (SE)	Slope3 β (SE)	Posterior Probability Mean (SD)
AS						
Apathetic	16 (19.8%)	20.7 (0.7)	-	-		0.90 (0.13)
Moderate Apathy	30 (37.0%)	28.2 (0.6)	-	-		0.84 (0.16)
Motivated	35 (43.2%)	33.9 (0.5)	-	-		0.90 (0.13)
R-AS						
Apathetic	13 (16.0%)	19.0 (1.0)	-	-		0.81 (0.19)
Moderate Apathy	26 (32.1%)	24.5 (1.0)	-	-		0.79 (0.18)
Improving	4 (5%)	21.0 (3.1)	5.3 (0.4)	-		0.74 (0.25)
Moderate Motivation	35 (43.2%)	29.3 (0.6)	-	-		0.87 (0.16)
Highly motivated	3 (3.7%)	58.0 (12.1)	-4.2 (3.1)	-1.5 (0.2)	3.5 (0.4)	0.93 (0.12)
RMA						
Apathetic	22 (27.2%)	21.4 (0.6)	-	-		0.86 (0.16)
Moderate Apathy	42 (51.8%)	25.8 (0.5)	-	-		0.79 (0.16)
Moderate Motivation	12 (14.8%)	29.6 (0.7)	-	-		0.87 (0.15)
Highly motivated and improving	5 (6.2%)	32.3 (14.8)	2.9 (0.2)	-0.6 (0.3)	3.9 (0.6)	0.99 (0.01)
SIS-Participation						
Poor	12(14.8%)	32.9 (4.8)	-	-		0.90 (0.16)
Fair	33 (40.7%)	51.2 (4.3)	4.7 (0.5)	-		0.88 (0.13)
Very Good	31.0 (38.3%)	79.2 (4.0)	2.8(0.5)	-		0.89 (0.12)
Excellent	5 (6.2%)	89.2 (14.0)	8.5 (3.3)	2.1 (0.3)		0.93 (0.08)

Table 6.4: Prevalence (%) of concordance between AS, R-AS, and RMA, and Participation Trajectories

		PARTICIPATION			
		Poor	Fair	Very Good	Excellent
APATHY SCALE(high is worse)	**Group Membership**	**1: 12.3%**	**2: 37.6%**	**3: 33.6%**	**4: 16.5%**
Apathetic	**1: 19.73%**	11.5	74.6	13.9	0
Moderate apathy	**2: 38.4%**	21.8	31.9	46.3	0
Motivated	**3: 33.9%**	4.0	25.3	31.3	39.4
MOTIVATION (high is better)	**Group Membership**	**1: 14.0%**	**2: 38.7%**	**3: 34.8%**	**4: 12.5%**
Apathetic	**1: 12.5%**	0	100	0	0
Moderate Apathy	**2: 32.7%**	29.9	30.0	40.0	0
Improving	**3: 6.0%**	0	23.5	76.5	0
Moderate Motivation	**4: 44.7%**	9.4	33.4	35.1	21.9
Highly motivated	**5: 4.0%**	0	0	34.5	65.4
Apathy (RMA**) (high is worse)	**Group Membership**	**1: 18.8%**	**2: 39.6%**	**3: 35.3%**	**4: 6.2%**
Apathetic	**1: 26.8%**	38.7	61.3	0	0
Moderate apathy	**2: 49.7%**	17.0	38.6	41.8	2.5
Moderate Motivation	**3: 17.0%**	0	23.5	76.5	0
Highly motivated and improving	**4: 6.5%**	0	0	22.7	77.3

*The labelling of the trajectories does not follow the scoring, but rather the interpretation of the score with respect to the construct under study, apathy or motivation. When apathy is the construct, high refers to high apathy. When motivation is the construct, high refers to high motivation; **Rasch Measure of Apathy

Table 6.5: Concordance of participants assigned to each trajectory groups of the R-AS, AS, and RMA.

R-AS	Apathetic 13	Moderate Apathy 3 / 23	Improving 4	Moderate Motivation 6 / 29	Highly Motivated 3
AS	**Apathetic** **16**	**Moderate Apathy** **30**			**Motivated** **35**
RMA	Apathetic 16 / 6	Moderate Apathy 24 / 18	Moderate Motivation 12		Highly motivated and Improving 5

Figure 6.1: Three group trajectory model for Apathy Scale over the first year of stroke

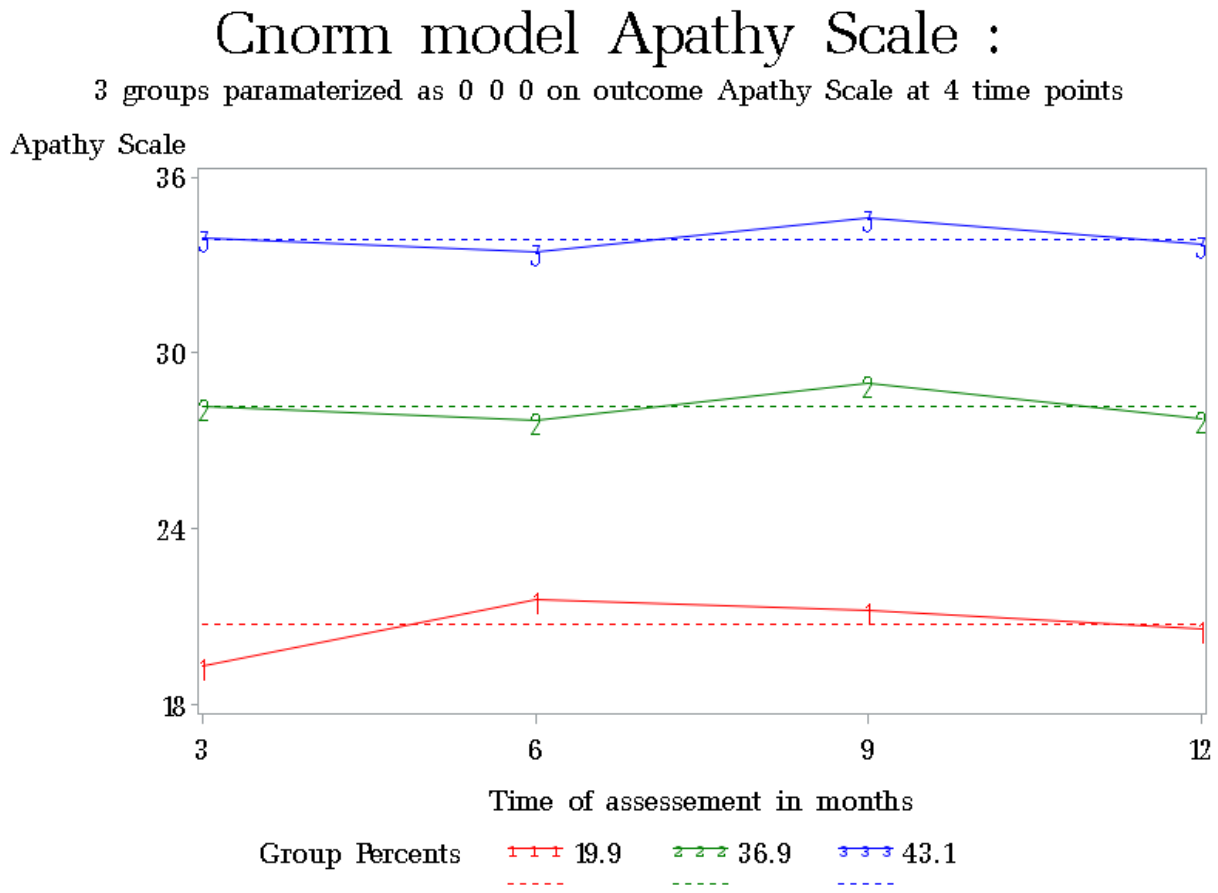


Figure 6.2: Five group trajectory model for Motivation over the first year of stroke

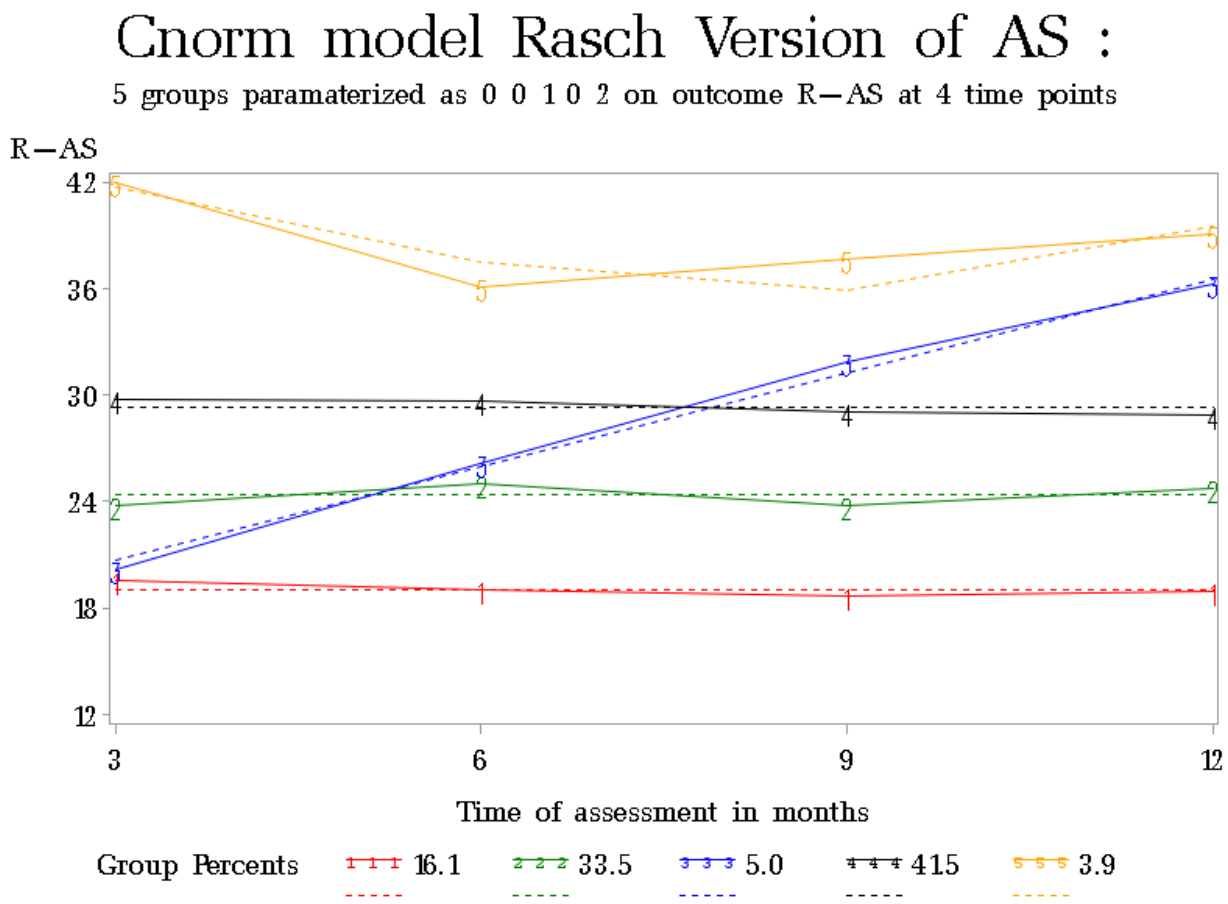


Figure 6.3: Four group trajectory model for Apathy measured by the Rasch Measure of Apathy over the first year of stroke

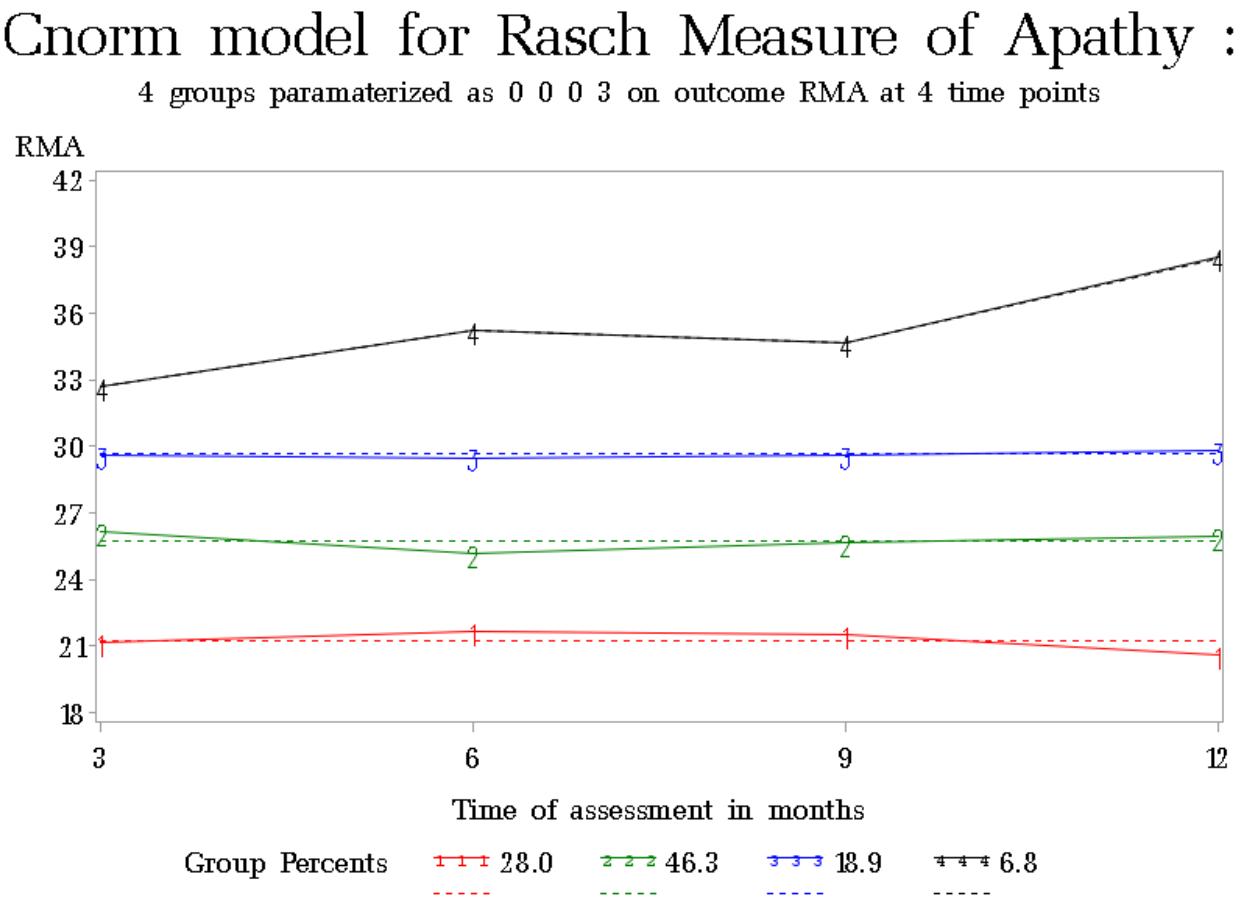
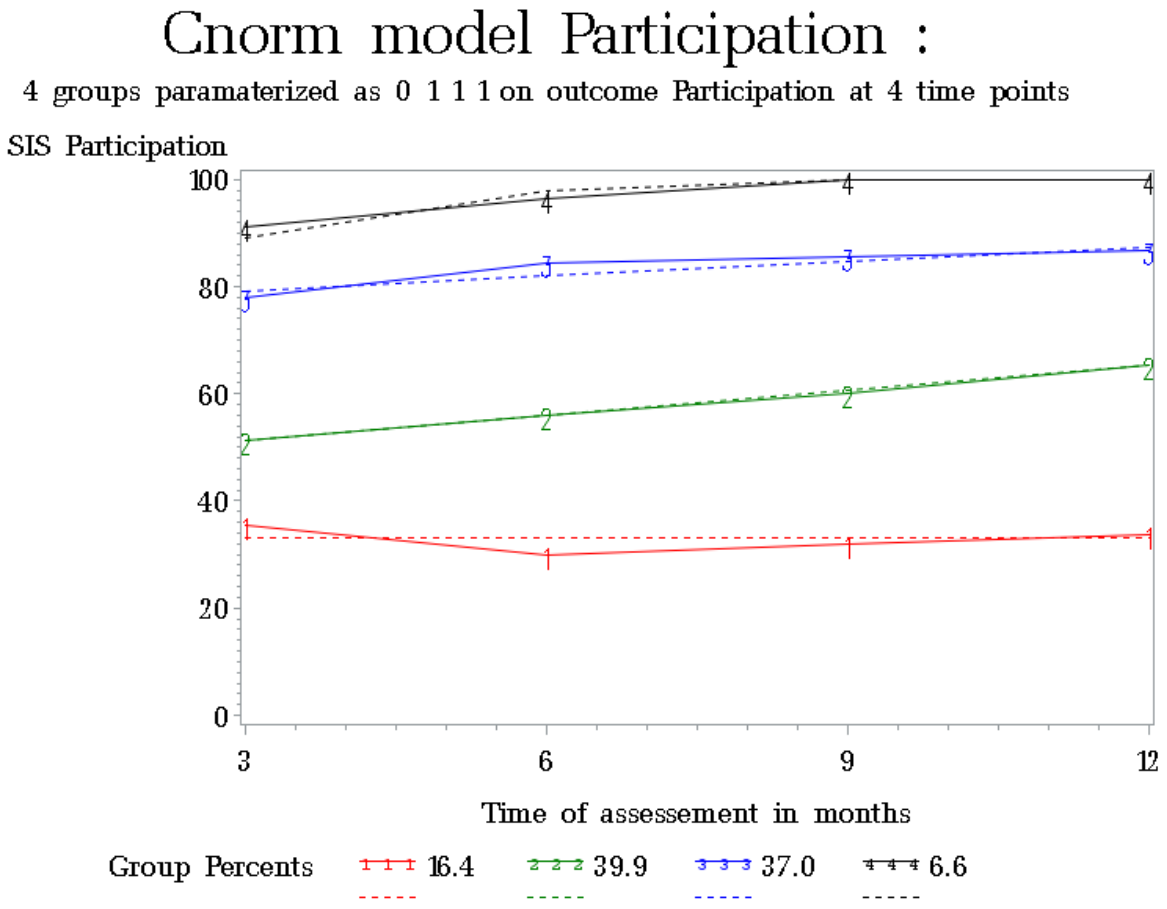


Figure 6.4: Four group trajectory model for Participation over the first year of stroke



CHAPTER 7 – Relationship Between Brain Lesion Location and the Apathy Motivation Continuum

At this stage, two versions of the AS original measure were proposed, one for the motivation continuum comprising 10 of the original 14 AS items (with rescoring); and one for the apathy-motivation continuum adding 3 additional items to reflect the emotional component necessary for apathy. Because stroke produces neurological damage which affects all aspects of function and disrupts neurobiological structures which are implicated in motivation and apathy, additional evidence for the conceptualization of the apathy-motivation continuum would be provided if a link to the brain can be identified.

As outlined in the Literature review in Chapter 1, the limbic system and basal ganglia are known to regulate motivation while the control of real-life decision making is a function of the prefrontal cortex, in which the left prefrontal cortex is associated with positive goal context and the right prefrontal cortex with withdrawal behaviour [63]. The literature provides consistent evidence that the apathy syndrome is associated with those areas regulating motivation and decision making, with predominant involvement of the frontal lobe [37, 38, 60] and basal ganglia [36, 41, 65]. Other structures have also been associated, such as internal capsule [30] and the temporal lobe [37].

It would be of value to the field of stroke and stroke recovery to identify whether lesions in specific locations predispose an individual to have apathy or low motivation. The studies to date are not conclusive owing to methodological differences.

The methods available to study the relationship between brain lesion location and function have changed significantly over the past few years [169]. The lesion method using computerized tomography (CT) and magnetic resonance imaging (MRI) enabled to identify regions that are commonly damaged in different individuals with the same deficit, by overlapping individual lesions in a brain template and reporting most commonly damaged, without the use objective statistics [170, 171]. With the advancement of technology, a new method denominated the activation method was created and includes techniques such as single-photon emission computed tomography (SPECT), position emission tomography (PET), and functional magnetic resonance imaging (fMRI). They allowed detecting task-related brain regions based on blood-oxygenation levels and have refined the classical lesion method.

The lesion method and activation method differ in a number of aspects and both have important limitations, which make them complementary methods rather than competitive techniques [169, 172]. The use of convergent methods to investigate the association of a lesion location and function only increases the power of cognitive neurosciences [170]. Therefore, a third method which is based on statistical lesion analysis has been proposed to strengthen even more lesion location studies. There are three statistical methods that analyse on a voxel-by-voxel basis: voxel-based morphometric (VBM),

voxel-based lesion symptom method (VLSM), and voxel-based analysis of lesions (VAL). They provide highly spatial precision and use the same spatial coordinates of fMRI which allows comparisons between studies using fMRI and PET [170].

In the VAL approach, the entire brain is mapped as a volume of small 3D 'voxels' (each voxel ranging from $1mm^3$ to $27mm^3$), and then an independent statistical test is conducted for each voxel (i.e. for each voxel, one computes a group of subjects who show disorder of interest and a group of control patients with brain damage who do not show the disorder). One of the advantage of this method is that it offers better spatial precision as the entire brain is mapped as a volume of small 3D "voxels" and then the subject's performance is compared to voxels located in areas with lesion and without lesion, which allows for prediction of poor performance when a brain region is injured and of good performance when it is not. Another advantage of this approach it that it uses all available information and eliminates reliance on cut off scores, clinical diagnoses, or special regions of interest, allowing for additional areas to emerge in the exploration of networks that support a given behaviour (e.g. apathy) [172]. Therefore, this approach is considered a bridge between classic lesion method and modern functional imaging [173].

The literature review in Chapter 1 identified 8 studies that have investigated the association of lesion location and apathy syndrome after stroke. As can be seen in Table 1.2, the imaging methods differed considerably among studies, including both lesion methods and activations methods (CT, MRI, and SPECT), but there were no

studies using the statistical lesion method such as the VAL, which might explain why there is still no strong conclusions on which region may be involved with apathy syndrome.

Therefore, the purpose of this study was to identify the extent to which location along the apathy-motivation continuum as defined using Rasch Analysis was associated with a specific lesion location using a VAL approach. The specific objectives were to identify communality in lesion locations among those subjects defined within the apathetic range of apathy-motivation continuum. Based on the literature review, the hypothesis was that lesions located on the frontal/striatal regions of the brain would be associated with apathy symptoms. Evidence from this exploratory study would provide preliminary data to support the development of a research protocol for a comprehensive evaluation of this relationship.

METHODS

Participants

A secondary data analysis was conducted from a prospective longitudinal study on stroke depression. Subject and procedures have been described thoroughly in Manuscript 2. For the current analysis, only the subset of patients who had a MRI (n=57).

Image processing

MRI data were obtained for 57 patients with a 1.5T GE Signa Excite MR scanner (General Electric, Milwaukee) using a standard 8-channel head coil. Using the stroke imaging protocol, the following sequences were acquired: 3-D, T1-weighted spoiled gradient recalled (SPGR) echo sequence of 124 contiguous, 1.0 mm axial slices acquisition in steady state, TR=40ms, TE=7ms, flip angle 30º, 1.02 × 1.02 × 1.5 mm3, T1-weighted gadolinium-enhanced image (T1-GdH), T2-weighted (T2-W) and proton density (PD) multi-slice spin-echo sequence of 124 contiguous, 1.0 mm thick axial slices (spin echo, TR= 40 ms, TE= 7 ms, 1.0 × 1.0 × 5.0 mm3). Pulse sequence characteristics were as follows: 256 x 256 matrix, FOV of 25.6 cm.

Measurement

Measures of the apathy-motivation continuum

The development of the Rasch version of AS (R-AS) and the Rasch Measure of Apathy (RMA) have been previously described. See Manuscript 2 and 3. Briefly, the R-AS comprises 10 of the original AS items but rescored into three multi-item sub-tests in order that the data fit the Rasch Model to derive a meaningful total score ranging from 0-42 with high score indicating greater motivation. The Rasch measurement psychometric criteria were supported with person separation index 0.70 and adequate scale-to-sample targeting of the data to the Rasch model (X^2=6.4, DF=12, p=0.73).

The RMA builds from the R-AS by adding an important fourth dimension for emotional function; it comprises 13 items reflecting all four domains that fit with the apathy construct. The Rasch measurement psychometric criteria were supported with person separation index 0.68 and adequate scale-to-sample targeting of the data to the Rasch model (X^2=27.8, DF=18, p=0.06). A total score ranges from 0-42 with higher scores indicating lower apathy. For both measures the cut-off score ≤ 29 was used. This cut-off value was supported by data presented in Manuscript 4 be a valid value to classify subjects within the apathetic/low motivation range.

Lesion mapping

The Lesions of these patients were manually delineated by a trained rater (CS) using an interactive visualization software package, Display (McConnell Brain Imaging Centre, MNI), which provides three-dimensional (coronal, axial and sagittal) localization on MRI. Lesions were identified as being intra-axial and hyperintense on T2 FLAIR or PD images. When in doubt, radiologist reports and diffusion-weighted images were used to confirm lesion location. Corresponding delineated lesion maps were subsequently transformed to standardized space using the same transformation matrix calculated on the native MRI volume.

Statistical analysis

The working hypothesis was that lesions located on the frontal/striatal regions of the brain would be associated with lower scores on the apathy-motivation continuum (i.e. more apathetic). Three statistical approaches were taken. First, the mean scores on the R-AS and RMA were compared between subjects with lesions on the frontal/striatal regions and subjects with lesions elsewhere (Table 7.1). In this analysis the outcome, scores on the R-AS and RMA, were continuous and the exposure (independent variable) was lesion location a binary variable. The independent t-test was used and assumptions held normality across level of exposure, and homoscedasticity.

Second, the proportion of people classified as apathetic or motivated, based on the cut-off score of 29 on the R-AS and RMA, was calculated. As both the outcome, apathetic or motivated, and the exposure, lesion location, are binary a chi-square test was used.

The third analysis contrasted extreme groups of highly apathetic and highly motivated people on lesion location. These groups were selected based on their assigned trajectory, as described in Manuscript 3 with the highest trajectory representing the most highly motivated group and the lowest trajectory the group with the lowest motivation. Because the sample size did not warrant statistical testing and the purpose of the study was to generate preliminary data for future research, a descriptive approach was taken. Brain images for these two extreme groups were compared by overlaying images of apathetic subjects on images of motivated subjects. Using the number of voxels as the metric for lesion size, the two sets of images were subtracted and the results displayed

in a colour bar spectrum. Images of highly motivated subjects were subtracted from images of highly apathetic subjects. In the colour bar spectrum, blue represents negative result (i.e. lesions of motivation subjects were dominant) and red reflects positive results (i.e. lesions of apathetic subjects were dominant).

RESULTS

Table 7.1 shows the comparison of the two measures of interest with the hypothesized lesion location. There were no differences between the average scores of the R-AS and RMA in subjects with lesion on the frontal/striatal region in comparison to subjects with lesions in other areas, however, the mean score on the R-AS for subjects with lesions in the frontal/striatal region (26.0) was 3.5 points lower than for subjects with other lesions (29.5) yielding an effect size of 0.55.

The same pattern can be observed in Table 7.2. Even though there were no significant differences in the chi-square analysis, the frequency of lesions in the frontal/striatal region for apathetic subjects measured by the R-AS and RMA was much higher than of lesions in other areas. Motivated patients seem to have equal frequencies across lesion locations.

Figure 7.1 illustrates the lesion overlapping of all subjects in a standard brain template. No common lesion pattern was observed; lesions were fairly spread. Figure 7.2 shows a comparison of images of highly apathetic subjects with images of highly motivated

subjects (measured by both R-AS and RMA). In picture A, it appears that subjects with lesions on the left and posterior side of the brain have their motivation spared (probably without apathy). On the other hand, in picture B, it is clearer pattern that lesions in the right frontal lobe are more likely to be associated with apathy. However, the small sample size does not give enough power to make strong conclusions.

DISCUSSION

The only evidence for a link between the apathy-motivation continuum came from comparing extreme groups of people, those at the high motivation end of the continuum to those at the apathy end. Despite the limitations of the analyses, there is visual evidence that the apathy syndrome may be associated with lesions on the right frontal lobe, while motivation does not seem to be related to any specific lesion location.

This finding concurs with Brodaty et al.[68] who also compared stroke subjects in the highest and lowest end of the spectrum of the Apathy Evaluation Scale (AES) showing a significant hyperintensity score for the right-sided and a trend for right fronto-subcortical circuit hyperintensity scores. There seems to be a consistency among studies in stroke associating frontal lobe lesions and apathy syndrome [37-39, 68], however, there was one study showing a predominant association between apathy and lesions in the right prefrontal cortex[39]; and one study suggesting that post stroke apathy may correlate with left frontal dysfunction [37].

The most important limitation of VAL statistics methods is that the statistical method offers very low power and, therefore, requires examining a relatively large group of subjects, a process that typically requires years of data collection. Therefore, the small sample size of 57 participants did not allow for making strong conclusions but, as an exploratory study, the data would be useful in designing an appropriately powered study. Another important limitation was using clinical MRIs. Although this has some advantages as they are available on all people and not just a sub-sample of those who agree for further research investigations, clinical MRIs are not optimally standardized for research purposes which can introduce variability across subjects.

The findings that lesions in the left side seem to spare motivation and that lesions in the right frontal areas impair motivation, if confirmed, would be informative for rehabilitation potential and to help carers better understand some of the behavioural changes associated with stroke that they find frustrating [174].

In addition, if symptoms of low motivation are not associated with the lesion location, then this would indicate a temporary state that could be improved by implementing motivational approaches during the recovery phase of stroke. In contrast, people with lesion-induced apathy may need more focused interventions using pharmacological agents and/or cognitive behavioural therapy. These have some evidence of inducing change in apathy as shown in the systematic review in Chapter 2.

CONCLUSION

This exploratory study provides preliminary data in support of distinctive features between apathy syndrome and symptoms of low motivation. Apathy may be a brain related construct but low motivation without emotional blunting may be a temporary state induced by the stroke experience. Further study of these important constructs is warranted.

Table 7.1: Comparison of AS, R-AS, and RMA on lesion location

	Apathy-motivation(0-42)		
	AS	R-AS	RMA
Lesion Location	(Mean ± SD)	(Mean ± SD)	(Mean ± SD)
Frontal / Striatal	29.0 ± 6.8	26.0 ± 5.2	26.0 ± 4.5
Other	29.8 ± 6.6	29.5 ± 7.3	26.4 ± 6.5
Effect size (Cohen's d)	0.11	0.55	0.07

Apathetic ≤ 29

Table 7.2: Frequency of striatal/frontal lesion among people classified according to apathy status.

		Apathetic	Motivated	Total
AS	Striatal/Frontal	22 (64.7%)	15 (65.2%)	37
	Other	12 (35.3%)	8 (34.8%)	20
	Total	34 (100%)	23 (100%)	57
R-AS	Striatal/Frontal	27 (71.0%)	10 (55.5%)	37
	Other	11 (29.0%)	8 (44.4%)	19
	Total	38 (100%)	18 (100%)	57
RMA	Striatal/Frontal	31 (68.8%)	6 (50.0%)	37
	Other	14 (31.2%)	6 (50.0%)	20
	Total	45 (100%)	12 (100%)	57

AS – original Apathy Scale; R-AS – Rasch version of AS; RMA- Rasch measure of Apathy; Apathetic ≤ 29; all not significant ($p > 0.22$)

Figure 7.1: Overlay of all subjects MRI images.

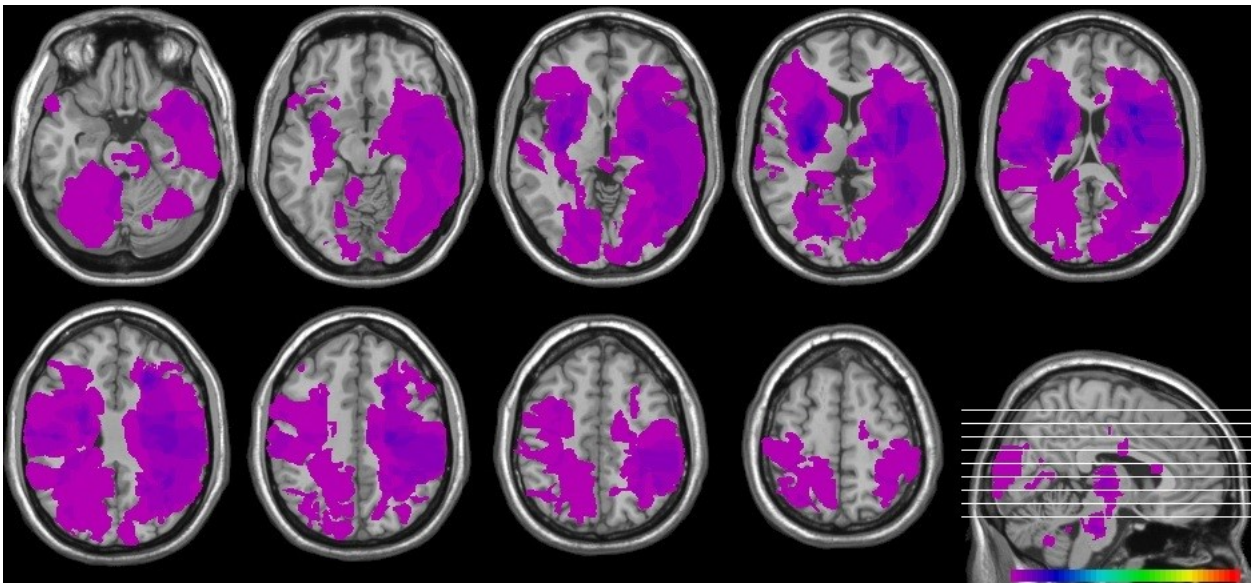
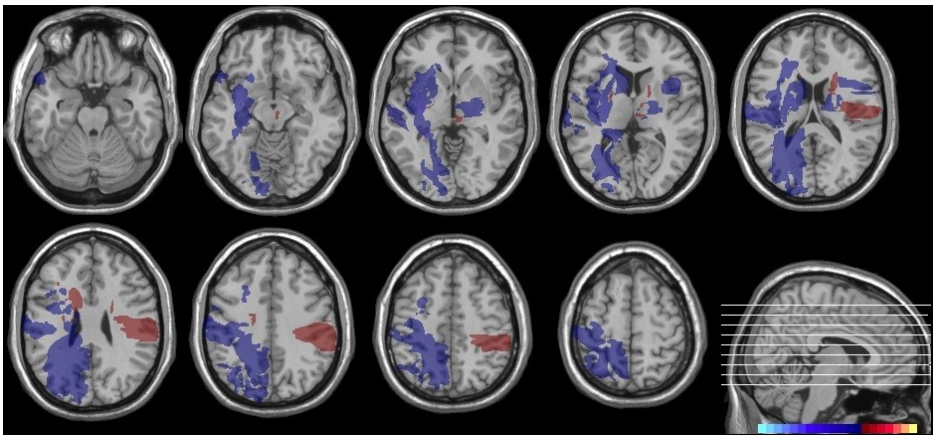
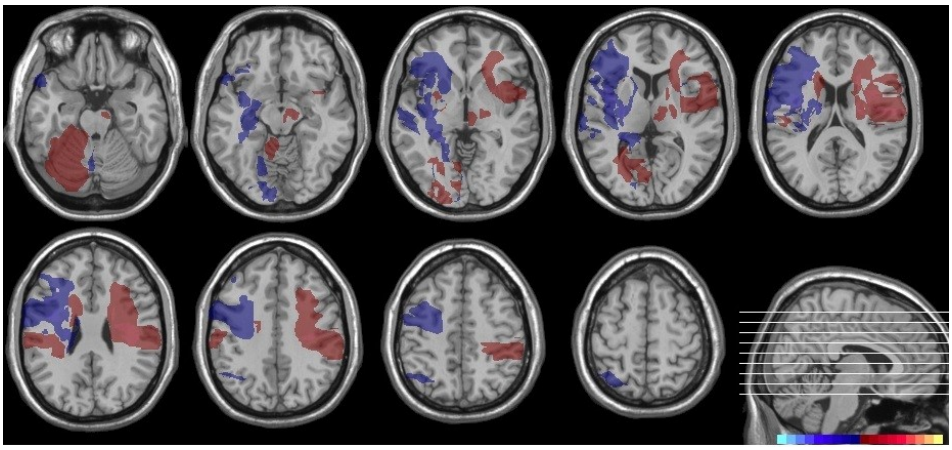


Figure 7.2: MRI images overlapping of apathetic subjects (in red) versus motivated subjects (in blue) as measured by R-AS and RMA.

A: Motivation (R-AS)



B: Apathy Syndrome (RMA)



CHAPTER 8 - SUMMARY OF RESULTS, DISCUSSION AND CONCLUSION

SUMMARY OF RESULTS

The overall objective of this thesis is to contribute to the understanding of the role of motivation in stroke rehabilitation by taking a longitudinal view of the apathy-motivation continuum. The studies presented in this thesis followed three phases: (i) construct conceptualization (literature review and Manuscript 1); (ii) construct measurement (Manuscript 2 and 3); and (iii) construct impact (Manuscript 4).

The Literature review and Manuscript 1 contributed to the understanding of the construct of apathy: how it is defined, how it has been measured, what are the clinical manifestations and associations, and what interventions have been tried. This conceptual work was the foundation for the other three manuscripts. Manuscript 2 contributed to the understanding of the conceptual and psychometric properties of the most used measure of the apathy syndrome in Stroke, the Apathy Scale (AS) [9], and provided a new version of the measure, with better psychometric properties, and reflecting the motivation continuum. Manuscript 3 solved the AS conceptual limitations identified in Manuscript 2 and by including items for emotional function created a measure of the apathy-motivation continuum with good psychometric properties. Preliminary mapping of the estimates derived from the motivation measure (R-AS) and

from the apathy-motivation measure (RMA) to the brain provided evidence that the symptoms of the apathy-motivation continuum, rather than only symptoms of the motivation continuum have some association with brain lesions. There was a trend for lesions in the right frontal hemisphere to be associated with apathy.

Finally, the group-based trajectory analysis carried in Manuscript 4 revealed five trajectories of motivation and four trajectories of apathy. The majority of participants showed moderate levels of apathy that remained stable over time, but with an important impact on participation.

DISCUSSION

While the findings from each study have been discussed separately, this overall discussion aims to emphasize the relevance of these finding in the context of rehabilitation, summarize the lessons learned from undertaking this work, and provide implications for the advancement of knowledge.

Manuscripts 2 and 3 emphasized the importance of construct conceptualization and developing good items. Lack of clarity about the definition of apathy had an important effect on the psychometric properties of the original AS.

A conceptual model harmonizing all the terminology was presented (see Figure 4.1) showing a clear distinction between apathy symptoms and the apathy syndrome. The measurement aspects of symptoms also indicates that it is possible to measure a spectrum of motivation (R-AS) ranging from low-motivation to high motivation; it is also possible to measure a spectrum of apathy symptoms (RMA) by including items for emotional response with a potential range from apathetic to motivated. Without strong conceptualization, these distinctions would not have been possible, weakening the measurement model.

The literature agrees that impaired motivation is necessary (but not sufficient) for the apathy syndrome. Emotional response differentiates a temporary lack of motivation arising from a major life change event from a potential apathy syndrome. For example, a person who has lost his or her job may lack interest and effort to look for a new job, but he or she will probably also be sad about it, or be anxious about the future. As emotional “caring” is present, this scenario likely describes a temporary state of apathetic symptoms. On the other hand, if there was no emotional reactivity associated with the event such that the person was indifferent to the job loss, this could indicate a more severe clinical manifestation. When nothing is capable of stimulating a person’s interest and motivation, there will probably be impairment in function and limitation on personal, social, or occupational aspect of his life; the literature indicates this to reflect a more severe form state along the apathy-motivation continuum.

The results presented in Manuscript 3 did not support that emotional function changed the severity of the construct. The inclusion of items reflecting emotion (presence or absence of emotional response, like feeling sad or anxious) did not yield significant change in estimations in the RMA in comparison to the motivation continuum measure (R-AS).We expected that the inclusion of the emotional items would make a difference to the scores of people with emotional blunting and that some thresholds of these items would be situated at the apathy end of the spectrum. But instead, the three emotional items were evenly distributed along the continuum, as shown by the location map (Table 5.5).

This may be explained by the fact that these items were not specifically developed for this purpose. But it also raised concern about the definition of the emotional component. Looking back at the original AS items, the emotional component was captured by items reflecting concern, indifference, and awareness. The ICF mapping linked these items to “insight” and “interest”.

Similarly, in the Apathy Evaluation Scale developed by Marin [31], there are five items related to “importance” (e.g. “Seeing a job through the end is important to him”) and “concern” (e.g. “he is less concerned about his problem than he should be”) and only one item actually questioning about emotional response. Based on ICF conceptualization, importance and concern is related to the “insight” category which is a different category than “emotional function”. Future research should focus on clarifying

the conceptual definition of this “emotion-insight” component so that better items can be developed to measure the symptomatology of the apathy syndrome.

But why is it relevant to make a distinction between temporary apathy symptoms and apathy syndrome? Because, even though the literature on apathy is still inconclusive, there is evidence suggesting that the presence of the apathy syndrome can have a great impact in all aspects of the patient’s life. Apathetic patients will also cause more burdens on caregivers and make the rehabilitation process more difficult and less effective. The results of Manuscript 4 showed that apathetic patients, measured by both measures, only achieved “fair” participation at the end of the first year of stroke. The only evidence that apathy as measured by the RMA may be more severe than low motivation as measured by the R-AS, was that apathetic subjects defined by the RMA required higher scores to reach excellent participation compared to subjects with low motivation defined by R-AS. An improvement in RMA score from 32 to 38.5 at 12 months was necessary to achieved excellent participation. On the other hand, a score of 30 in the R-AS, that remained stable over time, was sufficient to achieve excellent participation. Therefore, apathy syndrome, measure by RMA, may be more disabling than impaired motivation, measured by R-AS.

The preliminary brain analyses further supported distinctive features between apathy syndrome and symptoms of low motivation. Despite the limitations of the analyses, subjects with low motivation (as measured by the R-AS) showed no commonalities in

brain lesions, suggesting that location along the motivation continuum may not be a brain related construct, and might be a temporary adaptation to all the changes imposed by the stroke. If this is the case, low motivation maybe improved by implementing motivational approaches during the recovery phase of stroke.

In contrast, the symptomatology of the apathy syndrome (RMA) had a more evident relationship with lesions on the right frontal-striatal region. This finding is consistent with the literature showing that right hemispheric stroke may develop an "indifference reaction"[20]. Frontal lobe lesions have also been frequently associated with apathy syndrome in Stroke [37, 38, 68], Alzheimer's disease [45-47, 50, 175-177], and Parkinson's disease [54]. Depression on the other hand, has been shown to be associated with lesions of the right temporal-parietal lobe [124, 125], which strengthens the assumption that apathy and depression are different constructs.

There is a need for an adequately powered study to confirm this preliminary finding. Future work is needed to test the hypothesis that lesions in specific locations produce apathy. To do this would require that apathy symptoms are assessed using psychometrically sound methods, and that that some information should be obtained on pre-stroke personality.

The main clinical message of this thesis is that rehabilitation professionals need to understand and be able to differentiate between a temporary lack of motivation caused by the stroke sequelae and a potential apathy syndrome caused by the brain lesion. A

temporary diminished level of motivation is expected after experiencing a disabling health event as stroke. Using a measure of motivation symptoms, like the R-AS, can help to quantify patients' level of motivation and guide tailored interventions to modify this state. Therapists can stimulate motivation in a number of ways: setting specific attainable goals; creating simple tasks that are challenging and meaningful to the patient; and providing clear instructions on how to complete the task so that the patient has confidence that he or she has some control over the task [178]. Performance feedback can also enhance interest and curiosity with the rehabilitation process and functional recovery [179].

For example, if the patient presents a low score in the R-AS, in addition to signs of indifference and emotional blunting, he or she should be referred to further diagnostic evaluation. The impaired motivation might be explained by more than external factors, and if so motivational approaches may have limited effect. These patients may benefit from pharmacological therapy or cognitive-behaviour therapy as demonstrated in Manuscript 1. But most importantly, patients, caregivers, and health professionals can learn to adapt to this condition, instead of being misunderstood and judged as lazy, challenging, passive, pessimists, depressed, or even demented.

Another lesson learned from this work was the power of Rasch analysis as a framework for assessing the structure of items, scales and the underlying construct. That location along the apathy- motivation continuum did not change over time in this cohort is of

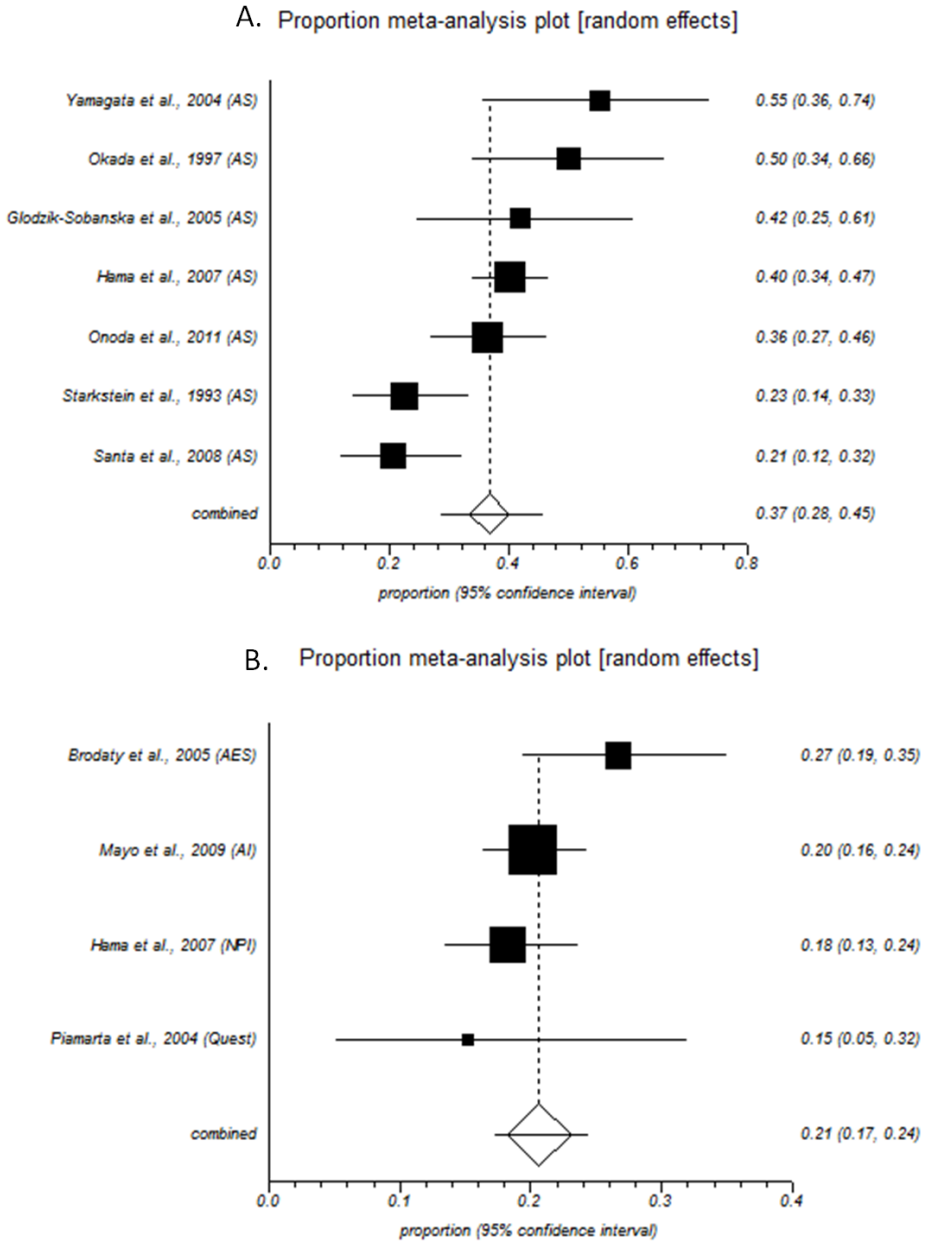
concern. Rehabilitation professionals spend more time on therapies targeting walking and function for basic activities of daily living, doing things for patients when they do not wish to or cannot do for themselves. This work may be a way forward for rehabilitation professionals to adapt therapies to enhance motivation or limit the negative effects of apathy.

CONCLUSION

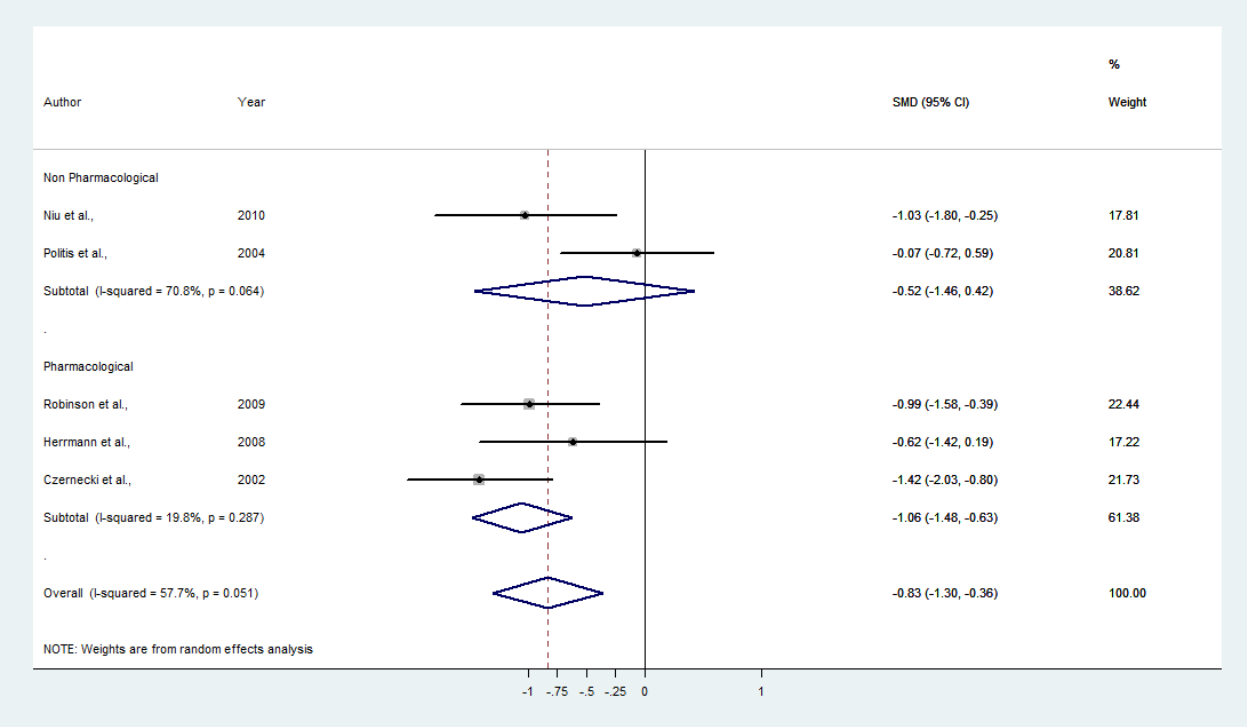
Apathy is a difficult construct to understand. It is no wonder that it has not received the attention it needs in the rehabilitation field. Of all the professions dealing with people recovering from stroke, rehabilitation professionals are the best placed to take on the challenge of enhancing motivation in their patients. This work has put some tools, albeit with residual limitations, in their hands.

APPENDICES

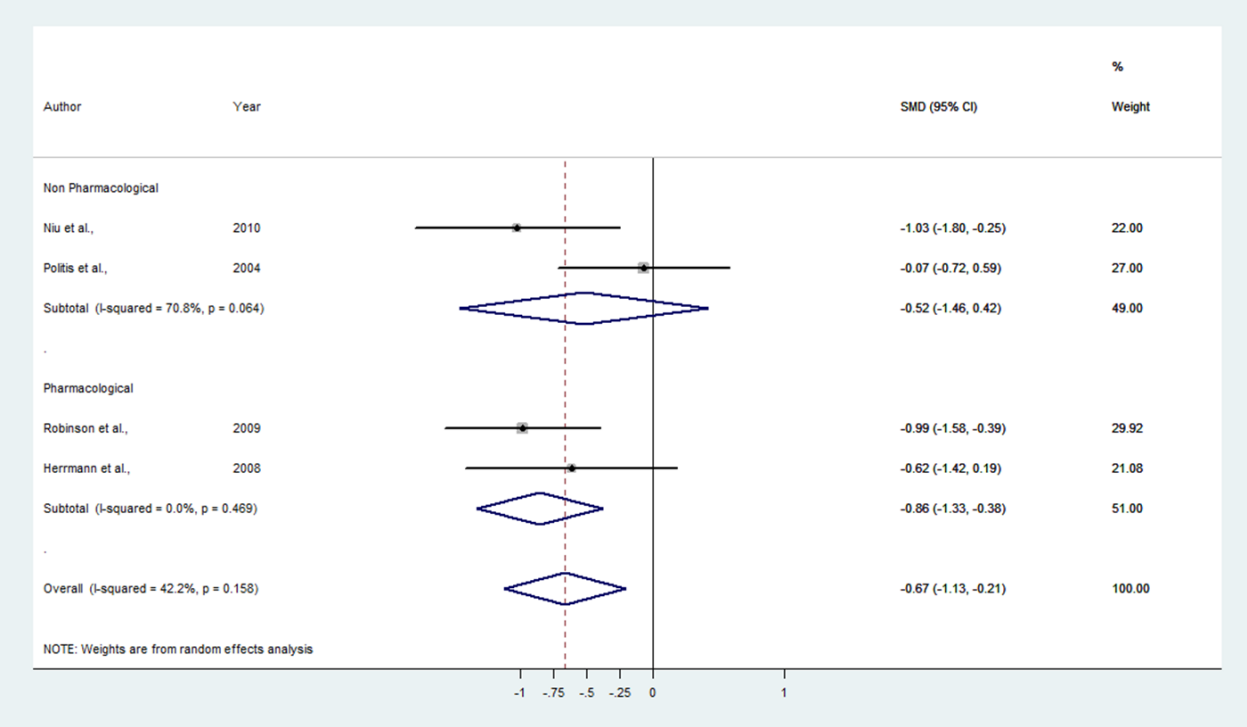
A1. Forest plot of the proportion of apathy in stroke population for studies that used the AS as outcome measure; and from studies that used other measures and indices



A2. Effect sizes of interventions on Apathy with overall effect



A3: Effect size of intervention on apathetic neurological patients



A4: Starkstein's Apathy Scale

Apathy Scale Scoring: For questions 1-8, not at all = 3 points; slightly = 2; some = 1; a lot = 0. For questions 9-14, not at all = 0; slightly = 1; some = 2; a lot = 3.

1. Are you interested in learning new things?

not at all slightly some a lot

2. Does anything interest you?

3. Are you concerned about your condition?

4. Do you put much effort into things?

5. Are you always looking for something to do?

6. Do you have plans and goals for the future?

7. Do you have motivation?

8. Do you have the energy for daily activities?

9. Does someone have to tell you what to do each day?

10. Are you indifferent to things?

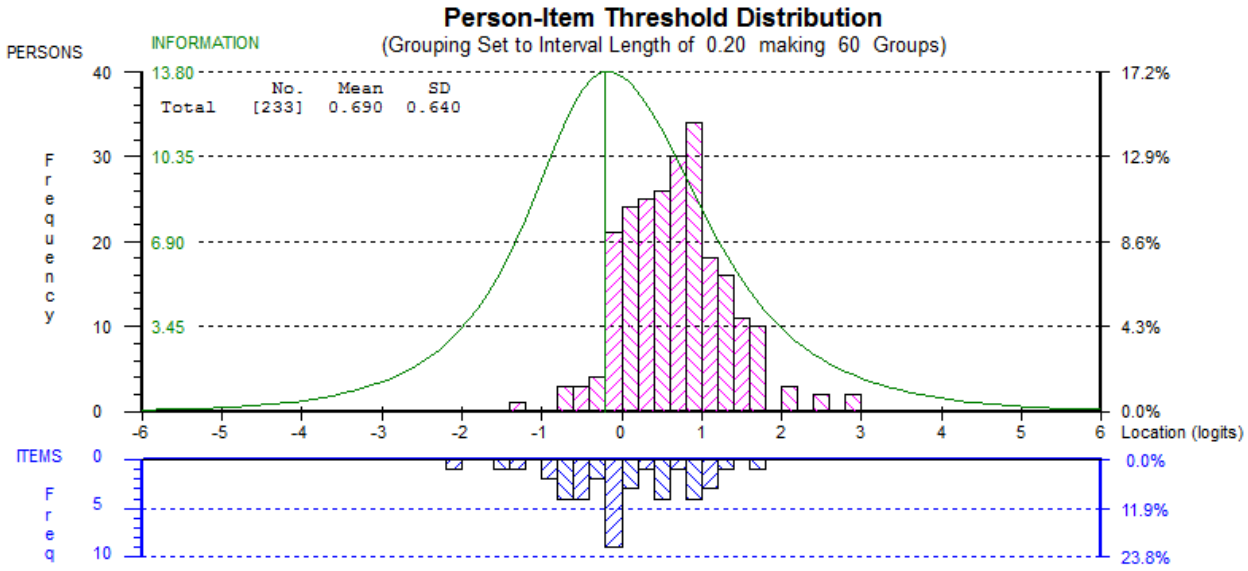
11. Are you unconcerned with many things?

12. Do you need a push to get started on things?

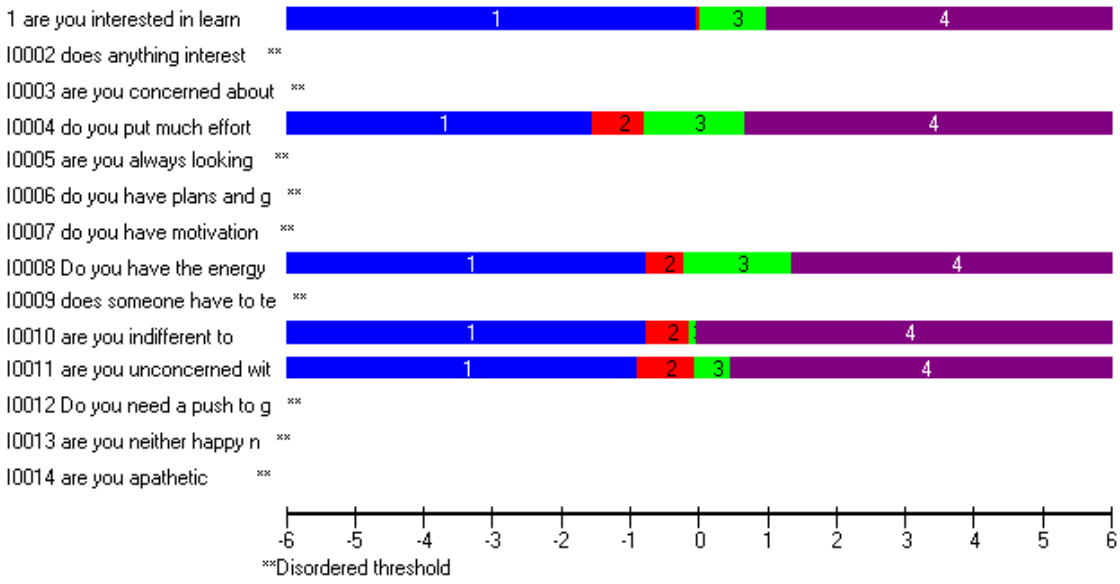
13. Are you neither happy nor sad, just in between?

14. Would you consider yourself apathetic?

A5: The item-person threshold distribution and test information function before analysis



A6: Threshold map for the 14-item Apathy Scale



Items are listed in order of increasing item number (item 1 to item 14) on the y-axis. Persons' measure of ability (logits) is on the x-axis ranging from -6 to +6.

A7: Differential item functioning by personal factors

Personal Factors		Items	MS	F-stat	df	p
Age	5	Are you always looking for something to do?	5.02	5.89	3	0.00055
Gender	8	Do you have energy for daily activities?	20.07	23.96	1	0.00003
	9	Does someone have to tell you what to do each day?	16.78	11.71	1	0.00074
Education	5	Are you always looking for something to do?	4.78	5.25	4	0.00046
	6	Do you have plans and goals to the future?	3.85	5.60	4	0.00023
	9	Does someone have to tell you what to do each day?	7.38	5.19	4	0.00052
Native language	6	Do you have plans and goals to the future?	6..51	11.86	2	0.00001
	7	Do you have motivation	5.55	9.22	2	0.00014
	8	Do you have energy for daily activities?	15.07	18.87	2	0.00000
	10	Are you indifferent to things?	8.59	12.88	2	0.00000
	11	Are you unconcerned with many things?	7.00	7.62	2	0.00064
	12	Do you need a push to get started on things?	9.91	9.32	2	0.00012
	13	Are you neither happy nor sad, just in between?	8.57	7.72	2	0.00058
Test language	6	Do you have energy for daily activities?	9.34	16.95	1	0.00005
	7	Do you have motivation	6.54	10.80	1	0.00117
	8	Do you have energy for daily activities?	18.57	21.83	1	0.00000
	10	Are you indifferent to things?	9.43	13.37	1	0.00032
	12	Do you need a push to get started on things?	18.78	17.65	1	0.00004

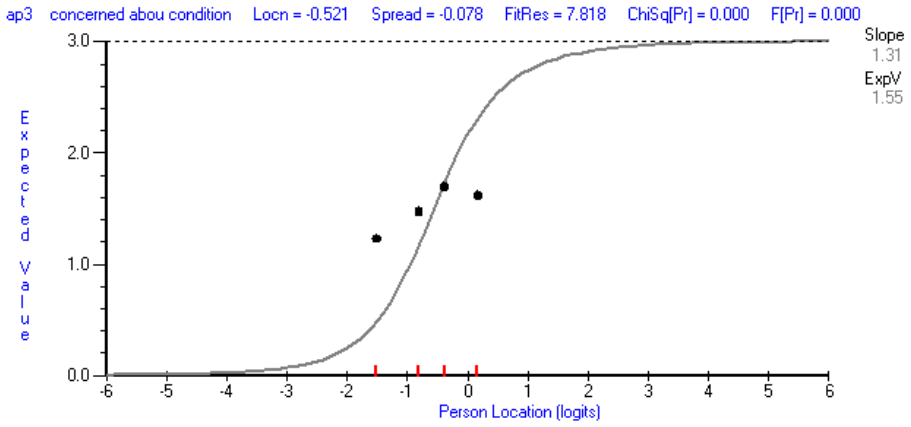
A8: Summary of Global Fit Statistics for the Rasch Model of the original AS overall time points and for each time point separately

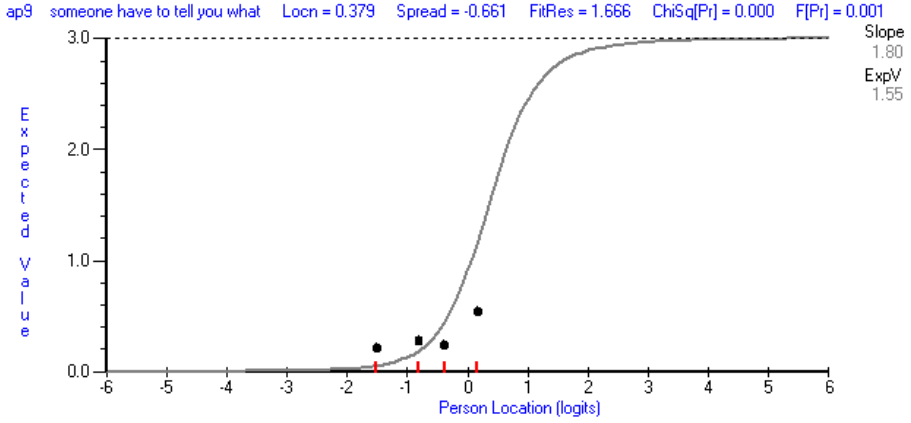
	All time points combined (n=232)	3 months (n=53)	6 months (n=57)	9 months (n=54)	12 months (n=68)
Item-Trait Interaction					
Total item Chi-Square	102.0	49.5	56.8	52.8	92.17
Total Deg of Freedom	42	42	42	42	42
Total Chi-Square Probability	0.00000	0.19	0.06	0.12	0.0000
Item-Person Interaction					
Item					
Difficulty	0.00 ± 0.98	0.00 ± 0.93	0.00 ± 0.96	0.00 ± 1.24	0.00 ± 1.88
Fit Residual	0.10 ± 2.42	0.16 ± 0.89	0.14 ± 1.41	0.06 ± 1.17	0.02 ± 1.83
Person					
Measure of ability	0.97 ± 1.01	0.84 ± 0.94	0.89 ± 1.10	1.21 ± 1.00	1.49 ± 1.03
Fit Residual	-0.33 ± 1.13	-0.20 ± 1.14	-0.24 ± 1.01	-0.18 ± 0.75	-0.32 ± 1.14
Reliability					
PSI	0.719	0.69	0.75	0.70	0.72

A9: 14 original AS items difficulty and standard error overall time points and for each time point separately

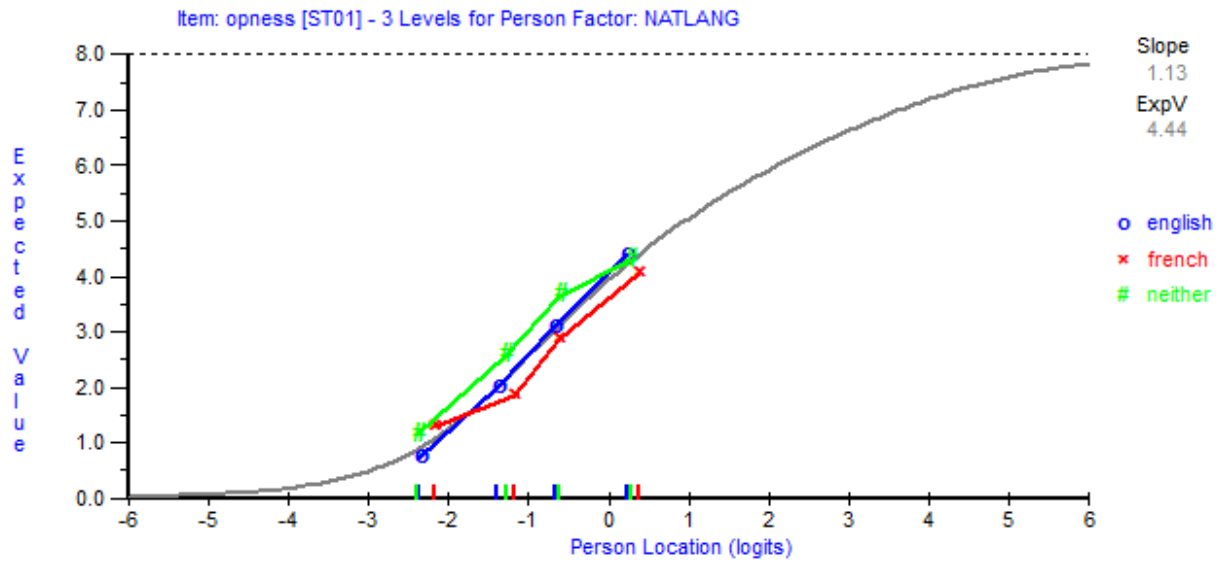
Item #	All time points		3 months		6 months		9 months		12 months	
	Difficulty	SE	Difficulty	SE	Difficulty	SE	Difficulty	SE	Difficulty	SE
9	-2.18	0.29	-1.56	0.46	-2.44	0.61	-3.21	1.02	-1.66	0.52
14	-0.92	0.13	-0.68	0.25	-0.68	0.24	-0.71	0.28	-4.26	0.34
4	-0.86	0.13	-1.07	0.28	-0.44	0.25	-0.83	0.29	-3.47	0.27
10	-0.51	0.12	-0.55	0.25	-0.60	0.25	-0.68	0.28	0.10	0.22
12	-0.48	0.12	-0.26	0.25	-0.53	0.24	-0.93	0.28	0.05	0.22
11	-0.25	0.12	-0.62	0.26	-0.06	0.23	-0.08	0.26	0.22	0.23
2	-0.05	0.12	-0.32	0.25	0.25	0.24	0.08	0.27	0.29	0.22
8	0.17	0.13	0.25	0.27	-0.16	0.28	0.14	0.30	0.73	0.21
13	0.25	0.11	-0.16	0.24	0.41	0.23	0.60	0.25	0.70	0.21
1	0.46	0.11	0.41	0.23	0.30	0.24	0.82	0.22	0.84	0.21
6	0.53	0.11	0.56	0.23	0.55	0.21	0.63	0.24	0.90	0.20
3	1.06	0.14	0.79	0.30	0.99	0.29	1.57	0.30	1.47	0.27
5	1.16	0.11	1.38	0.24	0.92	0.22	1.19	0.22	1.79	0.21
7	1.61	0.15	1.82	0.32	1.49	0.30	1.41	0.30	2.31	0.28

A10: Item characteristic curves of items 3 and 9

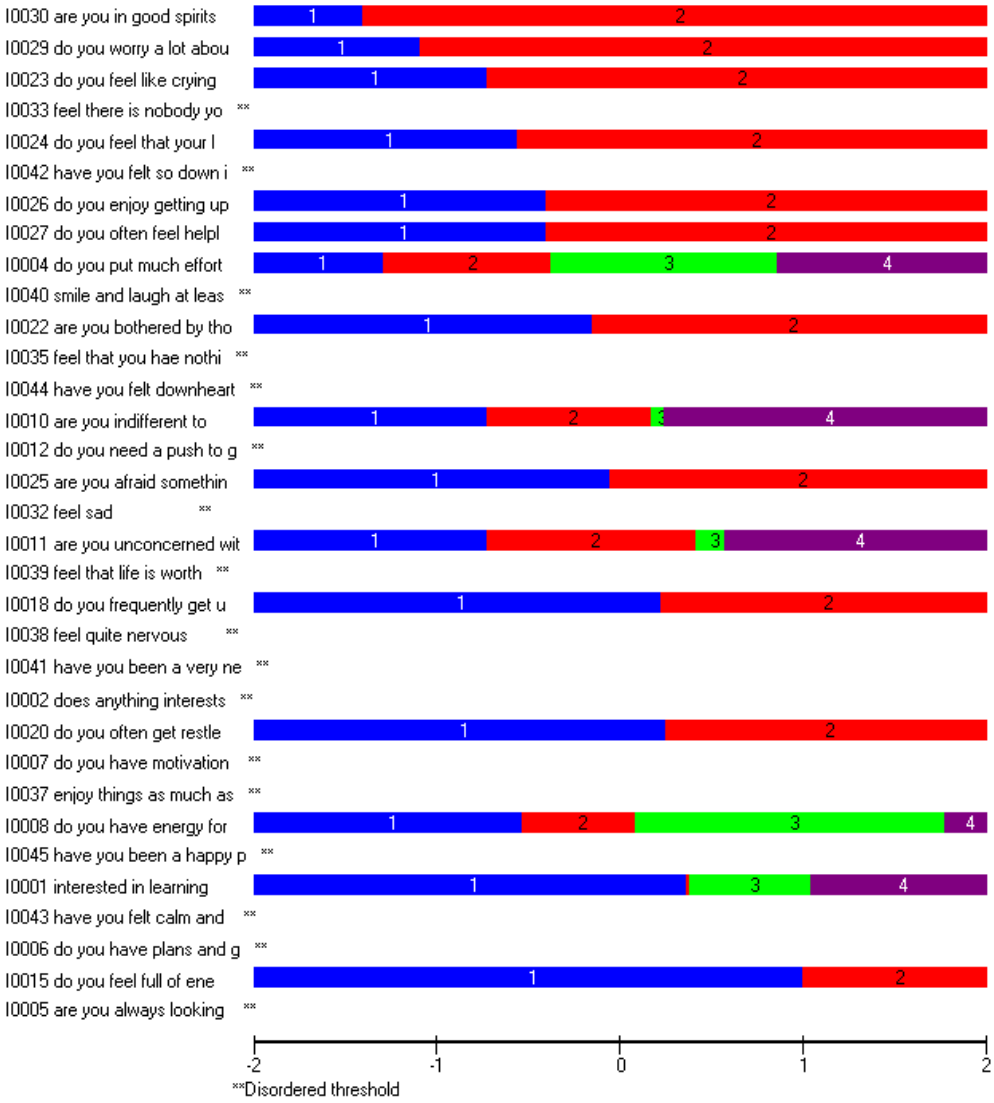




A11: Item characteristic curve plotted by language spoken for subtest 1, openness to experience.



A12: Threshold map for the 33-item Apathy Scale

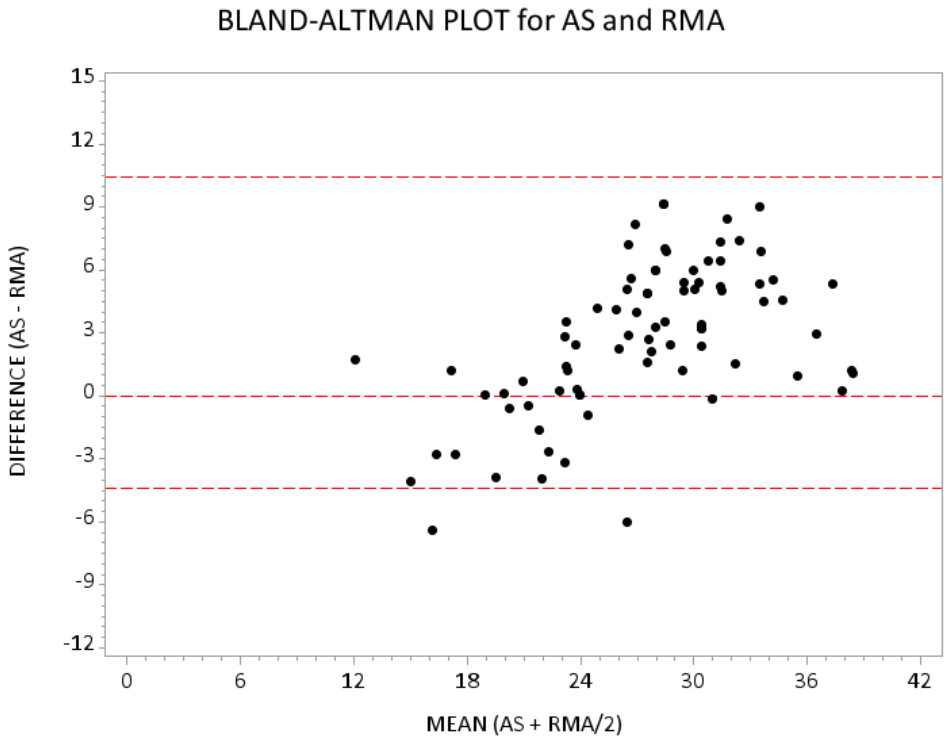


A13: Results of item analysis of the 33-item model

Index code	N° of Threshold	Item Fit	Local dependency	DIF	PCA
R_AS 1.1	3	3.9	1.2	Language	+
R_AS 1.2	4	2.6			+
R-AS 1.3	**		R-AS 1.4	Education	+
R-AS 1.4	3	2.6	R-AS 2.4	Age	+
R-AS 2.1	4				+
R-AS 2.3	4			Gender, language	+
R-AS 2.4	**	6.3		Language	+
R-AS 3.1	**		1.3, 1.4	Language	+
R-AS 3.2	**		1.3, 1.4, 2.1, 2.4, 3.1	Language	+
GDS12	2				+
SIS3f	**			Language	+
SIS3h	**		SIS3i		+
R-AS 12	**				
GDS 1	2				
SIS3i	**				
GDS16	2				
GDS4	2		GDS6, GDS9		-
GDS6	2				-
GDS8	2				-
GDS9	2				-
GDS10	2	<0.05		Gender	-
GDS11	2				-
GDS13	2				-
GDS15	2				-
SIS3a	**		MHIc		-
SIS3b	**				-
SIS3d	**		GDS10, SIS3b		-
SIS3g	**		GDS11, MHIb,		-
MHIb	**				-
MHIc	**				-
MHId	**				-
MHIf	**		SIS3a, SIS3b, MHIc		-
MHIh	**				-

** Disordered thresholds; * fit residuals value; PCA – Principal Component Analysis; DIF – Differential Item Functioning; Language – language first spoken at birth (English, French or Neither); Test Language – Language of the measure (English or French)

A14: Bland-Altman plot of the difference between original AS and RMA estimations



A15: Descriptive statistics over four time points for each group trajectory of the AS, R-AS, RMA, and SIS-participation

Group Membership		3 months	6 months	9 months	12 months
	N (%)	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)
AS					
Apathetic	16 (19.8%)	20.7 (19.2, 22.2)			
Moderate apathy	30 (37.0%)	28.1 (26.8, 29.4)			
Motivated	35 (43.2%)	33.8 (32.7, 35.0)			
R-AS					
Apathetic	13 (16.0%)	19.0 (17.1, 20.9)			
Moderate motivation	26 (32.1%)	24.4 (22.4, 26.3)			
Improving	4 (5%)	20.7 (16.6, 24.8)	26.0 (23.3, 28.6)	31.3 (28.5, 34.0)	36.5 (31.0, 42.0)
Motivated	35 (43.2%)	29.3 (28.2, 30.4)			
Highly motivated	3 (3.7%)	41.7 (38.8, 42)	37.5 (32.5, 42)	35.9 (31.2, 42)	39.5 (34.9, 42)
RMA					
Apathetic	22 (27.2%)	21.1 (20.0, 22.4)			
Moderate apathy	42 (51.8%)	25.7 (24.7, 26,7)			
Moderate motivation	12 (14.8%)	29.6 (28.1, 31.1)			
Highly motivated and improving	5 (6.2%)	32.7 (28.8, 36.5)	35.2 (31.7, 38.7)	34.6 (31.7, 35.6)	38.5 (35.0, 41.8)
SIS- Participation					
Poor	12 (14.8%)	32.9 (22.1, 43.8)			
Fair	33 (40.7%)	51.1 (44.0, 58.2)	55.8 (49.3, 62.4)	60.6 (53.3, 67.9)	65.3 (55.9,74.7)
Very Good	31.0 (38.3%)	79.2 (71.2, 87.1)	82.0 (75.4, 88.7)	84.7 (77.4, 92.1)	87.3 (78.2, 96.4)
Excellent	5 (6.2%)	89.2 (68.5, 100)	97.7 (93.0, 100)	99.8 (89.9,100)	100 (98.7, 100)

REFERENCES

1. Feibel J and Springer C. Depression and failure to resume social activities after stroke. *Arch Phys Med Rehab*. 1982; 63: 276.
2. Bacher Y, Korner-Bitensky N, Mayo N and Becker R. A longitudinal study of depression among stroke patients participating in a rehabilitation program. *Canadian Journal of Rehabilitation*. 1990.
3. Wade DT, Legh-Smith J and Hewer R. Depressed mood after stroke. A community study of its frequency. *The British Journal of Psychiatry*. 1987; 151: 200.
4. Sinyor D, Amato P, Kaloupek D, Becker R, Goldenberg M and Coopersmith H. Post-stroke depression: relationships to functional impairment, coping strategies, and rehabilitation outcome. *Stroke*. 1986; 17: 1102.
5. Nelson LD, Cicchetti D, Satz P and Sowa M. Emotional sequelae of stroke: A longitudinal perspective. *Journal of clinical and experimental neuropsychology*. 1994.
6. Mayo NE, Korner-Bitensky NA and Becker R. Recovery time of independent function post-stroke. *Am J Phys Med Rehabil*. 1991; 70: 5.
7. Parikh RM, Robinson RG, Lipsey JR, Starkstein SE, Fedoroff JP and Price TR. The impact of poststroke depression on recovery in activities of daily living over a 2-year follow-up. *Archives of Neurology*. 1990; 47: 785.
8. Morris P, Raphael B and Robinson RG. Clinical depression is associated with impaired recovery from stroke. *The Medical Journal of Australia*. 1992; 157: 239.
9. Statistics. Heart & Stroke Foundation of Canada, 2012.
10. Statistics. Heart & Stroke Foundation of Canada.
11. Leading Cause of Death, 2009. The Daily: Statistics Canada, 2012.
12. Umphred D and Carlson C. *Neurorehabilitation for the Physical Therapist Assistant*. Slack Incorporated, 2006.
13. Bogousslavsky J, Regli F, Delaloye B, Delaloye Bischof A, Assal G and Uske A. Loss of psychic self activation with bithalamic infarction. *Acta neurologica scandinavica*. 1991; 83: 309-16.

14. Kelley R and Borazanci A. Stroke rehabilitation. *Neurological research*. 2009; 31: 832-40.
15. Kaufman S and Becker G. Stroke: Health care on the periphery. *Social Science & Medicine*. 1986; 22: 983-9.
16. Maclean N, Pound P, Wolfe C and Rudd A. Qualitative analysis of stroke patients' motivation for rehabilitation. *Bmj*. 2000; 321: 1051-4.
17. Maclean N, Pound P, Wolfe C and Rudd A. The concept of patient motivation a qualitative analysis of stroke professionals' attitudes. *Stroke*. 2002; 33: 444-8.
18. Maclean N, Pound P, Wolfe C and Rudd A. The concept of patient motivation: a qualitative analysis of stroke professionals' attitudes. *Stroke*. 2002; 33: 444-8.
19. Landes AM, Sperry SD, Strauss ME and Geldmacher DS. Apathy in Alzheimer's disease. *Journal of the American Geriatrics Society*. 2001; 49: 1700-7.
20. Marin RS. DIFFERENTIAL-DIAGNOSIS AND CLASSIFICATION OF APATHY. *American Journal of Psychiatry*. 1990; 147: 22-30.
21. Marin RS. APATHY - A NEUROPSYCHIATRIC SYNDROME. *Journal of Neuropsychiatry and Clinical Neurosciences*. 1991; 3: 243-54.
22. Starkstein SE. Apathy and withdrawal. *Int Psychogeriatr*. 2000; 12: 135-7.
23. Mizrahi R and Starkstein SE. Epidemiology and management of apathy in patients with Alzheimer's disease. *Drugs & Aging*. 2007; 24: 547-54.
24. Robert P, Onyike CU, Leentjens AFG, et al. Proposed diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders. *Eur Psychiatry*. 2009; 24: 98-104.
25. Starkstein SE and Leentjens AFG. The nosological position of apathy in clinical practice. *Journal of Neurology, Neurosurgery & Psychiatry*. 2008; 79: 1088-92.
26. Morse JM, Hupcey JE, Mitcham C and Lenz ER. Concept analysis in nursing research: a critical appraisal. *Research and Theory for Nursing Practice*. 1996; 10: 253-77.
27. Cummins RA. Proxy responding for subjective well-being: A review. *International Review of Research in Mental Retardation, Vol 25*. 2002; 25: 183-207.

28. Schwartz CE and Sprangers MA. *Adaptation to changing health: Response shift in quality-of-life research*. American Psychological Association, 2000.
29. Starkstein SE, Mayberg HS, Preziosi TJ, Andrezejewski P, Leiguarda R and Robinson R. Reliability, validity, and clinical correlates of apathy in Parkinson's disease. *Journal of Neuropsychiatry and Clinical Neurosciences*. 1992; 4: 134.
30. Starkstein SE, Fedoroff JP, Price TR, Leiguarda R and Robinson RG. APATHY FOLLOWING CEREBROVASCULAR LESIONS. *Stroke*. 1993; 24: 1625-30.
31. Marin RS, Biedrzycki RC and Firinciogullari S. Reliability and validity of the apathy evaluation scale* 1. *Psychiatry Res*. 1991; 38: 143-62.
32. Robert PH, Claire S, Benoit M, et al. The apathy inventory: Assessment of apathy and awareness in Alzheimer's disease, Parkinson's disease and mild cognitive impairment. *Int J Geriatr Psychiatry*. 2002; 17: 1099-105.
33. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi D and Gornbein J. The neuropsychiatric inventory. *Neurology*. 1994; 44: 2308.
34. Clarke DE, Ko JY, Kuhl EA, van Reekum R, Salvador R and Marin RS. Are the available apathy measures reliable and valid? A review of the psychometric evidence. *Journal of Psychosomatic Research*. 2011; 70: 73-97.
35. ANDERSSON S, GUNDERSEN PM and FINSET A. Emotional activation during therapeutic interaction in traumatic brain injury: effect of apathy, self-awareness and implications for rehabilitation. *Brain Injury*. 1999; 13: 393-404.
36. Santa N, Sugimori H, Kusuda K, Yamashita Y, Ibayashi S and Iida M. Apathy and functional recovery following first-ever stroke. *Int J Rehabil Res*. 2008; 31: 321.
37. Okada K, Kobayashi S, Yamagata S, Takahashi K and Yamaguchi S. Poststroke apathy and regional cerebral blood flow. *Stroke*. 1997; 28: 2437-41.
38. Yamagata S, Yamaguchi S and Kobayashi S. Impaired novelty processing in apathy after subcortical stroke. *Stroke*. 2004; 35: 1935-40.
39. Glodzik-Sobanska L, Slowik A, Kieltyka A, et al. Reduced prefrontal N-acetylaspartate in stroke patients with apathy. *Journal of the Neurological Sciences*. 2005; 238: 19-24.

40. Hama S, Yamashita H, Shigenobu M, et al. Depression or apathy and functional recovery after stroke. *Int J Geriatr Psychiatry*. 2007; 22: 1046-51.
41. Onoda K, Kuroda Y, Yamamoto Y, et al. Post-stroke apathy and hypoperfusion in basal ganglia: SPECT study. *Cerebrovasc Dis*. 2011; 31: 6-11.
42. Cano S, Barrett L, Zajicek J and Hobart J. Beyond the reach of traditional analyses: using Rasch to evaluate the DASH in people with multiple sclerosis. *Multiple Sclerosis Journal*. 2011; 17: 214-22.
43. Cano S, Klassen AF, Scott A, Thoma A, Feeny D and Pusic A. Health outcome and economic measurement in breast cancer surgery: challenges and opportunities. *Expert Review of Pharmacoeconomics & Outcomes Research*. 2010; 10: 583-94.
44. Hobart JC, Cano SJ, Zajicek JP and Thompson AJ. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *The Lancet Neurology*. 2007; 6: 1094-105.
45. Craig AH, Cummings JL, Fairbanks L, et al. Cerebral blood flow correlates of apathy in Alzheimer disease. *Archives of Neurology*. 1996; 53: 1116-20.
46. Benoit M, Koulibaly PM, Migneco O, Darcourt J, Pringuey DJ and Robert PH. Brain perfusion in Alzheimer's disease with and without apathy: a SPECT study with statistical parametric mapping analysis. *Psychiatry Res*. 2002; 114: 103-11.
47. Lanctot KL, Moosa S, Herrmann N, et al. A SPECT study of apathy in Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2007; 24: 65-72.
48. Starkstein S, Jorge R, Mizrahi R and Robinson R. A prospective longitudinal study of apathy in Alzheimer's disease. *Journal of Neurology, Neurosurgery & Psychiatry*. 2006; 77: 8.
49. Starkstein SE, Petracca G, Chemerinski E and Kremer J. Syndromic validity of apathy in Alzheimer's disease. *American Journal of Psychiatry*. 2001; 158: 872.
50. Benoit M, Clairet S, Koulibaly P, Darcourt J and Robert P. Brain perfusion correlates of the Apathy Inventory dimensions of Alzheimer's disease. *Int J Geriatr Psychiatry*. 2004; 19: 864-9.

51. Starkstein SE, Ingram L, Garau M and Mizrahi R. On the overlap between apathy and depression in dementia. *Journal of Neurology, Neurosurgery & Psychiatry*. 2005; 76: 1070.
52. Isella V, Melzi P, Grimaldi M, et al. Clinical, neuropsychological, and morphometric correlates of apathy in Parkinson's disease. *Movement Disorders*. 2002; 17: 366-71.
53. Pluck GC and Brown RG. Apathy in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*. 2002; 73: 636-42.
54. Reijnders JSAM, Scholtissen B, Weber WEJ, Aalten P, Verhey FRJ and Leentjens AFG. Neuroanatomical correlates of apathy in Parkinson's disease: A magnetic resonance imaging study using voxel based morphometry. *Movement Disorders*. 2010.
55. Dujardin K, Sockeel P, Devos D, et al. Characteristics of apathy in Parkinson's disease. *Movement Disorders*. 2007; 22: 778-84.
56. Pedersen KF, Larsen JP, Alves G and Aarsland D. Prevalence and clinical correlates of apathy in Parkinson's disease: a community-based study. *Parkinsonism Relat Disord*. 2009; 15: 295-9.
57. Drijgers RL, Dujardin K, Reijnders JSAM, Defebvre L and Leentjens AFG. Validation of diagnostic criteria for apathy in Parkinson's disease. *Parkinsonism Relat Disord*. 2010; 16: 656-60.
58. Kant R, Duffy J and Pivovarnik A. Prevalence of apathy following head injury. *Brain Injury*. 1998; 12: 87-92.
59. Piamarta F, Iurlaro S, Isella V, et al. Unconventional affective symptoms and executive functions after stroke in the elderly. *Archives of Gerontology and Geriatrics*. 2004; 38: 315-23.
60. Brodaty H, Sachdev PS, Withhall A, Altendorf A, Valenzuela MJ and Lorentz L. Frequency and clinical, neuropsychological and neuroimaging correlates of apathy following stroke - The Sydney Stroke Study. *Psychological Medicine*. 2005; 35: 1707-16.

61. Mayo NE, Fellows LK, Scott SC, Cameron J and Wood-Dauphinee S. A Longitudinal View of Apathy and Its Impact After Stroke. *Stroke*. 2009; 40: 3299-307.
62. Tang W-K, Lau CG, Mok V, Ungvari GS and Wong K-S. Apathy and Health-Related Quality of Life in Stroke. *Arch Phys Med Rehab*. 2013.
63. Shulman LM. Apathy in patients with Parkinson's disease. *International Review of Psychiatry*. 2000; 12: 298-306.
64. Ghaffar O and Feinstein A. Mood, affect and motivation in rehabilitation. *Cognitive neurorehabilitation: Evidence and Application*. 2nd ed. New York, NY: Cambrigde University Press; US, 2008, p. 205-17.
65. Hama S, Yamashita H, Shigenobu M, et al. Depression or apathy and functional recovery after stroke. *Int J Geriatr Psychiatry*. 2007; 22: 1046-51.
66. van Reekum R, Stuss DT and Ostrander L. Apathy: why care? *J Neuropsychiatry Clin Neurosci*. 2005; 17: 7-19.
67. Hama S, Yamashita H, Kato T, et al. 'Insistence on recovery' as a positive prognostic factor in Japanese stroke patients. *Psychiatry and Clinical Neurosciences*. 2008; 62: 386-95.
68. Brodaty H, Sachdev PS, Withall A, Altendorf A, Valenzuela MJ and Lorentz L. Frequency and clinical, neuropsychological and neuroimaging correlates of apathy following stroke - The Sydney Stroke Study. *Psychological Medicine*. 2005; 35: 1707-16.
69. Hama S, Yamashita H, Shigenobu M, et al. Post-stroke affective or apathetic depression and lesion location: left frontal lobe and bilateral basal ganglia. *Eur Arch Psychiatry Clin Neurosci*. 2007; 257: 149-52.
70. Santa N, Sugimori H, Kusuda K, Yamashita Y, Ibayashi S and Iida M. Apathy and functional recovery following first-ever stroke. *Int J Rehabil Res*. 2008; 31: 321-6.
71. Geen RG. *Human motivation: A social psychological approach*. Brooks/Cole Pub. Co., 1995.
72. Jorge RE, Starkstein SE and Robinson RG. Apathy following stroke. *Canadian Journal of Psychiatry*. 2010; 55: 350-4.

73. Marsh L and Berk A. Neuropsychiatric aspects of Parkinson's disease: recent advances. *Current Psychiatry Reports*. 2003; 5: 68-76.
74. Organization WH. Recommendation on stroke prevention, diagnosis and therapy. *Stroke*. 1989; 20: 1407.
75. Skidmore ER, Whyte EM, Holm MB, et al. Cognitive and affective predictors of rehabilitation participation after stroke. *Arch Phys Med Rehabil*. 2010; 91: 203-7.
76. Starkstein SE, Mizrahi R, Capizzano AA, Acion L, Brockman S and Power BD. Neuroimaging Correlates of Apathy and Depression in Alzheimer's Disease. *Journal of Neuropsychiatry and Clinical Neurosciences*. 2009; 21: 259-65.
77. StataCorp LP. Stata 11. *Base Reference Manual TX: Stata Corporation, University Dr East*. 2009.
78. Higgins J and Thompson SG. Quantifying heterogeneity in a meta analysis. *Statistics in medicine*. 2002; 21: 1539-58.
79. Higgins JPT and Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1. 0. The Cochrane Collaboration, 2011. 2011.
80. Egger M, Smith GD and Phillips AN. Meta-analysis: principles and procedures. *Bmj*. 1997; 315: 1533.
81. Maher CG, Sherrington C, Herbert RD, Moseley AM and Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Physical Therapy*. 2003; 83: 713.
82. Tate L, Mcdonald S, Perdices M, Togher L, Schultz R and Savage S. Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychol*. 2008; 18: 385-401.
83. Moher D, Liberati A, Tetzlaff J and Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*. 2009; 6: e1000097.
84. Czernecki V, Pillon B, Houeto J, Pochon J, Levy R and Dubois B. Motivation, reward, and Parkinson's disease: Influences of dopatherapy. *Neuropsychologia*. 2002; 40: 2257-67.

85. Politis AM, Vozzella S, Mayer LS, Onyike CU, Baker AS and Lyketsos CG. A randomized, controlled, clinical trial of activity therapy for apathy in patients with dementia residing in long-term care. *Int J Geriatr Psychiatry*. 2004; 19: 1087-94.
86. Herrmann N, Rothenburg LS, Black SE, et al. Methylphenidate for the treatment of apathy in Alzheimer disease: prediction of response using dextroamphetamine challenge. *J Clin Psychopharmacol*. 2008; 28: 296-301.
87. Bonfanti AB, Etcheverry JL, Persi GG, Zezza H, Starkstein S and Gatto EM. Apathy in parkinsons disease impairment in quality of life. [Spanish]
Apatia en enfermedad de parkinson impacto sobre la calidad de vida. *Medicina*. 2009; 69: 253-8.
88. Niu Y-X, Tan J-P, Guan J-Q, Zhang Z-Q and Wang L-N. Cognitive stimulation therapy in the treatment of neuropsychiatric symptoms in Alzheimer's disease: a randomized controlled trial. *Clin Rehabil*. 2010; 24: 1102-11.
89. Whyte EM, Lenze EJ, Butters M, et al. An open-label pilot study of acetylcholinesterase inhibitors to promote functional recovery in elderly cognitively impaired stroke patients. *Cerebrovascular Diseases*. 2008; 26: 317-21.
90. Newburn G and Newburn D. Selegiline in the management of apathy following traumatic brain injury. *Brain Inj*. 2005; 19: 149-54.
91. Skidmore ER, Holm MB, Whyte EM, Dew MA, Dawson D and Becker JT. The feasibility of meta-cognitive strategy training in acute inpatient stroke rehabilitation: Case report. *Neuropsychol*. 2011; 21: 208-23.
92. Lane-Brown A and Tate R. Evaluation of an intervention for apathy after traumatic brain injury: a multiple-baseline, single-case experimental design. *J Head Trauma Rehabil*. 2010; 25: 459-69.
93. Robinson RG, Jorge RE, Clarence-Smith K and Starkstein S. Double-blind treatment of apathy in patients with poststroke depression using nefiracetam. *Journal of Neuropsychiatry and Clinical Neurosciences*. 2009; 21: 144-51.
94. Newburn G and Newburn D. Selegiline in the management of apathy following traumatic brain injury. *Brain Injury*. 2005; 19: 149-54.

95. Rothstein HR, Sutton AJ and Borenstein M. Publication Bias in Meta-Analysis. *Publication Bias in Meta-Analysis*. 2005: 1-7.
96. Cerniauskaite M, Quintas R, Boldt C, et al. Systematic literature review on ICF from 2001 to 2009: its use, implementation and operationalisation. *Disability and rehabilitation*. 2011; 33: 281-309.
97. Jette AM. Toward a common language for function, disability, and health. *Physical Therapy*. 2006; 86: 726-34.
98. Cieza A, Brockow T, Ewert T, et al. Linking health-status measurements to the international classification of functioning, disability and health. *Journal of Rehabilitation Medicine*. 2002; 34: 205-10.
99. Stucki G, Cieza A and Melvin J. The international classification of functioning, disability and health: A unifying model for the conceptual description of the rehabilitation strategy. *Journal of Rehabilitation Medicine*. 2007; 39: 279-85.
100. Cieza A and Stucki G. Content comparison of health-related quality of life (HRQOL) instruments based on the international classification of functioning, disability and health (ICF). *Quality of Life Research*. 2005; 14: 1225-37.
101. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B and Stucki G. ICF linking rules: an update based on lessons learned. *J Rehabil Med*. 2005; 37: 212-8.
102. Weigl M, Cieza A, Harder M, et al. Linking osteoarthritis-specific health-status measures to the International Classification of Functioning, Disability, and Health (ICF). *Osteoarthritis and Cartilage*. 2003; 11: 519-23.
103. Geyh S, Cieza A, Kollerits B, Grimby G and Stucki G. Content comparison of health-related quality of life measures used in stroke based on the international classification of functioning, disability and health (ICF): a systematic review. *Quality of Life Research*. 2007; 16: 833-51.
104. Geyh S, Kurt T, Brockow T, et al. Identifying the concepts contained in outcome measures of clinical trials on stroke using the International Classification of Functioning, Disability and Health as a reference. *Journal of Rehabilitation Medicine*. 2004; 36: 56-62.

105. Sigl T, Cieza A, Brockow T, Chatterji S, Kostanjsek N and Stucki G. Content comparison of low back pain-specific measures based on the International Classification of Functioning, Disability and Health (ICF). *The Clinical journal of pain*. 2006; 22: 147-53.
106. Moriello C, Byrne K, Cieza A, Nash C, Stolee P and Mayo N. Mapping the stroke impact scale (SIS-16) to the International Classification of Functioning, Disability and Health. *Journal of Rehabilitation Medicine*. 2008; 40: 102-6.
107. Mayo NE, Poissant L, Ahmed S, et al. Incorporating the International Classification of Functioning, Disability, and Health (ICF) into an electronic health record to create indicators of function: proof of concept using the SF-12. *J Am Med Inform Assoc*. 2004; 11: 514-22.
108. Tessier A, Mayo NE and Cieza A. Content identification of the IWQOL-Lite with the International Classification of Functioning, Disability and Health. *QualLife Res*. 2011; 20: 467-77.
109. Kuspinar A and Mayo NE. Do generic utility measures capture what is important to the quality of life of people with multiple sclerosis? *Health QualLife Outcomes*. 2013; 11: 71.
110. Mayo NE, Moriello C, Asano M, van der Spuy S and Finch L. The extent to which common health-related quality of life indices capture constructs beyond symptoms and function. *QualLife Res*. 2011; 20: 621-7.
111. Barbic SP, Bartlett SJ and Mayo NE. Emotional vitality: concept of importance for rehabilitation. *Arch Phys Med Rehabil*. 2013; 94: 1547-54.
112. Wade DT. Outcome measures for clinical rehabilitation trials: impairment, function, quality of life, or value? *Am J Phys Med Rehabil*. 2003; 82: S26-S31.
113. Noonan VK, Kopec JA, Noreau L, et al. Comparing the content of participation instruments using the International Classification of Functioning, Disability and Health. *Health Qual Life Outcomes*. 2009; 7: 93.
114. Teixeira-Salmela LF, Neto MG, Magalhães LC, Lima RC and Faria CD. Content comparisons of stroke-specific quality of life based upon the international classification of functioning, disability, and health. *Quality of Life Research*. 2009; 18: 765-73.

115. Alghwiri AA, Marchetti GF and Whitney SL. Content comparison of self-report measures used in vestibular rehabilitation based on the international classification of functioning, disability and health. *Physical Therapy*. 2011; 91: 346-57.
116. Prieto L, Alonso J and Lamarca R. Classical test theory versus Rasch analysis for quality of life questionnaire reduction. *Health and Quality of Life Outcomes*. 2003; 1: 27.
117. Nijsten T, Unaeze J and Stern R. Refinement and reduction of the Impact of Psoriasis Questionnaire: classical test theory vs. Rasch analysis. *British Journal of Dermatology*. 2006; 154: 692-700.
118. Hobart J and Cano S. *Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods*. Prepress Projects, 2009.
119. Andrich D. Rating scales and Rasch measurement. *Expert review of pharmacoeconomics & outcomes research*. 2011; 11: 571-85.
120. Rasch G. On general laws and the meaning of measurement in psychology. *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*. University of California Press Berkeley, CA, 1961, p. 321-33.
121. Siegert RJ, Tennant A and Turner-Stokes L. Rasch analysis of the Beck Depression Inventory-II in a neurological rehabilitation sample. *Disabil Rehabil*. 2010; 32: 8-17.
122. Hawthorne G, Densley K, Pallant JF, Mortimer D and Segal L. Deriving utility scores from the SF-36 health instrument using Rasch analysis. *Quality of Life Research*. 2008; 17: 1183-93.
123. Bateman A, Teasdale TW and Willmes K. Assessing construct validity of the self-rating version of the European Brain Injury Questionnaire (EBIQ) using Rasch analysis. *Neuropsychol*. 2009; 19: 941-54.
124. Duncan PW, Bode RK, Min Lai S and Perera S. Rasch analysis of a new stroke-specific outcome scale: the stroke impact scale1. *Arch Phys Med Rehab*. 2003; 84: 950-63.

125. Antonucci G, Aprile T and Paolucci S. Rasch analysis of the Rivermead Mobility Index: a study using mobility measures of first-stroke inpatients. *Arch Phys Med Rehab*. 2002; 83: 1442-9.
126. Itzkovich M, Tripolski M, Zeilig G, et al. Rasch analysis of the Catz-Itzkovich spinal cord independence measure. *Spinal Cord*. 2002; 40: 396-407.
127. Avery LM, Russell DJ, Raina PS, Walter SD and Rosenbaum PL. Rasch analysis of the Gross Motor Function Measure: validating the assumptions of the Rasch model to create an interval-level measure. *Arch Phys Med Rehab*. 2003; 84: 697-705.
128. Lourenco C. Apathy in degenerative and acquired disorder. 2011.
129. van Reekum R, Stuss DT and Ostrander L. Apathy: Why care? *Journal of Neuropsychiatry and Clinical Neurosciences*. 2005; 17: 7-19.
130. Hama S, Yamashita H, Kato T, et al. 'Insistence on recovery' as a positive prognostic factor in Japanese stroke patients. *Psychiatry Clin Neurosci*. 2008; 62: 386-95.
131. Mayo NE, Fellows LK, Scott SC, Cameron J and Wood-Dauphinee S. A longitudinal view of apathy and its impact after stroke. *Stroke*. 2009; 40: 3299-307.
132. Maclean N, Pound P, Wolfe C and Rudd A. Qualitative analysis of stroke patients' motivation for rehabilitation. *Bmj*. 2000; 321: 1051-4.
133. Rasch G. *Probabilistic models for some intelligence and attainment tests*. ERIC, 1993.
134. Tang WK, Wong E, Chiu HF, Lum C and Ungvari GS. The Geriatric Depression Scale should be shortened: results of Rasch analysis. *Int J Geriatr Psychiatry*. 2005; 20: 783-9.
135. Andrich D, Sheridan B and Luo G. RUMM 2030 (Beta Version for Windows) Perth. *Western Australia: RUMM Laboratory Pty Ltd*. 2009.
136. Kreiner S. Validity and objectivity: Reflections on the role and nature of Rasch models. *Nordic Psychology*. 2007; 59: 268.
137. Kreiner S and Christensen KB. Graphical rasch models. *Statistical methods for quality of life studies*. 2002: 187-203.

138. Tennant A and Conaghan PG. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care & Research*. 2007; 57: 1358-62.
139. Fisher W. The cash value of reliability. *Rasch Measurement Transactions*. 2008; 22: 1160-3.
140. Siegel S. Nonparametric statistics. *The American Statistician*. 1957; 11: 13-9.
141. Psychometric Laboratory for Health Sciences. Introductory Rasch Analysis using RUMM2030. University of Leeds, 2012, p. 29-34.
142. Cella D, Lai Js, Chang CH, Peterman A and Slavin M. Fatigue in cancer patients compared with fatigue in the general United States population. *Cancer*. 2002; 94: 528-38.
143. Teresi JA and Fleishman JA. Differential item functioning and health assessment. *Quality of Life Research*. 2007; 16: 33-42.
144. Norman S, Davis E, Goldstone LW, Villa L and Warholak TL. A Rasch analysis of a questionnaire designed to evaluate psychiatric patient attitudes and knowledge after attending a pharmacist-led patient medication education group. *Research in Social and Administrative Pharmacy*. 2014.
145. Mayo NE, Bronstein D, Scott SC, Finch LE and Miller S. Necessary and sufficient causes of participation post-stroke: practical and philosophical perspectives. *Quality of Life Research*. 2014; 23: 39-47.
146. Allison PD. Missing data: Quantitative applications in the social sciences. *British Journal of Mathematical and Statistical Psychology*. 2002; 55: 193-6.
147. Duncan PW, Wallace D, Lai SM, Johnson D, Embretson S and Laster LJ. The Stroke Impact Scale Version 2.0 evaluation of reliability, validity, and sensitivity to change. *Stroke*. 1999; 30: 2131-40.
148. Hays RD, Sherbourne CD and Mazel RM. The rand 36-item health survey 1.0. *Health economics*. 1993; 2: 217-27.
149. Cinamon JS, Finch L, Miller S, Higgins J and Mayo N. Preliminary evidence for the development of a stroke specific geriatric depression scale. *Int J Geriatr Psychiatry*. 2011; 26: 188-98.

150. Yesavage JA, Brink T, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of psychiatric research*. 1983; 17: 37-49.
151. Munro BH. *Statistical methods for health care research*. Lippincott Williams & Wilkins, 2005.
152. O'connor BP. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior research methods, instruments, & computers*. 2000; 32: 396-402.
153. Ghaffar O, Feinstein A, Stuss IDT, Winocur G and Robertson IH. Mood, affect and motivation in rehabilitation. *Cognitive neurorehabilitation: Evidence and application*. 2008; 2: 205-17.
154. Marin RS and Wilkosz PA. Disorders of diminished motivation. *J Head Trauma Rehabil*. 2005; 20: 377-88.
155. Pallant JF and Tennant A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*. 2007; 46: 1-18.
156. Shea TL, Tennant A and Pallant JF. Rasch model analysis of the Depression, Anxiety and Stress Scales (DASS). *BMC psychiatry*. 2009; 9: 21.
157. Riley WT, Pilkonis P and Cella D. Application of the National Institutes of Health Patient-Reported Outcomes Measurement Information System (PROMIS®) to Mental Health Research. *The journal of mental health policy and economics*. 2011; 14: 201.
158. Food U. Drug Administration: Guidance for industry: Patient-reported outcome measures—Use in medical product development to support labeling claims. *Federal Register*. 2009; 74: 65132-3.
159. Snijders TA. *Multilevel analysis*. Springer, 2011.
160. Nagin D. *Group-based modeling of development*. Harvard University Press, 2005.
161. Xie H, Mayo N and Koski L. Predictors of future cognitive decline in persons with mild cognitive impairment. *DementGeriatrCogn Disord*. 2011; 32: 308-17.

162. Xie H, Mayo N and Koski L. Identifying and characterizing trajectories of cognitive change in older persons with mild cognitive impairment. *DementGeriatrCogn Disord*. 2011; 31: 165-72.
163. Ahmed S, Mayo N, Scott S, Kuspinar A and Schwartz C. Using latent trajectory analysis of residuals to detect response shift in general health among patients with multiple sclerosis. [corrected]. *QualLife Res*. 2011; 20: 1555-60.
164. Bartlett SJ, Ling SM, Mayo NE, Scott SC and Bingham CO, III. Identifying common trajectories of joint space narrowing over two years in knee osteoarthritis. *Arthritis Care Res(Hoboken)*. 2011; 63: 1722-8.
165. Mayo NE, Bronstein D, Scott SC, Finch LE and Miller S. Necessary and sufficient causes of participation post-stroke: practical and philosophical perspectives. *QualLife Res*. 2014; 23: 39-47.
166. Soicher JE, Mayo NE, Gauvin L, et al. Trajectories of endurance activity following pulmonary rehabilitation in COPD patients. *Eur Respir J*. 2012; 39: 272-8.
167. Caeiro L, Ferro JM and Costa J. Apathy secondary to stroke: a systematic review and meta-analysis. *Cerebrovasc Dis*. 2013; 35: 23-39.
168. Nagin DS and Tremblay RE. Analyzing developmental trajectories of distinct but related behaviors: a group-based method. *Psychological methods*. 2001; 6: 18.
169. Fellows L, Heberlein A, Morales D, Shivde G, Waller S and Wu D. Method matters: an empirical study of impact in cognitive neuroscience. *Cognitive Neuroscience, Journal of*. 2005; 17: 850-8.
170. Rorden C and Karnath HO. Using human brain lesions to infer function: a relic from a past era in the fMRI age? *Nature Reviews Neuroscience*. 2004; 5: 812-9.
171. Rorden C, Karnath HO and Bonilha L. Improving lesion-symptom mapping. *J Cogn Neurosci*. 2007; 19: 1081-8.
172. Rorden C, Fridriksson J and Karnath HO. An evaluation of traditional and novel tools for lesion behavior mapping. *Neuroimage*. 2009; 44: 1355-62.
173. Bates E, Wilson SM, Saygin AP, et al. Voxel-based lesion–symptom mapping. *nature neuroscience*. 2003; 6: 448-50.

174. Williams A. What bothers caregivers of stroke victims? *J Neurosci Nurs*. 1994; 26: 155-61.
175. Robert PH, Darcourt G, Koulibaly MP, et al. Lack of initiative and interest in Alzheimer's disease: a single photon emission computed tomography study. *Eur J Neurol*. 2006; 13: 729-35.
176. Marshall GA, Fairbanks LA, Tekin S, Vinters HV and Cummings JL. Neuropathologic correlates of apathy in Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2006; 21: 144-7.
177. Apostolova LG, Akopyan GG, Partiali N, et al. Structural correlates of apathy in Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2007; 24: 91-7.
178. Lourenço CB, Azeff L, Sveistrup H and Levin MF. Effect of environment on motivation and sense of presence in healthy subjects performing reaching tasks. *Virtual Rehabilitation, 2008*. IEEE, 2008, p. 93-8.
179. Sansone C, Sachau DA and Weir C. Effects of instruction on intrinsic interest: The importance of context. *Journal of Personality and Social Psychology*. 1989; 57: 819.